16. THE COSMOLOGICAL FRAMEWORK

Index | Questions | References | PDF

Under Current Construction : last update Jan 12 2013

Next | Top

(1) Introduction

(a) The Good News

Cosmology is in a truly golden era: now is a great time to learn it!

- Some recent **observational developments** have been:
  - the ¾-century long uncertainty in H_o is now over: $H_o = 72 \pm 5$ km s^{-1} Mpc^{-1}
  - the long search for a decent standard candle is also over: SNIa seem to work well
  - COBE & WMAP have now measured the CMB spectrum & anisotropy with high accuracy
  - 2dFGRS & SDSS have recently mapped cosmologically significant local structures
  - observations at z ~1-6 are almost routine, and probe a significant slice of cosmic history

- In close partnership, **theoretical developments** and their applications are increasingly robust:
  - The cosmic **constituents** and their **relative proportions** have been ascertained
    $\Omega_\Lambda = 0.73 \pm 0.02$: dark energy dominates the current universe
    $\Omega_{DM} = 0.26 \pm 0.02$: dark matter is very important, particularly in structure formation
    $\Omega_b = 0.04 \pm 0.01$: baryons are a trace (though vital!) component
    $\Omega_\gamma = 5.0 \times 10^{-5}$: CMB photons reveal an early hot phase
    $\Omega_\nu = 3.4 \times 10^{-5}$: CNB neutrinos have been observed indirectly in the CMB power spectrum.
  - An "ordinary" FRW cosmological **world model** has emerged as being completely adequate
    The FRW parameters have been measured → **"The Concordance Model"**
    This yields a reliable framework for charting cosmic history -- ie we now know t(z)
    A number of puzzles are removed by invoking an early period of inflation.
  - A detailed theory exists tracing the **average** conditions from very early times
    a fairly full description of t < 1s exists, though it is not yet well tested
    nucleosynthesis at t ~ 1-5 mins nicely recovers the observed light element abundances
    (in fact, this measures conditions at t ~ 1s, when the neutron/proton ratio was fixed)
    the theory accurately describes recombination at ~½Myr and the origin of the CMB
  - A detailed theory now exists which describes the **growth of perturbations**
    Starting from inflation, a natural **spectrum of fluctations** can be followed to z~1000
    here it matches the CMB anisotropies in great detail

it can then be followed to z=0, where it accurately matches **local structure**

- After thousands of years of wild speculation, the true story of creation is finally emerging
  → we are living through (and participating in!) a historic period of intellectual growth
  → in the future, our time will be recalled much like that of Copernicus, Newton, or Darwin.

(b) The Not-So-Good News

- Before we get too dizzy with euphoria, let's regain some humility by recalling:
  - we have **no idea** what the dark matter or dark energy are actually made of (ie 96% !)
  - although inflation is a promising idea, it is **far from proven** and its cause is unknown
  - the origin of the baryon asymmetry is only guessed at
  - why particular cosmological values are favoured is unknown beyond anthropic arguments.
  - why there is **something** and not **nothing** is as unknown now as it was in Aristotle's day.
- Of course, these (and many other) puzzles are not really bad news at all:
  they signify a rich subject in good health, with more fascinating work to be done.
  In any case, it would be unseemly to fully understand Creation after only a couple of generations.

(c) Our Path Through the Subject

- Understandably, "The New Cosmology" has attracted enormous interest and effort
  → The subject is now mature and sophisticated -- much is well beyond our/my range
  → Our aim will be to **outline the overall framework**, while **ignoring details**
- Here is our path through the subject:
  - Following tradition, we begin with a discussion of some **global cosmic properties**:
    isotropy, homogeneity, expansion, structure, composition
  - Introduce the five cosmic constituents: Baryons; Photons; Neutrinos; Dark Matter; Dark Energy
    Properties: densities; pressures; cooling; evolution.
  - We need a **General Relativistic framework**, comprising two inter-dependent parts:
    geometry: metrics define the spacetime
    dynamics: contents define coefficients of the metric and hence its time evolution
  - We clarify some confusing aspects of geometry on an **expanding coordinate grid**:
    various types of distance and horizon
  - We introduce a **toolkit** for measuring intrinsic properties at cosmological distances.
    these allow us to convert **redshift** (our prime observable) to distance; time; etc.
  - Such relations lead to the **classic tests** of the world model parameters.
  - As you may know, several deep **conundrums/puzzles** emerge from the standard picture:
    the so-called flatness; horizon; structure; anti-matter; problems
  - These are "explained" by introducing an early period of **rapid inflation**.
    we note how inflation solves some of these problems, and how it might be caused
  - We briefly outline the evolution of the **hot early universe**:
    the various "eras" : quark; hadron; lepton; radiation; matter; dark energy
    a brief treatment of cosmic nucleosynthesis, and recombination
  - We see recombination **directly** as the Cosmic Microwave Background (CMB)
    its primary characteristics are its remarkable isotropy and purity of black body form
    it provides the strongest evidence for the hot big bang
    it concludes our review of **homogeneous** Cosmology.
- Following homogeneous Cosmology, we are ready to start discussing **inhomogeneities**
  These provide the starting point for our next Topic: **structure formation**.

Next Prev Top

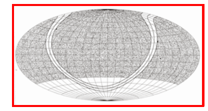
(2) Global Properties

(a) Large Scale Isotropy

- To humans, the universe seems **highly anisotropic**

down is solid, up is the sky, with the sun, moon, & stars in specific directions
even statistically, stars prefer the milky way while bright galaxies cluster in Virgo and Coma

- Only at **much fainter levels** and **much greater distances** does isotropy begin to emerge
  Here are some nice examples: [images]
  2 million faint ($m_J < 17$) galaxies cover ~1 sr with only slight structure visible
  31,000 radio sources (typical z ~ 1) uniformly cover the northern hemisphere
  the CMB with the galaxy & dipole removed is isotropic to one part in 10^5



- At faint levels (i.e. large scales) the Universe seems **remarkably isotropic**

(b) The Cosmological Principle

- Ever since Copernicus, we are loath to assign Earth a special location
  Going one step further, we promote egalitarianism to a **"Cosmological Principle"**
  → **The Universe looks, statistically, the same from all locations**
  go to **any** galaxy -- you will witness an isotropic universe.

- One consequence of this is that the Universe **has no edge or center**
  for a Euclidean space this also means the universe is **infinite**
  Note, this need **not** be true for other, closed, geometries (see later).

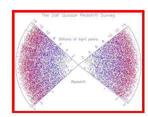
- There is also a "**Strong** Cosmological Principle" = Cosmological Principle + **"at all times"**
  → in addition to uniformity, it explicitly states the universe **does not evolve**
  this underpins the **Steady State Cosmology** advocated by Hoyle, Bondi & Gold in the 50s & 60s
  The evidence **for** evolution first came in the 1960s and has grown stronger ever since [song].
  → we won't consider the Steady State cosmology further.

(c) Large Scale Homogeneity

- It is easy to show that: isotropy + cosmological principle = **homogeneity**
  → all locations are (statistically) equivalent (e.g. have the same mean density)

- This is sometimes extended to postulate that the laws of physics are also global
  The observations of familiar spectral features in distant galaxies certainly supports this

- Homogeneity only sets in on **large enough scales**, somewhere between 100 Mpc & 1 Gpc [images]
  On smaller scales, of course, one encounters much **inhomogeneity** (see below)

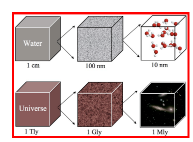


(d) Small Scale Structure

- The Universe's small scale **inhomogeneity** is much more obvious to us than its global **homogeneity**.
  The inhomogeneity appears as a **hierarchy of structures**: stars; galaxies; clusters; tapestry

- There is also a **maximum scale** (~150 Mpc) for inhomogeneity before homogeneity sets in [images]

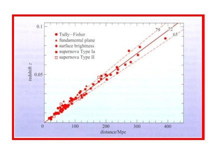
  Our cosmology must explain the **origin, development, and maximum scale of all this structure**.



(e) Expansion

(i) The Hubble Law

- As soon as galaxy spectra were measured it became clear that most were **redshifted**
  In 1929 Hubble found a "roughly linear" relation between redshift and distance: $cz \propto d$
  as data improved, this relation was confirmed and has strengthened ever since [image]



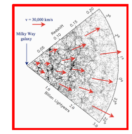
- Don't confuse establishing the linear relation with measuring its gradient:
  It took ~75 years to achieve $\lesssim$10% uncertainty in the gradient, H_o (see below)
  This is primarily because calibrating the distance scale is notoriously difficult.
  The current best estimate for H_o is 72 ± 5 km/s/Mpc (72×10^{-6} Myr^{-1} in psm units)
  Note, it is still customary to quote measurements scaled to h = 100 km/s/Mpc
  For example, "The luminosity of M87 is $2.3 \times 10^{11}\ h^{-2}\ L_\odot$" [Topic 1.3k]

- Note also that there is always some small scatter (~few 100 km/s) in the Hubble relation
  → galaxy redshifts have **two** components:
  - the global "Hubble Flow" (arising from cosmic expansion)
  - a small "peculiar" velocity (arising as the galaxy responds to the gravity of its neighbors).

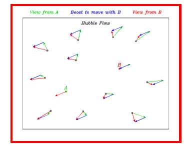
  ["peculiar" comes from Latin: "peculium", meaning "private property" - specific to it alone]

(ii) Everyone Sees the Same Law

- At first glance the Hubble Law seems to violate the Cosmological Principle:
  → galaxies appear to move radially away from **us** suggesting we are somehow central [image]
  However, its linear form ensures that **all** locations witness the same relation:



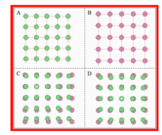
  Consider two (vector) locations **k** and **p** and the vector field: **v** = H **r** centered on us (at O)
  We see p move at $\mathbf{v_p} = H\,\mathbf{p}$; so how does an observer located on **k** see p move?
  use primes to denote values measured by k [image]:



$$\mathbf{p'} = \mathbf{p} - \mathbf{k} \quad \text{and} \quad \mathbf{v'_p} = \mathbf{v_p} - \mathbf{v_k} = H\mathbf{p} - H\mathbf{k} = H(\mathbf{p} - \mathbf{k}) = H\mathbf{p'}$$

  Hence, for any point **p**, k sees a Hubble Law: $\mathbf{v'_p} = H\,\mathbf{p'}$
  → if we see v = Hr then so does everyone else!
  → the cosmological principle still holds true

- Since everyone witnesses the same Hubble Law, we conclude that:
  → the **Universe itself** is undergoing **isotropic expansion** with form **v** = H **r** [image]



  This is a remarkable and profound result.

(f) The Big Bang

- The linear nature of the Hubble law has a second fascinating consequence:
  Tracing the expansion backwards, there was a moment when **all galaxies were together**

  This singular event is called "**The Big Bang**"

  The term was coined in 1952 by Hoyle to gently mock opponents of his Steady State Cosmology.

- If no forces act to accelerate or decelerate the expansion, then v(t) = const,
  In that case, the singular event ocurred at a time $t_{H,o} = r/v = 1/H_o$ in the past

  $t_{H,o} = H_o^{-1} = 9.76\ h^{-1}$ Gyr is called **The Current Hubble Time**

  If the average expansion rate isn't too different from its current value, then $t_{H,o} \approx$ current age.
  For $H_o = 72$ km/s/Mpc, $t_{H,o} = 13.5$ Gyr → a ballpark figure for the age of the Universe.

(g) Cosmic Time

- Being physics graduates, we are often nervous when time is introduced:
  does it depend on our location or motion, i.e. is it different for different observers?

  In this case, you can relax: the time we've introduced is **a proper time**
  → it is measured by inertial observers, and can be agreed upon by everyone.

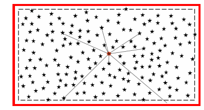
- In fact, the cosmological principle, coupled with global expansion, guarantees a universal time:
  → fundamental observers may synchronize clocks when the local density reaches some value
  → since these observers are at rest w.r.t. their local frame, they measure **proper time**

- If we start the clocks at the Big Bang (ie set t_{BB} to 0 at $\rho = \infty$), we have **cosmic time**:
  → we can all agree on the age of the universe, and the time/age at a given redshift: t(z).

(h) Olbers' Paradox

- Olbers (~1820) is credited with first noticing that a dark night sky has cosmological implications.

Imagine an **infinitely** large, old, static, homogeneous universe. [image]
Every sight line intersects a star's surface, whose surface brightness is independent of its distance
→ hence the entire sky should shine with the surface brightness of stars!
However it doesn't, so one or more of the assumptions must be **wrong**.



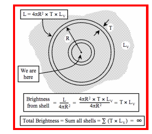
(Absorption doesn't help: dust ultimately heats to reach equilibrium with the radiation field).

- Let's use modern numbers to quickly assess how this paradox might constrain cosmology.
  The mean local cosmic luminosity density is $\rho_L \sim 10^8\ L_{B,\odot}\ \mathrm{Mpc}^{-3}$ with $M/L \sim 10$ [Topic 4.3]
  → $\rho_L \sim 10^{-32}$ erg s^{-1} cm^{-3}; $n_\star \sim 10^8$ Mpc^{-3}; $\rho_\star \sim 10^{-31}$ gm cm^{-3} $(\equiv \Omega_\star \sim 0.01)$; $\Sigma_\star \sim R_\odot^2$

- There are several ways to cast these numbers, all confirming the Universe's emptiness & darkness:

  - mfp = $1 / n_\star \Sigma_\star \sim 10^{18}$ Mpc : a typical sight line terminates at an **enormous** distance.

  - the sky brightness, J, out to distance D is $\int \rho_L\, r^2\, d\Omega\, dr / 4\pi r^2 = \rho_L D/4\pi$ erg s^{-1} cm^{-2} sr^{-1} [image]
    for D_{CMB} (now) ~ 40 Glyr we find $J \sim 10^{-4.5}$ erg s^{-1} cm^{-2} sr^{-1} $\equiv$ $\mu \sim 24^m$ arcsec^{-2} ie very dark!
    to get $J \sim 10^{10}$ erg s^{-1} cm^{-2} sr^{-1} (5000K black body) we need $D \sim 10^{18}$ Mpc (as before).
    Clearly, given its current emptiness, a dark sky only rules out **gigantic** old static Universes.



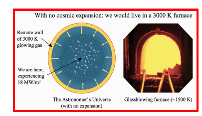
  - A 5000K radiation field has energy density $u \sim 30$ erg cm^3. Can stars generate this?
    At $\rho_L \sim 10^{-32}$ erg s^{-1} cm^{-3} it would take them 10^{26} years !

    A bright sky requires our static universe to be both gigantic **and immensely old**.

  - But real stars cannot shine that long anyway, nor can they create so much light:
    Converting **all baryons** to light: $u = 0.04\ \rho_{crit}\ c^2 \sim 10^{-8}$ erg cm^{-3}, well below what is needed
    Stated differently: in a bright sky universe the radiation alone would contribute $\Omega_{rad} \sim 10^9$!

    Stars alone **cannot** yield Olbers' bright sky, even in an old enough and big enough Universe.

- There is, of course, **another** source of radiation: the 3000K photosphere of the CMB.
  The omnidirectional CMB is uncannily like the classic Olbers' bright sky, except for one detail....
  Redshift kills its surface brightness by a factor $(1 + z)^4 \sim 10^{12}$ [sec 8d]
  In the absense of redshift, we would indeed witness a lethal sky with $J \sim 10^6$ Watt m^{-2} sr^{-1}
  (roughly equivalent, of course, to filling the sky with solar disks, as Olbers' originally conceived)



  For this light source, then, it is **expansion** which keeps the sky dark.

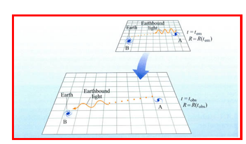
Next Prev Top

(3) More on Expansion

(a) Expanding Coordinate Grids

- When thinking about cosmic dynamics, it helps to think of an **expanding coordinate grid**
  Imagine a (flat) sheet of graph paper which is slowly getting bigger, carrying the grid lines with it [image].

- One can imagine **three** kinds of motion:



  - Points (ink dots) **in the paper** participate in its motion. They are fixed on the grid.
    Such points are called **fundamental observers**: their motion **is** the Hubble flow.
    As already shown in [sec 2eii] they all witness the same Hubble law.
    In introductory texts, this is the expanding raisin bread analogy.

  - Points (ants, if you like) can also "walk" **over** the paper, **crossing** grid lines.
    This motion is the **peculiar velocity**: it is a true space velocity.

  - Now imagine a speedy ant which always runs quickly at a constant rate in a straight line
    this represents a photon, crossing the expanding space, always moving locally at c.

- From General Relativity, we know that in Cosmology we **may** need to consider a **curved** grid. For this, one imagines our dots and ants on an **expanding sphere** with longitude/latitude lines. Fortunately, it turns out that expanding **flat** graph paper is all that's needed for our Universe.

(b) The Scale Factor a(t) & Comoving Coordinates

How do we treat this kind of coordinate system mathematically?
Easy, we consider the grid and its expansion separately.

- First notice that the **Hubble law preserves shapes**: patterns of galaxies becomes bigger patterns
  → for a set of i points, cosmic expansion gives: $\mathbf{r}_i(t) = a(t)\,\mathbf{r}_i(t_o)$, for a fiducial time t_o
  Here, a(t) is a universal scalar function of time, and is called **the scale factor**
  a(t) simply tracks how the separation of galaxies changes over time, starting at a=0 at the big bang.

  → **Finding the form for a(t) is a holy grail in cosmology.**

- Sensibly, we assign the **current** time special status: t_o = now; $a(t_o) = 1$ and $\mathbf{r}(t_o) = \mathbf{r_o}$.
  Hence:
  - the current values, $\mathbf{r_o}$, provide the coordinate grid, and are called **comoving coordinates**
    as the grid expands, the comoving coordinates **do not change**
  - at any time, the **physical coordinate** of an object is: $\mathbf{r} = a(t)\,\mathbf{r_o}$
  - by setting $a(t_o) = 1$, we ensure that in the past, $a < 1$ while in the future $a > 1$.

  For example, at the time of recombination, $a \approx 0.001$
  The **comoving distance** to proto-M87 is still 15 Mpc, but its **physical distance** is only 15 kpc.

- Notice that **r** and $\mathbf{r_o}$ are both **proper distances**:
  they tell us how many **non-expanding** rulers fit end-to-end from here to the galaxy.

  Later we introduce several **pseudo-distances**: eg luminosity & angular diameter distance; D_L, D_A.
  these are not true (proper) distances, but convenient **functions** of distance.

- **Warning**: symbol conventions for physical/comoving/pseudo distances varies greatly
  You must be **careful** when reading texts: "is this r or d physical or comoving?"

  → In these notes I will try to be consistent: r = physical; r_o = comoving; D = pseudo.

(c) The Hubble Parameter: H(t)

- With our new expanding coordinate grid, how does H fit in?
  To find out, take time derivatives of $\mathbf{r}(t) = a(t) \times \mathbf{r}(t_o)$:

  $$d\mathbf{r}/dt = \mathbf{v}(t) = da/dt \times \mathbf{r}(t_o) = (1/a)\, da/dt\, \mathbf{r}(t)$$

  But this is simply: $\mathbf{v}(t) = H(t)\,\mathbf{r}(t)$ with $H(t) = (1/a)\, da/dt$
  → we have found that the Hubble relation applies **at all times**

  $$H(a) \equiv H(t) = 1/a\; da/dt \quad \text{and} \quad dr/dt = v = H\,r$$

- In general, H(t) and a(t) both vary with time
  For the current epoch, we write $H(t_o) \equiv H_o$ and it has units of inverse time
  Our best current estimate for H_o is [sec 3g] 2.34×10^{-18} Hz (or 72 km/s/Mpc)

(d) The Velocity-Distance Relation & The Hubble Law

- At this point, we need to clarify something: there are **two** apparently similar relations.
  - the **theoretical "proper" velocity-distance** relation: v = H r
  - the **observational redshift-distance** "Hubble Law": cz = H D
  These are, in fact, somewhat different.

- The theoretical relation v = H r is **globally exact**, though it is observationally inaccessible:

- both v and r are "proper" quantities, ie as measured in a local rest frame. For example
  for r: how many **non-expanding** rulers must be laid down between us and the galaxy?
  after 1 second, v additional rulers must be laid down, where v = H r.
- notice that the values are all measured **at the same cosmic time**:
  we deal with distant galaxies **as they <u>are</u>**, right now, not **as we see them**.

- The Hubble Law is strictly observational:
  - $1 + z = \lambda_{obs}/\lambda_{em}$; and cz rather than v(z) is sometimes taken as a "Doppler velocity"
  - D is usually a luminosity distance, which matches the **proper distance** only at low z.
  - both z and D apply to the **time when the light set out**, not the **current time**
    during the light's journey, the galaxy moved further away and, possibly, slowed down

  At high z, several factors **break linearity**, indeed this deviation is used to measure q_o.

- At low redshift the Hubble Law and the velocity-distance relation **look the same**
  Cosmic expansion is best described by v = Hr; it is exact and holds everywhere at all times
  the Hubble Law, cz = H D, only provides imperfect observational access to this cosmic expansion.

(e) The Hubble Sphere

- Let's push the velocity-distance law to great distances:
  For H_o = 100 km/s/Mpc, we have :

  at r = 10 Mpc, $v = 10^3$ km/s
  at r = 10 Gpc, $v = 10^6$ km/s
  at r = 1000 Gpc, $v = 10^8$ km/s

  Yes, these velocities are **faster than light**: special relativity **does not apply** to this motion:
  → it arises from **expansion <u>of</u> space**, not **motion <u>through</u> space**.

- Can we see the galaxies which recede faster than light?
  The light they emit moves **through space** at speed c towards us
  but over time the wavefronts get **further from us**, at speed v - c > 0 → we will never see them!
  There is a critical distance $r_{H,o} = c/H_o = 3.0\ h^{-1}$ Gpc which is now receeding at $v = H_o\, r_{H,o} = c$

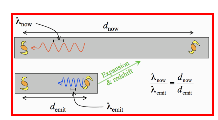
$r_{H,o} = c/H_o$ is called the **Hubble distance**; where, right now, galaxies recede at c

  For constant rate of expansion, we will ultimately see everything inside a sphere of radius $r_{H,o}$
  Only if v slows down significantly will we be able see beyond $r_{H,o}$.
  These, and other potentially confusing things, should become clearer later [sec 7].

(f) The Link Between Redshift and Scale Factor

- When we first encounter redshift, we view it as a **Doppler shift** due to a galaxy's motion
  while this **is** appropriate for peculiar velocities, it is **not** for expansion "velocities"

- A more mature view sees the light embedded in space, getting stretched **on its way to us**
  This kind of redshift is called a **Cosmological Redshift** [image]



- With this view, redshift clearly measures **the stretch factor between emission and observation**

$\lambda_o / \lambda_e = 1 + z = a(t_o) / a(t_e) = 1 / a(t_e)$

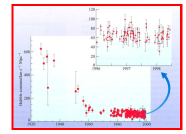
  since the current scale factor $a(t_o) = 1$
  This is a **fundamental relation & globally exact**
  It tells us the relative change in size since the light set out

  Some examples:

  at cz = 30,000 km/s, z = 0.1 → $a(t_e) = 1/1.1 = 0.909 \approx 90\%$
  at z = 1, a = 0.5 all galaxies were half their current separation (8 × the current density)
  at z = 6, a = 0.14 (currently, the most distant QSOs)

at z = 1000, a = 0.1%   (at recombination, everything is 10^3× closer; 10^9× denser)

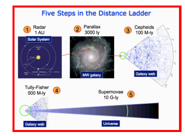
(g) Measuring Distances and H_o



The struggle to measure reliable astronomical distances and H_o has a long history   [image].
Rather than consider that history, we shall **briefly** summarize the current (1995 - 2005) situation.
A good review is Mould's EAA 1997 article (quoted H_o values are taken from this article).
Also Freedman and Madore's ARAA 2010 article.

- To derive H_o one must measure both **distances** and **redshifts**
  - Distance measurements take two rather different forms:
    - ladder methods which depend on nearby calibration such as parallax & Cepheids
    - one shot methods which need no calibration and rely only on known physical properties
  - Redshifts, while easily measured, need correction for peculiar motion:
    - our motion includes: the sun's orbit; MW motion in local group; Virgocentric infall.
    - the target galaxy is chosen to be a group/cluster member, & the group's mean redshift is used.

(i) Ladder Method

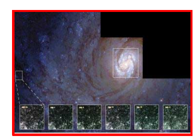


In elementary texts, the distance ladder is often presented with many rungs [image]
In practice, there are really only three:

- Use the Hipparcos satellite to get trigonometric parallaxes of nearby Cepheids
  → calibrate Period-Luminosity (PL) relation
- Use HST to get Cepheid distances to nearby (≲25 Mpc) galaxies
  → calibrate Tully-Fisher (TF) & Fundamental-Plane (FP) (& other) methods
- Use TF, FP (& other) distances to groups where peculiar velocities are unimportant.
  → group mean redshifts & distances now give H_o

Let's outline these various steps & techniques.

- **<u>Cepheid variables</u>**



This class of pulsating stars defines a tight period-luminosity(-color) relation [images]
→ measure **period** to get luminosity and hence **distance**

They are **luminous** stars (M_V: -2 to -6) and hence can be seen to considerable distances (~25 Mpc by HST)
However, they are also **rare**: the nearest is 250pc away, and only 10 have parallax distances measured to 10%.

Historically, the PL relation was calibrated by Main Sequence fitting to open clusters containing Cepheids
Now, Hipparcos provides direct **trigonometric calibration** (eg Perryman et al 1997)
However, this calibration still needs to be improved   (eg using future astrometric mission GAIA).

The distance to the LMC plays a **very important role**   (and also still needs to be improved)
→ it contains enough Cepheids to define the PL relation in m (not M)
→ hence extragalactic Cepheids yield **relative distances** to the LMC
the current best estimate for the LMC is: $m\text{-}M \;=\; 18.44 \pm 0.03 \;\equiv\; 48.7 \pm 0.7$ kpc   (uses E(B-V) = 0.1)

The HST Key Project has now measured 800 Cepheids in ~18 galaxies out to ~25 Mpc.
These galaxies were then used to calibrate the following methods:

- **<u>TF: Tully-Fisher Relation</u>**

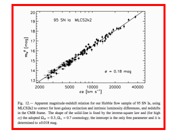
This is a luminosity-linewidth relation for **spirals**   [Topic 5.5b]
scatter is minimum in the near IR (I, H), hence the method is often referred to as "IRTF"
about 20 spirals now have Cepheid distances
about 25 groups/clusters out to 10,000 km/s have TF distances
→ $H_o = 71 \pm 8$ (eg Sakai et al 1999)

- **<u>FP: Fundamental Plane Relation</u>**

This is a refinement of the luminosity-linewidth (Faber-Jackson) relation for **ellipticals** [Topic 7.4b]
Either D_n-Σ (isophotal diameter/dispersion) or surface brightness/radius/dispersion relations
since no Cepheids in Es, calibration uses Es in **groups** with Cepheid distances (eg Virgo, Fornax, Leo)
many groups/clusters out to 10,000 km/s now have FP distances
→ $H_o = 78 \pm 10$ (eg Mould et al 1996; Kelson et al 1999)

- <u>**SNIa: WD binary thermonuclear detonation**</u>

These are very luminous, so well suited to q_o studies (high z), but also useful for H_o (lower z)
the light curves **aren't** all the same; but peak luminosity correlates with fading rate (and color) [image]
unfortunately, very few SNIa have ocurred in galaxies with Cepheid distances → calibration **not** ideal
→ $H_o = 68 \pm 6$ (eg Gibson et al 1999)



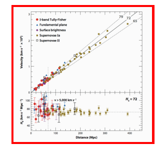
- <u>**SBF: Surface Brightness Fluctuations**</u>

Consider a set of CCD pixels recording the light from an E galaxy, each one with **perfect** S/N ratio
there is still **variation** between the pixels because of √N fluctuations in # stars
Although the **mean** surface brightness is independent of distance, the **variation** is not
→ nearer galaxies have **fewer** stars per pix → **larger** variation.

difficulties: contamination by globular clusters; color/population dependency; calibration.
HST can use this method out to about 7000 km/s
→ $H_o = 69 \pm 7$ (eg Ferrarese et al 1999)

- <u>**HST Key Project**</u> combines all these methods (plus GC & PN luminosity function methods)

→ $H_o = 72 \pm 7$ km s^{-1} Mpc^{-1} (eg Friedman et al 2001) [image]



(ii) Direct Methods

Let's briefly look at four direct distance measurement methods.

- <u>**EPM: Expanding Photospheres Method**</u> (originally from Baade & Wesselink)

This method applies to <u>all</u> pulsating/expanding photospheres -- particularly Type II (core collapse) SN
angular size is derived from flux, temperature and emissivity (black body = 1)
linear size is derived from integrating velocity (linewidth) over time
→ distance, by comparing angular and linear sizes.

- <u>**VLBI Masers in Nuclear Gas Disks**</u>

So far, only one good example of this method exists: NGC 4258, Miyoshi et al 1995 [Topic 14.4e]
a compact (~1pc) molecular disk orbits central black hole
VLBI of H_2O masers gives (Keplerian) velocities and proper motions
→ distance, by comparing linear and angular velocities.

- <u>**Gravitational Lensing Time Delays**</u>

2 QSO images have different light paths with different physical lengths
this path difference is given by the time delay between QSOs light curved (via cross-correlation).
the calculated path difference depends on projected mass density and **linear scale**
→ distance by comparing observed angular scale and calculated linear scale

- <u>**SZ: Sunyaev-Zeldovich Effect in Clusters**</u>

Hot electrons in galaxy cluster ICMs do two things:

  - they generate X-rays via bremsstrahlumg: $L_x \propto n_e^2 \, r_c^3 \, T_x^{1/2}$
  - they Compton scatter CMB photons: $(\Delta T/T)_{CMB} \propto n_e \, r_c \, T_x$

- Overall, these methods give distances that agree, statistically, with the ladder methods.

(4) Components, & Their Densities & Pressures

(a) The Stage and its Contents

- The current universe contains (we think) just five components:
  - Dark Energy
  - Dark Matter
  - Baryonic Matter
  - Photons
  - Neutrinos

- You should not think of these as "merely" components: like furniture sitting in a room
  In general relativity, the contents help to **define** the space-time in which they reside
  (the furniture actually influences the shape and size of the room!)

- To understand the global properties, we therefore need a **complete** itinerary of the components
  - Most important, we need to know their **densities**: ρ
  - We also need to know how densities **change with expansion:** $\rho = \rho_o a^{-x}$
    It transpires (see below) that this is equivalent to knowing a component's **Pressure**

  For several components, ρ and p are both additive: $\rho_{tot} = \Sigma\rho_i$ and $p_{tot} = \Sigma p_i$

(b) Critical Density & Spatial Geometry

- There is an extremely important link between cosmic density and spatial geometry.
  We will need a few equations which we justify later on [sec 6a], so just accept them for now.
  Here's the **cosmic energy (Friedmann) equation**, using scale factor "a" [sec 6a iv]

  $$(da/dt)^2 - (8\pi G/3)a^2\rho = -kc^2/R_0$$

  This matches the basic form KE + PE = E_{tot}, with E_{tot} appearing as a **spatial curvature**:

  k = +1 ≡ -ve E_{tot} ≡ closed geometry
  k = -1 ≡ +ve E_{tot} ≡ open geometry
  k = 0 ≡ zero E_{tot} ≡ flat geometry

- Hence, we can define a **critical density** ρ_c which yields a **flat geometry** (k=0):

  $$(1/a^2)(da/dt)^2 = H^2 = (8\pi G/3)\rho_c \quad \text{giving} \quad \rho_c = 3H^2/(8\pi G)$$

  $$\rho_c = 2.65 \times 10^{-7}\, h^2\, M_\odot pc^{-3} = 1.80 \times 10^{-29}\, h^2\, gm\, cm^{-3} = 10.8\, h^2\, m_p\, m^{-3}$$

- Notice ρ_c **varies with time** through H(t) (or h), and we'll write its current value as $\rho_{c,o}$
  However, once set, k (-1,0,+1) is fixed for all time.
  In particular, if k = 0 now (flat geometry) then we **always** have a flat geometry
  → if $\rho_o = \rho_{c,o}$ now, then $\rho = \rho_c$ always (worth stressing, since this seems to be the case).

- With such tiny values, it is convenient to express densities relative to ρ_c using $\Omega \equiv \rho / \rho_c$

- The table lists the measured **current** relative densities for the various cosmic components
  (values are for the concordence model, ~10% uncertainties)
  (other functions are given which we'll come to in a moment)

Component	$\Omega \equiv \rho_o / \rho_{c,o}$	w	x = 3(1+w) $\rho = \rho_o a^{-x}$	x = 2/(3+3w) $a \propto t^x$	1 + 3w sign accel
Dark Energy	0.73	-1	0	$a \propto e^t$	-2
Dark Matter	0.23	≈0	3	2/3	1
Baryons	0.04	≈0	3	2/3	1

Photons	5.0×10^{-5}	1/3	4	1/2	2
Neutrinos	3.4×10^{-5}	1/3	4	1/2	2

- The current best estimate for the **total** density is $\Omega_{tot} = \Sigma\, \Omega_i = 1.00 \pm 0.02$

  → **we live in a universe with "flat" spatial geometry**

  This is one of the most important discoveries in recent cosmology.

- **Warning**: please don't continue to view Ω_{tot} as defining the **future of the Universe**, it doesn't.
  this **was** appropriate in the pre-lambda days ($\Lambda = 0$), when it was common to state:
  if $\Omega_{tot} > 1$, the Universe will turn around, collapse, and end in a big crunch.
  if $\Omega_{tot} < 1$, the Universe will expand forever.
  with $\Lambda \neq 0$, one **cannot** infer the future simply from Ω_{tot}.
  Instead, Ω_{tot} only fixes **the spatial geometry** (open/flat/closed), not the future.

(c) Pressure & Equation of State Parameter, w

- In modern cosmology there is much talk of "w" the "equation of state" parameter.
  It specifies a (simple) relation between pressure and density: $p = w\,\rho\, c^2$.
  → pressure is a fraction w of a region's total energy density.
  We are told "pressures add to gravity" in GR, so we must consider them.

  I feel this is an unnecessarily obtuse approach.
  Here is a more straightforward approach (which amounts to the same thing).

- Focus instead on how cosmic densities **change with expansion**: $\rho(a) = \rho_o\, a^{-x}$
  For (simple, cold) matter, x = 3 : density $\propto$ 1 / volume (i.e. simple conservation of matter)

  But in cosmology, $x \neq 3$ is also possible.
  In such cases energy is either leaving or entering our expanding box.
  One can consider this as "work done" by a pressure.

- Let's use local conservation of energy to express this: $dQ = dU + p\, dV = 0$ so $dU = -p\, dV$
  or using $U = \rho\, c^2\, V$ and $V \propto a^3$ we have:

  $$d(\rho\, c^2\, a^3) = -p\, d(a^3)$$

  $$d\rho\, c^2\, a^3 + \rho c^2\, 3a^2\, da = -p\, 3a^2\, da \qquad \text{now divide by } 3a^2c^2\, da$$

  $$a/3\; d\rho/da + \rho = -p/c^2$$

  Now use $\rho = \rho_o\, a^{-x}$ so that $d\rho/da = -x\, \rho_o\, a^{-x} / a = -x\, \rho / a$, we get:

  $$-x\rho/3a + \rho = -p/c^2$$

  giving an expression for the pressure, and a relation between x and w:

  $$p = (x/3 - 1)\, \rho c^2 \equiv w\rho c^2$$

  Since $x = 3(1 + w)$ we recover our density change, but now in terms of w:

  $$\rho = \rho_o\, a^{-x} = \rho_o\, a^{-3(1+w)}$$

- The table above gives the values of w and x for the various components.
  Let's make sure we understand them:

  - For a **non-relativistc gas**: $p = nkT$ and $^3/_2 kT = ½\, m\langle v^2\rangle$
    so $p = (\rho_o/m)\,(m \langle v^2\rangle/3) = \rho_o c^2 \times \langle v^2\rangle/3c^2$
    → $w = \langle v^2\rangle/3c^2 \approx 0$ so that x = 3 giving $\rho_m = \rho_{m,o}\, a^{-3}$

Basically, the thermal energy is << rest energy
so cooling during expansion doesn't contribute to the change in ρ

- For a **photon gas**: $p = {}^1/_3\, u_r = {}^1/_3\, \rho_r c^2$
  $\rightarrow$ $w = {}^1/_3$ so that $x = 4$ giving $\rho_r = \rho_{r,o}\, a^{-4}$

  Basically, when a photon gas expands its pressure does work on the surroundings
  Energy is lost from the expanding box because the photons are redshifted.

- For a **vacuum energy**, its density doesn't change with expansion, so we have:
  $dU = -p\, dV$ $\rightarrow$ $\rho_v c^2\, dV = -p\, dV$ $\rightarrow$ $p = -\rho_v c^2$
  $\rightarrow$ $w = -1$ and we recover $x = 0$, giving $\rho_v = \rho_{v,o}\, a^0 = \rho_{v,o}$

  Basically, you must **provide** energy to create more vacuum
  Negative pressure is tension: imagine a strange piston with a little "strange water" in it.
  You pull **extremely hard**, with force $F = p \times A$, with $p = 9 \times 10^{20}$ dyne cm^{-2} (~10^{15} atm) and $A = 1$ cm^2
  the piston slowly moves out by d = 1 cm -- you have spent $F \times d = 9 \times 10^{20}$ erg of energy.
  To your surprise, the piston now contains an **additional cm^3 of water!**
  Your 9×10^{20} erg were converted to $9 \times 10^{20}/c^2 = 1$ gm of new water.
  (Note: since dark energy is 6.8×10^{-30} gm cm^{-3}, it only requires 6×10^{-9} dyne cm^{-2} tension to create.
  However, you can't verify this experimentally because there is vacuum on both sides of the piston!).

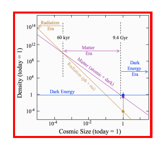
- You may ask: Where does the energy go to (photons) come from (vacuum)?
  The answer is rather subtle:
  In a Newtonian framework: it goes to (comes from) the gravitational binding energy of the Universe.
  In a GR framework: it appears as modifications to the geometry terms in $G_{\mu\nu}$
  We'll return to this later.

(d) Density Changes with Scale Factor

- Summarizing what we just found:

matter:	$\rho_m(a) = \rho_{m,o}\, a^{-3} = \rho_{m,o}\, (1+z)^3$	as expected by "conservation of mass"
radiation:	$\rho_r(a) = \rho_{r,o}\, a^{-4} = \rho_{r,o}\, (1+z)^4$	since $n_\gamma \propto a^{-3}$ and $E_\gamma \propto a^{-1}$ from redshift
vacuum:	$\rho_v(a) = \rho_{v,o} =$ const	space is space

So the **relative** densities of the three components **changes with expansion**: [image]
Depending on which component dominates the density (hence gravity), we find three different eras:



The Radiation Era;
The Matter Era;
The Dark Energy Era.

- Quantitatively, we can identify times when the various component densities are **equal**:
  (a_{eq} and z_{eq} are easily evaluated, while t_{eq} requires the integral relations from [sec 6c v])

density match	condition	a @ equality	z @ equality	t @ equality
$\rho_v = \rho_m$	$0.73 = 0.27\, a^{-3}$	0.72	0.39	9.43 Gyr
$\rho_v = \rho_{rel}$	$0.73 = 8.4 \times 10^{-5}\, a^{-4}$	0.103	8.3	615 Myr
$\rho_b = \rho_\gamma$	$0.04\, a^{-3} = 5.0 \times 10^{-5}\, a^{-4}$	1.25×10^{-3}	800	620 kyr
$\rho_m = \rho_{rel}$	$0.27\, a^{-3} = 8.4 \times 10^{-5}\, a^{-4}$	3.11×10^{-4}	3200	57 kyr
$\rho_m = \rho_\gamma$	$0.27\, a^{-3} = 5.0 \times 10^{-5}\, a^{-4}$	1.85×10^{-4}	5400	22 kyr

Note that here ρ_{rel} refers to the sum of photons and neutrinos (relativistic matter)
Likewise ρ_m refers to the sum of baryons and CDM (non-relativistic matter)

(e) Rates of Expansion In Each Era

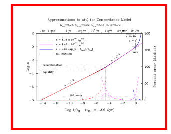
- In general **time** dependencies are complicated [sec 6c].
  However, for **flat geometries** with a **single component**, we find simple forms for a(t)
  From the cosmic energy equation, with k = 0, $\rho_o = \rho_{c,o}$ and $\rho = \rho_o\, a^{-3(1+w)}$, we have

$$(da/dt)^2 = (8\pi G/3)a^2\rho = (8\pi G/3)a^2\rho_o\, a^{-3(1+w)} = H_o^{\,2}\, \rho_o/\rho_{c,o}\, a^{-(1+3w)} = H_o^{\,2}\, a^{-(1+3w)}$$

$$da/dt = H_o\, a^{-(1+3w)/2} \rightarrow \int a^{(1+3w)/2}\, da = H_o \int dt \rightarrow a = [(3+3w)/2\ .\ t/t_H]^{2/(3+3w)}$$

- Now, as it happens the universe **is** flat, and there are times when just one component dominates:

during the radiation era:	$a \propto t^{1/2}$
during the matter era:	$a \propto t^{2/3}$
during the vacuum era:	$a \propto e^t$ (from $da/dt \propto a$)



- The function a(t) from zero to the present indeed matches these simple relations [see image and sec 6c iii]

(f) The Five Components

Let's now look briefly at the 5 cosmic components

(i) Dark Energy

- This component has had a checkered history:
  - Famously introduced in 1917 by Einstein to balance matter & yield a static universe
  - Dumped in 1929 when Hubble discovered cosmic expansion
  - Resurrected in 1930s-40s when large H_o implies $t_H < t_\oplus$ (loitering model)
  - Dumped in 1960s-70s when H_o revised down, giving $t_H > t_\oplus$
  - Resurrected in 1980s when $t_H < t_{GC}$; & $\Omega_M \sim 0.3$ but inflation suggests $\Omega_{tot} = 1$
  - Discovered in 1998 with accelerating expansion: $q_0 = ½\Omega_M - \Omega_v < 0$
  - Here to stay. Now (2011) 4 independent & consistent confirmations.

- Dark Energy is generic term; there are three main possibilities:
  A "cosmological constant" appearing in Einstein's equations as an integration constant, Λ.
  Quantum mechanical energy of the vacuum. Both vacuum energy and Λ have w=-1
  Quintessence: a quantum field with -1 < w < -1/3, evolving slowly over time.
  We dont yet know which of these dark energy is:
  Acceleration only demands w < -1/3, while current observations constrain w < -0.9
  → vacuum energy is currently favoured (which can be treated as a Λ term)

- Current theoretical understanding is ~zero
  a "natural" energy density has $\rho_v c^2 = m^4/(hc)^3$ for "natural mass" m
  if m = Planck mass = $(hc/G)^{½}$ → $\rho_v c^2 \sim 3 \times 10^{126}$ eV cm^{-3} (~10^{93} gm cm^{-3} !!)
  Observed $\rho_v c^2 \sim 10^3$ eV cm^{-3}, or ~10^{123} × smaller; maybe the worst guess ever.
  The particle mass which **does** give the correct value is m ~ 10^{-3} eV. No known candidates.
  Studying DE in the lab is very difficult - only energy **differences** can be measured
  (The Casimir effect, for example, shows the vacuum is "live" but not by how much).

(ii) Dark Matter

- Introduced by Zwicky in 1930s to explain high galaxy cluster dispersions. Then forgotten.
  Re-introduced in the 1970s to explain spiral rotation curves at large radii
  Increasingly important on larger scales: galaxies; binaries; clusters; large scale flows.
  Necessary to create large scale structure rapidly from CMB (baryons alone are inadequate).
  Combined methods yield $\rho_{DM} = 0.23 \pm 0.04$

- Theoretical understanding ~zero, though some properties are constrained.
  - Non-baryonic (ie neither protons nor neutrons); BBNS rules out high ρ_b
  - Non-relativistic after kT ~ 1 MeV → **cold** dark matter (CDM) → clusters efficiently

- Interacts only via weak and gravitational forces (**dark** matter)
- Possibly made of "WIMPS" -- Weakly Interacting Massive Particles
- WIMP candidates? not known; maybe lightest supersymmetric particle (LSP).

- Current laboratory searches underway. Nothing yet found.

(iii) Baryons

- In this context, baryons refers to neutrons, protons, **and electrons**
  Big Bang NucleoSynthesis (BBNS) and CMB power spectra together give: $\Omega_b = 0.044 \pm 0.005$

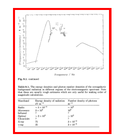
- Primordial H/He abundances are 75/25 by **mass** or 12/1 by **number**
  Since then, stars have converted ~2% (by mass) into heavier elements.

- Only 10% of baryons are **luminous**, i.e. are found in stars
  The other 90% is mosly in diffuse ionized gas surounding galaxies.

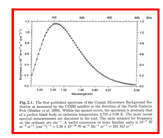
- Mass to light ratios help inter-convert: $\rho_m = j_X \times (M/L)_X$ in band X [Topic 1.3j]
  For the local universe, $j_B \approx 2.0 \times 10^8$ h $L_{\odot,B}$ Mpc^{-3} ($\approx$ 9 Watt AU^{-3})
  This yields several fiducial values:
  using ρ_{crit} we have $(M/L)_{crit} = 1400$ h $(M/L)_{\odot,B}$ $\rightarrow$ 7 million kg/Watt (very dark!)
  using $\rho_m = 0.27\ \rho_{crit}$ we have $(M/L)_m = 375$ $\rightarrow$ close to some clusters
  using $\rho_b = 0.04\ \rho_{crit}$ we have $(M/L)_b = 55$ $\rightarrow$ significantly more than most galaxies
  The universe is **optically** quite dark $\rightarrow$ on average, ~7000 tonnes/Watt.

(iv) Photons

- The CMB is by far the strongest of all the cosmic background radiation fields [images]
  This is true for both number density, $n_\gamma \sim 410\ cm^{-3}$, and energy density, $u_\gamma \sim 0.26$ eV cm^{-3}.
  It's energy density is even comparable to the photon and thermal components of the local ISM.



- The CMB spectrum spans 500μ to 1 cm, and peaks at 1.0 mm (B_λ) or 1.9 mm (B_ν)
  It is an exceedingly accurate black body, with $T_{CMB} = 2.725$ ($\pm$ 0.002 K = 0.07%) [image]



- Summarizing the integrated energy and number densities for a BB radiation field, we have:
  [note: here $a = 8\pi^5 k^4/15h^3c^3 = 4\ \sigma/c$ is the radiation constant and **NOT** the scale factor!].

Energy density	$u = a T^4 = 7.56 \times 10^{-15}\ T^4$ erg cm^{-3} $u_{cmb} = 4.17 \times 10^{-13}$ erg cm^{-3} = 0.26 eV cm^{-3}
Energy flux	$J = uc/4\pi = caT^4/4\pi = \sigma T^4/\pi = 1.80 \times 10^{-5}\ T^4$ erg s^{-1} cm^{-2} sr^{-1} $J_{cmb} = 9.94 \times 10^{-4}$ erg s^{-1} cm^{-2} sr^{-1}
Number density	$n = a T^3/2.7k_B = 20.3\ T^3\ cm^{-3}$ $n_{cmb} = 410\ cm^{-3}$
Number flux	$N = nc/4\pi = 4.84 \times 10^{10}\ cm^{-2}\ s^{-1}\ sr^{-1}$ $N_{cmb} = 9.78 \times 10^{11}\ cm^{-2}\ s^{-1}\ sr^{-1}$

(v) Neutrinos

- Like all particles, neutrinos were created out of energy in the early Universe.
  Since they are stable, they survive to today in great, though undetectable, numbers.
  There are three families, ν_e ν_μ ν_τ , each with particle/anti-particle pairs $\rightarrow$ six in all.

- They decoupled at t ~ 1s, z ~ 10^{10}, kT ~ 1 MeV, and have travelled unscattered ever since
  $\rightarrow$ **Cosmic Neutrino Background** (CNB) similar to, but much younger than, the CMB

- Since kT >> $m_\nu c^2$ ($\approx$ 0?) at decoupling, they were relativistic $\rightarrow$ **"hot" dark matter**.
  They behave like a relativistic gas (γ = 4/3; w = 1/3) similar to the CMB.

- Two facts make the CNB slightly different from the CMB
  - Shortly after ν decoupling the e^+e^- pairs annihilate (at kT ~ 0.5 MeV)

this puts energy & entropy into the CMB but not the CNB $\rightarrow$ $T_\gamma > T_\nu$
one can show [sec 12] that $T_\gamma = (11/4)^{1/3}\, T_\nu$ giving $T_\nu = 1.94$ K

- Neutrinos obey Fermi-Dirac statistics, whereas photons obey Bose-Einstein statistics. Summing over all six species, one can show [sec 12]:
  $u_{\nu,tot} = 0.68\, u_\gamma$
  $n_{\nu,tot} = 9/11\, n_\gamma = 335\ \text{cm}^{-3}$ (currently)

▪ The **total** energy density in γ's and ν's (after e^+e^- annihilation) is: $u_{rel} = 1.68\, u_\gamma = 1.68\, a\, T_\gamma^4$
The total relativistic contribution to Ω is: $\Omega_{rel} = \Omega_\gamma + \Omega_\nu = 8.4 \times 10^{-5}$ (today)
This is the value to use when evaluating the expansion rate during the radiation era.

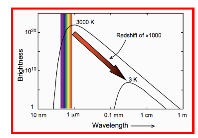
▪ What if neutrinos don't have zero rest mass?
Currently, $kT_\nu = 1.67 \times 10^{-4}$ eV, so only neutrino masses **larger** than this change things
For each species, one finds a current $\Omega_\nu = m_\nu(\text{eV}) / 94\, h^{-2}$
Hence, a mean mass of 8 eV per species will give $\Omega_{\nu tot} = 1$ and close the Universe.
Laboratory limits are presently well above this, though structure formation suggests Ω_ν is small.
(note: neutrino oscillation experiments only give mass **differences** ~ 0.05eV, not masses)

(g) Cosmic Cooling

Consistent with your intuition, cosmic contents cool with cosmic expansion
Ultimately, this occurs because particles move into receding regions so seem to have lower energy
There are two situations to consider: photons & matter (strictly: relativistic & non-relativistic).

(i) Photon (& Neutrino) Cooling

▪ The relativistic content includes both photons and neutrinos: both have a black body spectrum
Expansion preserves the black body shape but at a new temperature: $T_z = T_o(1+z) = T_o / a$
Let's quickly show that this is the case: [image]



- Currently, we have $a = 1$; $T_o = 2.725$K; & BB photon number density at frequency ν_o:
  $n(\nu_o)\, d\nu_o = (8\pi/c^3)\, \nu_o^2\, [\exp(h\nu_o/kT_o) - 1]^{-1}\, d\nu_o$

- At a time when $a \neq 1$, the corresponding frequency is $\nu = \nu_o / a$, giving $d\nu = d\nu_o / a$
  conserving photon numbers gives their density change: $n(\nu)\, d\nu = n(\nu_o) / a^3\, d\nu_o$
  Inserting these into the current BB spectrum, we get:
  $a^3\, n(\nu)\, d\nu = (8\pi/c^3)\, \nu^2\, a^2 [\exp(h\nu a/kT_o) - 1]^{-1}\, a\, d\nu$, giving
  $n(\nu)\, d\nu = (8\pi/c^3)\, \nu^2\, [\exp(h\nu/kT_o a^{-1}) - 1]^{-1}\, d\nu$

- This is another black body spectrum, with:

$$T = T_o / a = T_o\,(1 + z)$$

▪ This has been verified using absorption in the CI fine structure lines in damped Lyα systems:
e.g. at $z = 1.8$, the CI line ratios imply $T_{CMB} \sim 7.6$ K, as expected.

(ii) Matter (& Galaxy) Cooling

▪ A similar analysis can be applied to the cooling of non-relativistic matter particles
There are two key differences:
- The distribution is Maxwell-Boltzmann, in which $T \propto \langle v^2 \rangle$, while for BB: $T \propto \langle \nu \rangle$
- The number density is not set by the temperature, but enters as a separate variable.

▪ First let's derive the matter equivalent of the redshift relation $\nu = \nu_o / a$, ie $v = v_o / a$
In time dt, a peculiar velocity v traverses a distance $r = v\, dt$
at this new location, the peculiar velocity is reduced below v by the Hubble flow:
  $dv = -Hr = -H\, v\, dt = (-1/a)(da/dt)\, v\, dt = -(v/a)da$ $\rightarrow$ $dv/v = -da/a$ with solution:
  $v \propto a^{-1}$ or, for $v = v_o$ at $a = 1$, we have $v = v_o/a$
$\rightarrow$ as time passes a increases & v decreases; hence random motions **decrease** and T **drops**

▪ A gas of particles in thermal equilibrium has a Maxwell-Boltzmann (MB) velocity distribution:

$n(v_o)dv_o = 4\pi N_o\, v_o^2\, (m/2\pi kT_o)^{3/2} \exp(-mv_o^2 / kT_o)\, dv_o$ (N and n in number per cm^3)

at time when $a \neq 1$, we have $v = v_o/a$ and $dv = dv_o/a$
particle conservation also requires $n(v)dv = n(v_o)/a^3\, dv_o$ and $N = N_o/a^3$
Substituting these into the MB relation, we get:

$$a^3\, n(v)dv = 4\pi\, a^3 N\, a^2 v^2\, (m/2\pi kT_o)^{3/2} \exp(-mv^2 a^2 / kT_o)\, a\, dv$$
$$n(v)dv = 4\pi\, N\, v^2\, (m/2\pi k(T_o/a^2))^{3/2} \exp(-mv^2 / k(T_o/a^2))\, a\, dv$$

which is another MB distribution with temperature:

$$T = T_o / a^2 = T_o\, (1 + z)^2$$

- Thus, a gas undergoing passive cosmic expansion cools according to $T \propto a^{-2}$
which is **faster** than the CMB photons for which $T \propto a^{-1}$
→ At z ~ 1100 both have T ≈ 3000K; at z = 0, $T_\gamma \approx 2.725$ K, while $T_{matter} \approx 2.48$ mK.

  In practice, this cooling for the baryonic gas never occurs:

  - prior to recombination, the gas is coupled to the radiation and follows its temparature
  - after recombination other processes (virialization shocks, stars, AGN) heat the gas.

- However, one **can** consider the peculiar motion of **galaxies** in a similar manner
Cosmic expansion is continually "cooling" their peculiar velocities, while
gravitational interaction is continually heating them (e.g. cluster virial motions).

Next Prev Top

(5) Cosmic Geometry: Robertson-Walker Metric

With all these preliminaries now in place, we are ready to begin a full description of the Universe.

- **General Relativity** provides a framework for describing both geometry & dynamics
Einstein's equations of GR can be written in a deceptively compact form: $G_{\mu\nu} = -8\pi G/c^2\, T_{\mu\nu}$

- $G_{\mu\nu}$ & $T_{\mu\nu}$ are both 4 × 4 matricies, suggesting 16 equations, though symmetries reduce this to 10.
The indices $\mu\,\nu$ = 0,1,2,3 denote 1 time & 3 space coordinates (e.g. ct, x, y, z or ct, r, θ, ϕ).
G describes the **space-time geometry**; it incorporates the metric.
T describes the distribution of **mass-energy** ($\mu\,\nu$ = 0), and **momentum** ($\mu\,\nu$ = 1,2,3) in the system
Hence the famous saying: "Mass tells space how to curve, while space tells mass how to move".

- The derivation and solution of the GR equations is a huge subject unto itself.
Here we will extract only what we need.

  Loosely speaking, this section looks at the geometry part, G.
  The next section looks at the dynamical part, T, and how it relates to G.

(a) Space-Time Geometries

- Mankind's attempt to formally describe Nature began with **geometry**: the properties of **space**
Not surprisingly, Euclid's description arose out of our local "flat" space, characterised by:
  parallell lines never meet
  interior angles of a triangle sum to 180°
  a circle has circumference $2\pi r$ and area πr^2; a sphere has area $4\pi r^2$ and volume $4/3\, \pi r^3$
This is the geometry we first encounter in grade school, with rulers and flat graph paper.

  Adding Newton's intuitively plausible independent **time**, t, yields a 4 coordinate **space-time**

- Since we **evolved** within a "flat" space-time, Euclid's/Newton's descriptions feel deeply **TRUE**
However, they are not.
Like cultural norms, their self-evident correctness merely reflects our parochial experience.

Measure with great accuracy, visit a black hole, or traverse the Universe, and things are different.

- How do we describe other possibilities? Using **metrics**
  A metric defines the separation, ds, of two nearby points (or **events** if we include time)

(i) Example 1: "Flat" Geometry

- We call the Euclid space "Flat" because in 2-D it holds true on a flat surface.
  For 2-D and 3-D, we have metrics:

$$ds^2 = dx^2 + dy^2 = dr^2 + r^2\, d\theta^2 \quad \text{where } ds^2 \equiv (ds)^2, \text{etc.}$$

$$ds^2 = dx^2 + dy^2 + dz^2 = dr^2 + r^2\,(d\theta^2 + \sin^2\theta\, d\phi^2) = dr^2 + r^2\, d\psi^2$$

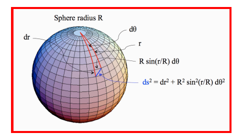
Note that the choice of cartesian or polar (or any other) coordinate system is unimportant
They are equivalent and define the **same** space.

- With these metrics, you can recover **all** the results of Euclidian geometry
  E.g. find the shortest lines between three points (minimize Σ ds) to construct a triangle
  $\rightarrow$ its interior angles sum to 180^o; and so on.

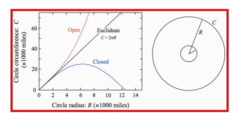
(ii) Example 2: Uniform Positive Curvature

- In 2-D, curved surfaces are everywhere (wineglasses, chairs, lampshades,....)
  Drawing triangles on them quickly reveals interior angle sums **different** from 180^o
  In general, sums **greater/smaller** than 180^o define **positive/negatively** curved surfaces [table 1].
  Such surfaces have metrics which are **different** from those above.

- Anticipating an isotropic universe, consider the simplest isotropic surface: a **sphere** of radius R.
  Take coordinates r, θ from a pole, where r is measured **along the surface** [image]
  $ds^2 = dr^2 + R^2 \sin^2(r/R)\, d\theta^2$
  Close to the pole, (r << R) we recover the "flat" 2-D metric ($ds^2 = dr^2 + r^2 d\theta^2$)
  Further from the pole, curvature reduces the $d\theta^2$ contribution to ds^2.



- If we draw a circle of radius r, the circumference is $\int$ds with r = const, $\theta = 0$ to $2\pi = 2\pi$ R sin(r/R).



$\rightarrow$ circumference = 2π R sin(r/R)

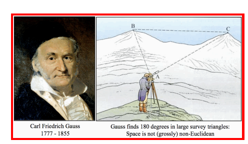
This is < 2πr, goes through a **maximum** at r = πR/2 (equator) and **goes to zero** at r = πR (antipode) [image]

- Now consider extending this metric to **3-D** giving an (un-imaginable) **hypersphere** of "radius" R
  $ds^2 = dr^2 + R^2 \sin^2(r/R)\, d\psi^2$ where $d\psi^2 = d\theta^2 + \sin^2\theta\, d\phi^2$, r is a "straight" line from origin.

To get a feel for the odd nature of this space, imagine holding a laser pointer with visible beam (r)
Turn it through 1^o ($d\psi = 1^o$):

at r = 100m it sweeps sideways by 100 × (1/57.3) = 1.74m (ds at fixed r = R sin(r/R)dψ)
at r = n × 100m it should sweep n × 1.74m, right? Wrong!
as n increases, the sweep drops below the predicted linear relation (ie R sin r/R < r)
in fact, beyond r = R × π/2 the sweep begins to **decrease** for larger r
waving the laser wildly, its rays **always** pass through the remote location r = πR.
A friend located at this "antipode" would see a laser "star" zooming around the sky!

If you explored geometry living in this kind of space, you would **not** recover the Euclidean results
In 1840 Gauss actually tried to measure the curvature of space by surveying big triangles [image].
Of course, his modestly accurate measurements only recovered the Euclidean value of 180^o
But in principle he could have discovered the non-Euclidean terrestrial Schwarzschild metric.
(in fact, R ~ 1AU, and a 30 km triangle deviates from 180^o by ~10^{-8} arcsec).



Carl Friedrich Gauss 1777 - 1855
Gauss finds 180 degrees in large survey triangles: Space is not (grossly) non-Euclidean

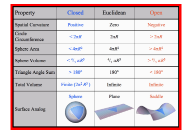
- Like the 2-D analog, the spherical 3-D space also has finite extent (π R) and volume ($2\,\pi^2 R^3$)

(iii) Example 3: Uniform Negative Curvature

- Replacing sin(r/R) by sinh(r/R) in the above metrics yields spaces of uniform **negative** curvature.
  Unlike the 2-D sphere, it isn't posible to construct a 2-D surface with this property,

however, a saddle does have a region of isotropic negative curvature in its center.
On this surface, circles have circumferences 2π R sinh(r/R) which is **less than** 2πr
Furthermore, in 3-D: two points can be infinitely far apart, and the total volume is infinite.

- This table summarizes the geometrical properties of the three spaces described above [image] :



Curvature	k	$d\psi^2$ coefficient	Parallel Lines	Triangle Angles	Circle Circumf	Sphere Area	Sphere Volume	Global Form
Positive	+1	$R^2 \sin^2(r/R)$	Converge	> 180	< $2\pi r$	< $4\pi r^2$	< $(4/3)\pi r^2$	Closed
Flat	0	r^2	Never meet	180	$2\pi r$	$4\pi r^2$	$(4/3)\pi r^2$	Open
Negative	-1	$R^2 \sinh^2(r/R)$	Diverge	< 180	> $2\pi r$	> $4\pi r^2$	> $(4/3)\pi r^2$	Open

(iv) Adding Time

- It is straightforward to add an independent, Newtonian, time t, as additional coordinate.
The 3-D + 1-t metrics for positive, flat, and negative space-times become:

$$ds^2 = -c^2\,dt^2 + dr^2 + R^2 \sin^2(r/R)\,(d\theta^2 + \sin^2\theta\, d\phi^2)$$

$$ds^2 = -c^2\,dt^2 + dr^2 + r^2\,(d\theta^2 + \sin^2\theta\, d\phi^2) \quad \text{(Minkowski space-time).}$$

$$ds^2 = -c^2\,dt^2 + dr^2 + R^2 \sinh^2(r/R)\,(d\theta^2 + \sin^2\theta\, d\phi^2)$$

One can also replace $(d\theta^2 + \sin^2\theta\, d\phi^2)$ with $d\psi^2$, the angle between the two events on the sky, and group them into a single expression, using $S_k(x)$ = sin(x), x, sinh(x) for k = +1, 0, -1:

$$ds^2 = -c^2\,dt^2 + dr^2 + R^2\, S_k^2(r/R)\ \ d\psi^2$$

[Note: I've adopted Peacock's notation for S_k, rather than Ryden's]

- Notice that $ds^2 \equiv -c^2 d\tau^2$, where $d\tau$ is the **Lorentz invariant proper time interval**.
Indeed, the second is the spacetime of special relativity, and is called **Minkowski spacetime**.

- Notice also that the flat metric involves **sums of quadratics** of coordinate differentials
→ flat geometry is rooted in the Pythagoraean relation (e.g. triangles have Σ ang = 180°).
Although more general geometries are curved, many are **locally flat** (e.g. in 3-D: a sphere)
→ expanding the metric to first order at any location gives a quadratic form.
Such geometries are called **Reimannian** and the spacetimes of GR are **all of this kind**. Why?
Because **locally** the Equivalence Principle demands a Minkowski spacetime of Special Relativity.

As you can see, the three metrics above are all of this kind: as $r \to 0$, they are locally flat.
Of course, the **second** derivatives do not vanish, and it is these that define the curvature.

(b) The Robertson-Walker Metric

- In the 1930s, Robertson and Walker (independently) considered metrics suitable for the Universe.
The required metrics need to describe a space-time which:
  - is always isotropic & homogeneous
  - expands (or contracts) with time

  These two conditions limit the possibilities **enormously**

- Note that in 1930s, the assumption of homogeneity & isotropy was fairly bold
But the algebraic advantages were so great it was the obvious/only option to pursue.
The degree to which these assumptions are valid **depends on scale**:

On **large scales** (> few Mpc): these assumptions are **excellent** [see sec 2a-c]

the RW-metric works well, and the treatment is valid.

On **intermediate scales**, where $|\Delta\rho/\langle\rho\rangle| < 1$ they are still **useful**:

the RW-metric can be applied **locally** with parameters slightly different from the global ones
this approach is integral to the analysis of the early stages of galaxy formation.

On **small scales**, where $\Delta\rho/\langle\rho\rangle \gg 1$ they are **poor** assumptions:

local matter dominates the metric which does **not** expand (planets; stars; galaxies; clusters)
electrostatic or quantum forces can also overcome the expansion (atoms, people, rulers)

- The form of the RW-metric is: [warning: other equivalent forms exist; this one is common].

$$ds^2 = -c^2\,dt^2 + a(t)^2\,[\,dr_o^2 + R_o^2\,S_k^2(r_o/R_o)\;d\psi^2\,]$$

where r_o is the **comoving proper distance** (ie as measured today) to an object.

This looks very familiar!

it has the **spatial** form introduced above
it has a **scale factor**, $a(t)$, which tracks expansion (usually, $a(t_{now})$ is set to 1)

- The primary parameters of the RW-metric are:
  - $k = +1, 0, -1$ specifies **+ve, flat, -ve curvature**, and the corresponding S_k
  - R_o = **current radius of curvature**; units of length (cf. Gaussian curvature: $\kappa \equiv 1/R_o^2$) .
  - $a(t)$ = a **time varying scale factor**, a dimensionless function of cosmic time, t.

  Much of 20th century cosmology was a struggle to find these three things.

- As the universe expands, its radius of curvature **increases** (just like a sphere).
  So why does the RW-metric only include the **current** radius of curvature, R_o, and not $R(t)$?
  In fact, $R(t)$ **is** present, implicitly:
  - one can show [6biii] that $R(t) = a(t)\,R_o$, and this factor is indeed present in the $d\psi$ coefficient
  - r_o is a **comoving** radius, ie as measured **today**, hence r_o/R_o in S_k is correct at **all** times.

  Please don't think of $R(t)$ as "the radius of the universe"; it is a measure of **spatial curvature**
  although for $k = +1$ it yields roughly the correct total volume, for $k = -1$ it is negative.

  Also, the limiting condition near $k = 0$ with $R \to \infty$ is well behaved, since $R\sin(r/R) \to r$.

- It is important to recognise that the RW-metric is **independent of General Relativity**.
  $\rightarrow$ it comes just from requiring isotropy & homogeneity at all times.
  GR enters by specifying how the curvature and scale factor depend on the universe's **contents**
  We'll turn to this topic in a moment, after briefly getting familiar with the RW-metric.

(c) Simple Properties of the RW-Metric

Let's just check a few simple results which we derived intuitively in sec 3:

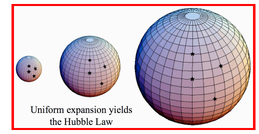
(i) Proper Time

- What time do fundamental observers witness?
  For them, $dr_o = d\psi = 0$ and so $ds^2 = -c^2 d\tau^2 = -c^2 dt^2$ giving $t = \tau$
  As expected, fundamental observers experience a proper time, on which all can agree.

(ii) Proper Distance

- At time t, how far away is a galaxy with **current** (comoving) coordiante r_o?
  by "how far away", we mean "what's the **proper distance**": how many rulers would span the gap?
  Between us and it is purely radial, so $d\psi = 0$, and "at time t" means $dt = 0$, so the RW-metric gives
  $\rightarrow$ $ds = a(t)\,dr_o$ and the total proper distance is $r = \int ds = \int a(t)\,dr_o = a(t)\int dr_o = a(t)\,r_o$
  As we suspected, the proper distance, r, is simply the current (comoving) distance scaled by a(t).

(iii) Velocity-Distance Law



- Let's recover the velocity distance (expansion) relation: [image]
  proper velocity, v, is the rate at which r increases: dr/dt. So, using $r = a(t)\, r_o$, we have:

  $$v = dr/dt = da/dt \times r_o = da/dt \times r/a = (1/a)\,(da/dt) \times r$$

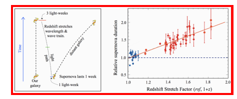
  which is our velocity distance law $v = H \times r$ with Hubble parameter $H = (1/a)\,(da/dt)$

(iv) Time Dilation & Redshift

- A photon emitted from a distant galaxy at r_e, travels to us with: $d\psi = 0$ and $ds = 0$
  (it travels along a radial null geodesic); so:

  $$0 = -c^2dt^2 + a^2\, dr_o^2 \rightarrow c\, dt/a = dr_o \rightarrow \int c\, dt/a = \int dr_o = r_o \quad \text{which is } \textbf{constant}$$

  Two photons are emitted at t_e and $t_e + \Delta t_e$ and arrive at times t_o and $t_o + \Delta t_o$
  During the time $t_e + \Delta t_e$ to t_o both photons are in flight and so $\int dt/a$ for this interval is the same.
  But for the full trip $\int dt/a$ is also the same, so the small start/finish contributions must be equal:

  $$\Delta t_e / a(t_e) = \Delta t_o / a(t_o) \rightarrow \Delta t_o / \Delta t_e = a(t_o) / a(t_e) = 1 / a(t_e) \quad (\text{since } a(t_o) \equiv 1)$$

  This tells us that the duration of any event we witness is **dilated** by a factor $a(t_e)^{-1}$
  This has been nicely confirmed in the longer apparent duration of high-z SNIa light curves: [image]

- Taking Δt_e and Δt_o to be the time between wavecrests of light, we have the cosmological redshift:

  $$\Delta t_o / \Delta t_e = \nu_e / \nu_o = \lambda_o / \lambda_e = (1 + z) = a(t_e)^{-1}$$

  Once again, it is best to think of redshift as a change in scale factor during the photon's journey.

(v) Total Cosmic Volume

- What's the total volume of a closed universe?
  At comoving coordinate r_o and time t, we have a radial vector of proper length: $r = a(t)\, r_o$
  however, swinging this radial vector by $d\psi$ only sweeps out a distance: $a(t)\, R_o \sin(r_o/R_o)\, d\psi$
  the full circumference is: $2\pi\, a(t)\, R_o \sin(r_o/R_o)$, and the full area is: $4\pi\, a(t)^2 R_o^2 \sin^2(r_o/R_o)$
  to get the volume, we must integrate this area over dr from 0 to the antipole at $r = \pi\, a(t)\, R_o$

  $$V = 4\pi\, a(t)^2 R_o^2 \int \sin^2(r_o/R_o)\, dr = 2\pi^2\, a(t)^3 R_o^3$$

  as expected, the volume grows with $a(t)^3$ and is close to the value for a 2-sphere of radius R
  for an open universe the integral diverges because (i) the area diverges, and (ii) $r \rightarrow \infty$

Next Prev Top

(6) Cosmic Dynamics: the Friedmann Equations

Homogeneity, isotropy and the Cosmological Principle alone have singled out the RW-metric as the only credible cosmological spacetime. However, the RW-metric still contains some arbitrary quantities: a(t), R_o, k. These quantities are determined by whatever generates the spacetime geometry. Einstein's General Relativity tells us this:

$$G_{\mu\nu} = 8\pi G/c^2\, T_{\mu\nu}$$

Spacetime is curved by the distribution of cosmic energy & momentum.
Here, μ and ν are four (1 time, 3 space) coordinate indices, (eg ct, x, y, z; or ct, r, θ, ϕ)

(a) Gravity's Field Equations

(i) $T_{\mu\nu}$: The Energy-Momentum Tensor

- $T_{\mu\nu}$ has 4 × 4 = 16 elements, arranged in a square,
  For an isotropic fluid, all the off-diagonal (cross coordinate) elements are zero: $T_{\mu\nu}$ is **diagonal**.

$T_{0,0} = E = \rho c^2$ is the total energy density

$T_{1,1}, T_{2,2}, T_{3,3} = < \mathbf{p_x p_x} > c^2 / E$ is the x-momentum density $\equiv p_x$ the x-pressure (y, z etc)

Now, pressure is isotropic, so $p_x = p_y = p_z = p$ (don't confuse momentum **p** with pressure p)
So $T_{\mu\nu} = \text{diag}(\rho c^2, p, p, p)$ (all off-diagonal elements are zero).

(ii) $G_{\mu\nu}$: The Curvature Tensor

- So much for $T_{\mu\nu}$, what about $G_{\mu\nu}$?
  On the LHS, it provides a description of the **spacetime curvature** in various "directions".
  In general, its components are complex combinations of 2nd order coordinate partial differentials (it is, in effect, replacing ∇^2 in the Newtonian Poisson equation).
  [The general curvature is the **Reimann tensor** of rank 4, $R_{\mu\nu,\gamma,\delta}$ with $4^4 = 256$ elements.
  Symmetry reduces this to the rank 2 **Ricci tensor**, $R_{\mu\nu}$, and its scaler trace, the **Ricci curvature**, R.
  Finally, the curvature is expressed as: $G_{\mu\nu} = R_{\mu\nu} - ½\, g_{\mu\nu} R$ where $g_{\mu\nu}$ are the metric coefficients]

- There are two very important constraints for the cosmological spacetime:
  - Everywhere the spacetime is **locally flat** → the geometry is **Reimannian**.
    recall, this is ultimately rooted in the Equivalence Principle
  - **Isotropy** ensures all off-diagonal elements are zero: $G_{\mu\nu}$ is **diagonal**
    furthermore, spatial isotropy demands $G_{1,1} = G_{2,2} = G_{3,3}$

Evaluating the elements for the RW-spacetime, one obtains:

$$G_{0,0} = 3/a^2\, [k\, c^2/R_o^2 + (da/dt)^2]$$

$$G_{1,1} = G_{2,2} = G_{3,3} = -1/a^2\, [\, 2\, a\, (d^2a/dt^2) + k\, c^2/R_o^2 + (da/dt)^2]$$

(iii) Two Cosmological Equations

- Now equate these elements in $G_{\mu\nu}$ to the same elements of $T_{\mu\nu}$ & add the proportionality constant:

$$G_{0,0} = 3/a^2\, [k\, c^2/R_o^2 + (da/dt)^2] = 8\pi G \rho = 8\pi G/c^2\, T_{0,0}$$

$$G_{j,j} = -1/a^2\, [\, 2\, a\, (d^2a/dt^2) + k\, c^2/R_o^2 + (da/dt)^2] = 8\pi G\, p/c^2 = 8\pi G/c^2\, T_{j,j}$$

Combining these, we arrive at two **fundamentally important cosmic equations**

$(da/dt)^2 = (8\pi G/3)\, \rho a^2 - k\, c^2/R_o^2$	The **Friedmann Equation**, or Energy Equation
$d^2a/dt^2 = -(4\pi G/3)\, a\, (\rho + 3p/c^2)$	The **Acceleration Equation**

where k = +1, 0, -1 follows the sign of the curvature radius R_o; and a(t) is the scale factor.

- In fact, the acceleration equation can be derived from the energy, using $p = w\, \rho c^2$ (i.e. $\rho = \rho_o a^{-3(1+w)}$)
  Differentiate the Friedmann equation and cancel through by da/dt:

$$d^2a / dt^2 = (4\pi G/3)\, [2\rho a + a^2\, d\rho/da\,]$$

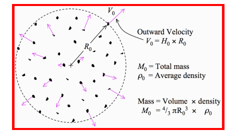
Substitute for $d\rho/da$ and rearrange:

$$d^2a / dt^2 = (4\pi G/3)\, [\, 2\rho a - 3(1+w) a \rho\,] = (-4\pi G/3)\, a\, [\, \rho + 3p/c^2\,]$$

- All we need, therefore, is Freedman's first equation, and our expressions for $\rho(a) = \rho_o\, a^{-3(1+w)}$.
  → These alone give the full expansion history a(t)
  Freedman's second equation simply makes explicit the form of the acceleration.

(iv) A Newtonian Analog

- It is remarkable that essentially all these results can be found in a simple Newtonian analysis.



Consider a huge sphere uniformly but sparsely filled with rubble [image]
Focus on a single rock at radius r_o: it feels only interior mass: $M = 4/3\,\pi r_o^3 \rho_o$
To follow the rock's motion, define radial coordinate $r = a(t)\, r_o$ (track using a scale factor)
Set the whole sphere expanding, with $dr/dt = r_o\, da/dt$ for our rock.

- What is the equation of motion for the rock?
  Its total energy (per unit mass) is the sum of its kinetic and potential energies:
  $$½(dr/dt)^2 \; - \; GM/r \; = \; TE \quad \text{(TE is also called the binding energy).}$$

  This is just the equation for **throwing a stone vertically upwards**:
  if TE > 0, then $v > v_{esc}$ and the rock escapes; if TE < 0, then $v < v_{esc}$ and the rock returns.

  Rewriting this equation using the changing scale factor and density, we get:

  $$(da/dt)^2 \; - \; (8\pi G/3)\, \rho\, a^2 \; = \; 2\,TE / r_o^2$$
  "KE" "PE" "TE"

  which is the Friedmann energy equation with $2\,TE / r_o^2$ standing in for curvature: $-k\, c^2/R_o^2$
  → positive TE ≡ open geometry
  → negative TE ≡ closed geometry

(b) Cosmological Parameters

Before we solve for a(t), let's draw together some useful cosmological parameter relations:

(i) Hubble Parameter: H

- Scale factor: $a = (1 + z)^{-1}$ (a = 1 now)

- Hubble Parameter: $H = (1/a)\,(da/dt)$ (varies with time)
  Hubble Constant: $H_o = (da/dt)_o = 72 \times 10^{-6}\ \text{Myr}^{-1}$ (today's value, psm units)

- Velocity factor: $v/v_o = H\, r / H_o\, r_o = a \times H/H_o$
  This gives the expansion velocity history for a given object, scaled to 1 today.

(ii) Density Parameters: ρ, Ω

- Critical density: $\rho_c = 3H^2 / 8\pi G$ (varies with time)
  Today's value: $\rho_{c,o} = 3H_o^2 / 8\pi G = 1.37 \times 10^{-7}\ M_\odot\ \text{pc}^{-3}$

- Density parameters: $\Omega = \rho / \rho_c = (8\pi G/3H^2)\, \rho$ (vary with time)
  Example of today's value: $\Omega_{m,o} = \rho_{m,o} / \rho_{c,o} = (8\pi G/3H_o^2)\, \rho_{m,o}$

- Summing gives the total density: $\rho_t = \rho_m + \rho_r + \rho_v$
  Alternatively: $\Omega_t = \Omega_m + \Omega_r + \Omega_v$ (all these vary with time)

- Including their dependence on scale factor [sec 4e] and inserting today's (known!) values gives:

  $$\rho_t(a) / \rho_{c,o} = \Omega_{m,o}\, a^{-3} + \Omega_{r,o}\, a^{-4} + \Omega_{v,o}$$

  This is an **exceedingly important relation**:
  it gives the full density evolution as a function of a or z using **today's** measured values.
  [warning: $\rho_t(a) / \rho_{c,o}$ is **not** the evolving density parameter, $\Omega_t(a) \equiv \rho_t(a) / \rho_c$, see sec 6cvii]

(iii) Curvature Parameters: R_o & Ω_k

- What about an expression for the curvature radius, R_o?
  Start with the Friedmann equation and divide by a^2:

  $$(1/a^2)(da/dt)^2 = H^2 = (8\pi G/3)\, \rho \; - \; kc^2/a^2 R_o^2$$

substitute for $\rho = 3H^2\Omega_t / 8\pi G$ to get $H^2 = H^2\Omega_t - kc^2 / a^2R_o^2$ which yields:

$$R_o = (c/H_o) / [k(\Omega_{t,o} - 1)]^{1/2}$$ [notice $k(\Omega_{t,o} - 1)$ is always positive]

This is as expected: $R_o \rightarrow \infty$ for $\Omega_{t,o} \rightarrow 1$
Current estimates put $\Omega_{t,o} - 1 < 0.02$, so $R_o > 7\ c/H_o = 30$ Gpc.
At earlier times, $R = a\,R_o$ is smaller.

- It is also useful to express the curvature term as an **effective density**:

$$\Omega_k \equiv -kc^2/R^2H^2 = -kc^2/a^2R_o^2H^2 \quad \text{and} \quad \Omega_{k,o} \equiv -k\,c^2/R_o^2H_o^2$$

Inserting into the Friedmann equation we find: $H^2 = H^2\Omega_t + H^2\Omega_k \rightarrow 1 = \Omega_t + \Omega_k$
Notice, Ω_k and k have **opposite** sign: +ve Ω_k means **open** geometry

We can also re-express R_o in terms of $\Omega_{k,o}$ and the Hubble distance $r_{H,o} = c / H_o$:

$$R_o = r_{H,o} / |\Omega_{k,o}|^{1/2}$$

(iv) Deceleration Parameter: q

- What about a parameter for acceleration/deceleration?
A dimensionless parameter is: $q = -a\,(d^2a/dt^2) / (da/dt)^2 = -(d^2a/dt^2) / (aH^2)$

To find an expression for this, start with the acceleration equation & divide by a:

$$1/a\,(d^2a/dt^2) = -H^2 q = -(4\pi G/3)(\rho + 3p/c^2) = -(4\pi G/3)\,\rho\,(1 + 3w)$$

so:

$$q = (4\pi G/3H^2)(3H^2/8\pi G)\,\Omega\,(1 + 3w) = \tfrac{1}{2}\,\Omega\,(1 + 3w)$$

- In practice, we should sum over the various components, each with their own w:

$$q = \tfrac{1}{2}[\Omega_m + 2\Omega_r - 2\Omega_m] = \tfrac{1}{2}\Omega_m + \Omega_r - \Omega_v$$ at any time, or

$$q_o = \tfrac{1}{2}\Omega_{m,o} + \Omega_{r,o} - \Omega_{v,o}$$ today.

- Notice that q is defined to be **+ve** for **deceleration**
In our pre-Ω_v cosmology days (1970s-90s), this was always +ve, with $q_o = \tfrac{1}{2}\,\Omega_o$ ($\Omega_{r,o}$ is negligible)
But with a vacuum term, q can be either -ve or +ve depending on Ω_m and Ω_v
Currently, with $\Omega_m = 0.26$ and $\Omega_v = 0.73$, we have $q_o = -0.6$: **we are accelerating**.

(v) Cosmological Constant: Λ

- What about **the** cosmological parameter: Λ
Recent fasion (& these notes) treat this term by simply adding a vacum component
Historically, however, it first appeared as a possible term in $G_{\mu\nu}$:
$$G_{\mu\nu} = R_{\mu\nu} - \tfrac{1}{2}R\,g_{\mu\nu} + \Lambda\,g_{\mu\nu}$$

- This propagates into the Friedmann & Acceleration equations:

$$(da/dt)^2 = (8\pi G/3)\,\rho a^2 - k\,c^2/R_o^2 + 1/3\,\Lambda\,a^2$$

$$d^2a/dt^2 = -(4\pi G/3)\,a\,(\rho + 3p/c^2) + 1/3\,\Lambda\,a$$

Since the new term appears with the same power of a, we can simply incorporate it in ρ and p:
this is done by defining Λ in terms of ρ_v, and discovering that $w = -1$:

$$\Lambda = 8\pi G\,\rho_v = 3\,H^2\,\Omega_v \quad \text{(units of time}^{-2}\text{)} \quad \text{and} \quad \rho_v = -p_v/c^2 \quad \text{(i.e. } w = -1\text{)}$$

(c) The Many FRW World Models

- Let's now use the Friedmann equation to study the **time development** of the scale factor: $a(t)$.
  → basically, derive those expansion/contraction graphs we've seen since high school.

  The form of $a(t)$ depends on two things:

  - the curvature: closed, critical, open, which depends on Ω_t
  - the form of the density relation, $\rho(a)$, which depends on the fluid.

  A specific solution is called an **FRW world model** (FRW = Friedmann, Robertson, Walker)

  Let's begin with the general case, then look at some special cases.

(i) The General Case

- Since the LHS of the Friedmann equation is a constant, it is equal to its value today ($a = 1$):

$$(da/dt)^2 - (8\pi G/3)\, a^2 \rho_t = (da/dt)_o^2 - (8\pi G/3)\, \rho_{t,o}$$

substitute for $8\pi G/3 = H_o^2/\rho_{c,o}$ and $(da/dt)_o = H_o$:

$$\begin{aligned}(da/dt)^2 &= (H_o^2/\rho_{c,o})\, a^2 \rho_t + H_o^2 - (H_o^2/\rho_{c,o})\, \rho_{t,o} \\ &= H_o^2 a^2 \, [\, \Omega_{m,o} a^{-3} + \Omega_{r,o} a^{-4} + \Omega_{v,o} + (1 - \Omega_{t,o})\, a^{-2} \,] \\ &= H_o^2 a^2 \, [\, \Omega_{m,o} a^{-3} + \Omega_{r,o} a^{-4} + \Omega_{v,o} + \Omega_{k,o} a^{-2} \,] \\ &= H_o^2 a^2 \, E^2(a) \qquad \text{giving:}\end{aligned}$$

$(da/dt) = H_o\, a\, E(a)$ for the evolution of the scale factor

where we use Peeble's notation for $E(a)$ [or $E(z)$], and the curvature "density" $\Omega_{k,o}$:

$E(a) \equiv [\, \Omega_{m,o} a^{-3} + \Omega_{r,o} a^{-4} + \Omega_{v,o} + \Omega_{k,o} a^{-2} \,]^{1/2}$
$E(z) \equiv [\, \Omega_{m,o} (1+z)^3 + \Omega_{r,o} (1+z)^4 + \Omega_{v,o} + \Omega_{k,o} (1+z)^2 \,]^{1/2}$
$\Omega_{k,o} \equiv 1 - \Omega_{t,o} = 1 - \Omega_{m,o} - \Omega_{r,o} - \Omega_{v,o}$

These are **exceedingly important functions**, and are central to all cosmological calculations.
Notice that they involve **current** (hence observable) values of Ω, and so can be evaluated directly.

- We are now ready to evaluate $a(t)$, the time evolution of the scale factor, by simple integration:

$$da/dt = H_o\, a\, E(a) \quad \text{giving} \quad \int da \,/\, a\,E(a) = H_o \int dt = t \,/\, t_{H,o} \qquad \text{(da from 0 to a; dt from 0 to t)}$$

In general, this and related integrals need to be done **numerically**.

Alternatively, one can solve the ODE with boundary conditions: ($a = 1$; $da/dt = H_o$) at $t = t_{now}$.
Note: for numerical work, it is sensible to express time in units of $t_{H,o} = 1/H_o$

- Notice we have three related variables here: cosmic time, t; scale factor, a; and redshift, z.
  you can move between them with the following **important** relations between their differentials:

$dt = t_{H,o}\, da \,/\, aE(a) = -t_{H,o}\, dz \,/\, (1+z)E(z)$ and $da = -dz/(1+z)^2 = -a^2 dz$

where $t_{H,o} = 1 / H_o$ is the Hubble time; one can also use the Hubble radius $c\, dt = r_{H,o}\, da \,/\, aE(a)$ etc.
For example, the relation for current cosmic age is simply:

$$\int dt \ \text{(0 to now)} = -t_{H,o} \int dz \,/\, (1+z)E(z) \ (\infty \text{ to } 0) = t_{H,o} \int da \,/\, aE(a) \ \text{(0 to 1)}$$

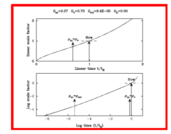
Let's now look at the time evolution of some specific FRW world models.

(ii) The Concordance Model

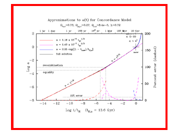
- The real Universe has several components; with current densities:

$$\Omega_{v,o} = 0.73; \quad \Omega_{m,o} = 0.27; \quad \Omega_{r,o} = 8.4 \times 10^{-5}; \quad \Omega_{t,o} = 1.00 \qquad (\Omega_{k,o} = 1 - \Omega_{t,o} = 0)$$

Integrating the general relation gives a current age 0.988 $t_{H,o}$ and a(t) curve shown here: [image]



- In practice, each term is dominant over a certain range in a, giving **three eras** :
  - during the radiation era: $E(a) \approx \Omega_{r,o}^{1/2} \; a^{-2}$
  - during the matter era: $E(a) \approx \Omega_{m,o}^{1/2} \; a^{-3/2}$
  - during the dark energy era: $E(a) \approx \Omega_{v,o}^{1/2}$



Taken individually, these yield reasonable approximations for a(t) over most of cosmic history.
This [image] shows the approximations and their errors, while the table gives the functional forms.
[The Toolbox includes a more extensive set, including Hubble radius, particle and event horizons.]

Approximations to the Concordance model
Radiation era: $E(a) \approx \Omega_{r,o}^{1/2} \; a^{-2}$
$t \; = \; t_{H,o} \int da \,/\, a\,E(a) \;\; (0 \text{ to } a) \;\approx\; \tfrac{1}{2}\, t_{H,o}\, \Omega_{r,o}^{-1/2}\, a^2 \;\approx\; 741\, a^2\, h_{72}^{-1}$ Gyr $= \; 2.34 \times 10^{19}\, a^2\, h_{72}^{-1}$ sec
$a \;\approx\; 1.16 \times 10^{-6}\, h_{72}^{1/2}\, t_{yr}^{1/2} \;\approx\; 2.09 \times 10^{-10}\, h_{72}^{1/2}\, t_s^{1/2}$
Matter era: $E(a) \approx \Omega_{m,o}^{1/2} \; a^{-3/2}$
$t \; = \; t_{H,o} \int da \,/\, a\,E(a) \;\; (0 \text{ to } a) \;\approx\; 2/3\; t_{H,o}\, \Omega_{m,o}^{-1/2}\, a^{3/2} \;\approx\; 17.4\, a^{3/2}\, h_{72}^{-1}$ Gyr
$a \;\approx\; 1.49 \times 10^{-7}\, h_{72}^{2/3}\, t_{yr}^{2/3}$
Dark Energy Era: $E(a) \approx \Omega_{v,o}^{1/2}$
$t - t_{now} \; = \; \Delta t \; = \; t_{H,o} \int da \,/\, a\,E(a) \;\; (1 \text{ to } a) \;\approx\; t_{H,o}\, \Omega_{v,o}^{-1/2} \ln(a) \;\approx\; 15.9 \ln(a)\, h_{72}^{-1}$ Gyr
$a \;\approx\; \exp[\, \Delta t_{Gyr}\, h_{72} \,/\, 15.9 \,] \qquad (t_{now} = 0.988\, t_{H,o})$

Two details:
Formally, the exponential term has infinite past, so it is sensible to integrate from a=1, the current time.
The matter solution is improved slightly by adding 39,000 years to correct the integral over the radiation era.

(iii) Velocity History and Newtonian Energy Diagrams

- There is a remarkably transparent way to present the solutions to the Friemdann equations.
  It mirrors the Newtonian analysis, using (dimensionless) kinetic, gravitational, and total energies.

  Follow the above analysis, starting from the first line:

$$(da/dt)^2 \; - \; (8\pi G/3)\, a^2 \rho_t \; = \; (da/dt)_o^2 \; - \; (8\pi G/3)\, \rho_{t,o}$$

$$(da/dt)^2 \; - \; (H_o^2/\rho_{c,o})\, a^2 \rho_t \; = \; H_o^2 \; - \; (H_o^2/\rho_{c,o})\, \rho_{t,o}$$

Introduce the velocity factor, $v/v_o \; = \; Hr \,/\, H_o r_o \; = \; a\, H/H_o \; = \; (da/dt) \,/\, H_o$
and express the time-varying density ρ_t in terms of today's Omegas, we get:

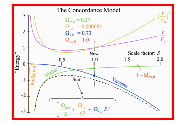
$$(v/v_o)^2 \; + \; [\, -\Omega_{m,o}/a \;\; -\Omega_{r,o}/a^2 \;\; -\Omega_{v,o} a^2 \,] \; = \; 1 \; - \; \Omega_{t,o}$$

"KE" "PE" "TE" (const)

Or using the previous notation: $v/v_o \; = \; a\, E(a)$.
Of course, to convert v(a) to a(t) we still need to integrate, just as before:

$v/v_o \; = \; (da/dt) \,/\, H_o \; = \; a\, E(a)$ giving:
$\int da \,/\, aE(a) \;\; (0 \text{ to } a) \; = \; H_o \int dt \;\; (0 \text{ to } t) \; = \; t \,/\, t_{H,o}$

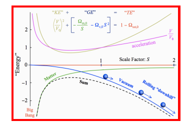
- Notice how our equation follows the Newtonian form, with kinetic, gravitational, and total energy terms.
  Plotting the three terms immediately gives us a feel for the behavior of the expansion: [image]
  → the dropping velocity matches the rising gravitational energy during the radiation/matter eras
  → the increasing velocity matches the dropping vacuum gravitational energy during the dark energy era.



- Let's try to understand the gravitational terms, since they are the key to the behavior.
  Returning to our upward moving rock at the surface of the expanding sphere, what is it's gravitational energy?
  There are three contributions: from the matter, radiation and vacuum within the sphere.
  The sphere's mass in matter is **constant**, so the rock's PE increases as $-GM_{m,o}/r$ giving $-\Omega_{m,o} / a$.
  The sphere's mass in radiation **decreases**, so the rock's PE increases faster, as $-GM_{r,o}/r^2$ giving $-\Omega_{r,o} / a^2$.
  The sphere's mass in vacuum **increases**, so the rock's PE actually falls, as $-GM_{v,o}/r \times r^3$ giving $-\Omega_{v,o}\, a^2$.

  Since the total energy term $(1 - \Omega_{t,o})$ is fixed, the expansion velocity must mirror the gravitational terms.

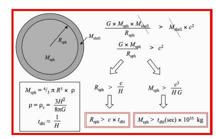
- We can understand dark energy's accelerating expansion by considering its **gravitational energy**. [image]
  Because larger spheres of vacuum have more mass, their gravitational energy is **more negative**
  Hence, expansion is "downhill" → vacuum Universe's **fall outwards**.



- One often hears that dark energy "makes gravity repulsive", which refers to the **force** of gravity.
  But force is really just the spatial gradient of energy: i.e. the gradient of the curve in the figure
  This curve $(-\Omega_{v,o}\, a^2)$ **slopes downwards to the right**, hence the force is **outwards**
  Notice, gravity is still "attractive", meaning its interaction still causes negative binding energy.

- The time of coasting is at the top of the total binding energy curve, with gradient (force) zero.
  → Matter force: $d/da\,(-\Omega_{m,o} / a) = \Omega_{m,o} / a^2$ (backwards, inverse square, as expected)
  → Vacuum force: $d/da\,(-\Omega_{v,o}\, a^2) = -2\Omega_{v,o}\, a$ (outwards, increasing with a).
  Notice, kg for kg vacuum generates **twice** the force of matter, so coasting occurs **before** density equality.
  We have: $\Omega_{m,o} / a_c^2 - 2\Omega_{v,o}\, a_c = 0$ → $a_c = (\Omega_{m,o} / 2\Omega_{v,o})^{1/3} = 0.57$ (z = 0.75, t = 7 Gyr).

- One may wonder how a vacuum sphere can simply gain mass -- where does that mass come from?
  Amazingly, the **gravitational energy** released as the sphere falls outwards **makes new vacuum!**.
  [Strictly, this only happens for spheres larger than the Hubble radius. image]
  This is remarkable: gravity makes the very space into which the Universe is expanding!



(iv) Flat Models: Single Component

- The special case of flat geometries with a single component are good pedagogical models.
  They also illustrate the behavior of the various "eras" in the multi-component models.

- We use $t = t_{H,o} \int da / aE(a)$ (from 0 to a); with E(a) containing just one term:

  1. Pure **matter**: $\Omega_{m,o} = 1$; $a\,E(a) = [a^{-1}]^{1/2}$ $t = (2/3)\, t_{H,o}\, a^{3/2}$ $a = [3/2\; t/t_{H,o}]^{2/3}$
  2. Pure **radiation**: $\Omega_{r,o} = 1$; $a\,E(a) = [a^{-2}]^{1/2}$ $t = (1/2)\, t_{H,o}\, a^2$ $a = [2\; t/t_{H,o}]^{1/2}$
  3. Pure **vacuum**: $\Omega_{v,o} = 1$; $a\,E(a) = [a^{2}]^{1/2}$ $t - t_o = t_{H,o} \log_e a$ $a = e^{[(t - t_o)/t_{H,o}]}$
  4. Pure **curvature**: $\Omega_{k,o} = 1$; $a\,E(a) = [\Omega_{k,o}]^{1/2}$ $t = t_{H,o}\, a$ $a = t / t_{H,o}$

  These recover the forms from [sec 4e]: $a \propto t^{2/3}, t^{1/2}, e^t$ for matter, radiation, vacuum.
  The pure curvature model is **open** (not flat), with simple linear expansion at all times.
  Pure matter & curvature models are also called: **Einstein-de Sitter** & **Milne** [sec 6cvii]
  Note: the coefficients are **different** from the concordance approximations because the current Ω's aren't unity.

- For a general equation of state parameter w, we have [sec 4e]:

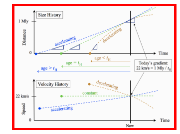
  $a\,E(a) = [a^{-(1+3w)}]^{1/2}$ → $t = 2/(3+3w)\; t_{H,o}\, a^{(3+3w)/2}$ → $a = [\,(3 + 3w)/2\; (t / t_{H,o})\,]^{2/(3+3w)}$

  you can quickly check this gives 1 & 2 above, for w = 0 and 1/3.
  Note also that the flat model with w = -1/3 behaves like an open pure curvature (Milne) model.

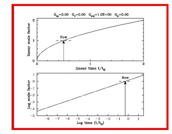
- It is easy to find the **current age** of all these models: just set a = 1 and solve for t:

$$t_{age} \; = \; 2/(3+3w)\, t_H \;\rightarrow\; 2/3\, t_H\,,\;\; 1/2\, t_H\,,\;\; \infty\,,\;\; t_H \quad \text{for 1 - 4 above.}$$

Here I chose to use t_H in place of $t_{H,o}$ since the relation for t_{age} is true at all times.
As we suspected from the outset, $t_{age} \approx t_H$ to within factors ~unity.
The exception is pure vacuum, with exponential expansion and infinite age ($t \rightarrow -\infty$ as $a \rightarrow 0$)
In general, deceleration gives $t < t_H$ while acceleration gives $t > t_H$ [image]



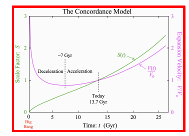
Note that $w = -1/3$ separates ac-/de-celerating models, and hence t_{age} greater/less than $t_{H,o}$
For example, for flat matter (Einstein-de Sitter; $w = 0$) we have $t_{age} = 2/3\, t_{H,o}$, a famous result.
By including a component with $w < -1/3$ (eg vacuum/lambda), then t_{age} can get **longer** than $t_{H,o}$
This was a key motivation for including Λ to solve $t_{H,o} < t_{\oplus}$ (1930s), or $t_{H,o} < t_{GC}$ (1990s).



- Some more a(t) curves for various models are given here [images]

(v) Flat Models: Matter + Vacuum

- As it happens, it **is** possible to derive an analytic form for a **flat** universe with both Ω_m and Ω_v
  Apart from a brief (~50 kyr) radiation period, **this is basically the Universe we live in** [image]



- As an exercise in integration, follow the analysis through to derive the following result:

$$a(t) \;=\; (\Omega_{m,o} / \Omega_{v,o})^{1/3} \;\; \sinh^{2/3} [\, (3/2)\, \Omega_{v,o}^{\;1/2}\, (t / t_{H,o})\,] \;=\; 0.712 \times \sinh^{2/3} (1.28 \times t / t_{H,o})$$

$$t_{age} \;=\; (2/3)\, t_{H,o}\, \Omega_{v,o}^{\;-1/2} \sinh^{-1} [\, (\Omega_{v,o} / \Omega_{m,o})^{1/2}\,] \;=\; 0.78 \times 1.27\, t_H \;=\; 0.993\, t_{H,o}$$

which recovers an age conveniently close to $t_{H,o} \equiv H_o^{-1}$
i.e. the decelerating & accelerating portions "average out" to approximate a constant expansion.
(the numbers here use: $\Omega_{v,o} = 1 - \Omega_{m,o} = 0.73$)

(vi) Curved Models: Matter Only

- For a long time (1960s-90s), curved matter models were **favoured**, justifying their inclusion here:
  $\rightarrow$ open with infinite future; closed ending in a big crunch; or "critical", balanced between the two

  There is, however, a more compelling reason: they are crucial in theories of **galaxy formation**
  $\rightarrow$ while the real Universe seems to be **globally** flat, this is not true **locally**:
  $\rightarrow$ voids form from **open** regions; while stars/galaxies/clusters form from **closed** regions
  FRW curved matter models apply since local evolution is independent of the surroundings.

- For a single component: $\rho_s = \rho_{s,o}\, a^{-3(1+w)}$ and $a\,E(a) = [\Omega_{s,o}\, a^{-(1+3w)} + (1 - \Omega_{s,o})]^{1/2}$
  For $w = 0$ (matter) we have $t(a) = t_{H,o} \int [\, \Omega_{m,o}\, a^{-1} + 1 - \Omega_{m,o}\,]^{-1/2}\, da$ (from 0 to a)

  This has a parametric solution which uses a "development angle", η
  The form of the solution depends on whether $\Omega_{m,o} > 1$ (closed) or $\Omega_{m,o} < 1$ (open):

- For **closed** models with $\eta = 0$ to 2π, the Universe expands, halts, and recollapses in a **cycloid**:

$$a(\eta) \;=\; ½\, a_{max}\, (1 - \cos \eta)$$

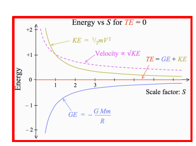
$$t(\eta) \;=\; ½\, t_H\, a_{max}\, (\eta - \sin \eta) / (\Omega_{m,o} - 1)^{1/2}$$

which turns at $a(\pi) = a_{max} = \Omega_{m,o} / (\Omega_{m,o} - 1)$ and collapses at $t(2\pi) = \pi\, t_{H,o}\, a_{max} / (\Omega_{m,o} - 1)^{1/2}$
for example, if $\Omega_{m,o} = 1.5$, we have $a_{max} = 3$ and $t_{crunch} = 6.7\, t_{H,o}$

- For **open** models, the Universe expands forever with similar solution, with $\eta = 0$ to ∞

$$a(\eta) \;=\; ½\, a_{par}\, (\cosh \eta - 1)$$

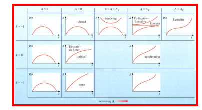
$$t(\eta) \;=\; ½\, t_{H,o}\, a_{par}\, (\sinh \eta - \eta) / (1 - \Omega_{m,o})^{1/2}$$



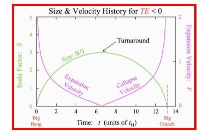
where $a_{par} = \Omega_{m,o} / (1 - \Omega_{m,o})$ plays the role of a_{max}

- Newtonian Energy curves and the Expansion Histories for these matter models are shown here: [images]

(vii) Specific & Named World Models



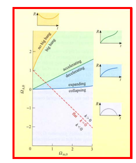
- This [figure] illustrates most classes of FRW world models.
- The two key parameters are
  - **curvature**: closed ($k = +1, \Omega_k$ -ve), flat ($k = 0, \Omega_k = 0$), open ($k = -1, \Omega_k$ +ve)
  - **lambda**: Ω_v: -ve (odd), 0 (simplest), +ve (what seems to be true).
- Full histories in linear time usually consider **pressure-less matter** along with vacuum.
  This is because radiation quickly becomes irrelevant when $\rho_r < \rho_m$ at $t \sim 50$ kyr.



- Referring to the figures, the overall forms are:
  negative Λ: unusual; simply adds to matter; all models **recollapse**
  zero Λ: most discussed until recently; **all** decelerate; future depends only on curvature.
  closed recollapses; flat halts at ∞; open expands forever. [images]
  positive Λ: early, matter dominates $\rightarrow$ decelerates; late, vacuum dominates $\rightarrow$ accelerates.

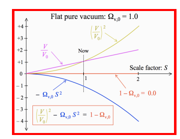
  For a **closed** geometry, additional possibilities include (see below):

  For $\Omega_v < \Omega_{v,E}$ both collapse and "bounce" solutions exist
  For $\Omega_v \approx \Omega_{v,E}$ vacuum repulsion and closed geometry nearly balance $\rightarrow$ long ~static periods.

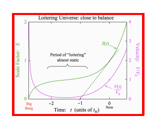


  Ages and **futures** for all these models can be nicely seen in these plots [image]

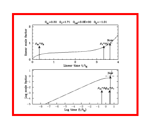
- Some of these models go by the names of those who advocated/studied them:
  - **Einstein-de Sitter**: pure matter, flat; "Critical" classical model, between open & closed.



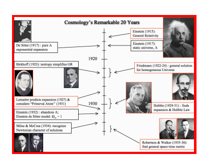
  - **de Sitter**: pure vacuum, flat **exponential expansion**: $a \propto e^t$; infinite past; H(a) = const [image]
    This model has renewed importance, since it is believed to occur during inflation.
    It also describes the Steady State model, which needs H(a) = const.
  - **Einstein**: His original model; curvature/matter/vacuum all **balance** to give **static** Universe.
    pre-dated Hubble's discovery of expansion; doesn't fit above analysis since $H_o = 0$.
    Instead, set $da/dt = d^2a/dt^2 = 0$ to find balance when: $\Omega_m + \Omega_v - 1 = 3/2\ (\Omega_v \Omega_m^{\ 2})^{1/3}$
    The model is unstable to small deviations from balance.
  - **Eddington-Lemaitre**: Two additional solutions to Einstein's model
    Infinitely long static history before starting expansion
    big bang in remote past, then expansion comes to gradual halt at infinite future.



  - **Lemaitre**: close to Einstein model, with Ω_v (Ω_m) slightly **below** (**above**) the critical values.
    Big bang $\rightarrow$ expansion $\rightarrow$ **"loitering"** $\rightarrow$ exponential expansion.
    The motivation for loitering was the too short (incorrect) Hubble time, $t_{H,o} = H_o^{-1}$
    The age can be increased by introducing Λ ; see this example: [image] [image]
    Lemaitre precienly also wanted a big bang to make elements



  - **Milne**: pure curvature (empty, not even vacuum energy); $\Omega_k = +1$, open, ($k = -1$).
    Similar to $\Omega_{m/r/v} << 1$ $\rightarrow$ no de- ac-celeration: linear expansion $a = t / t_H$; $t_{age} = t_H$
    Milne's (incorrect) motivation: moving galaxies subject only to special relativity.



- Some photos of many of the characters we've mentioned so far are shown here: [image]

Next Prev Top

(7) Distances & Horizons

- Discussing distances on an expanding, possibly curved, coordinate grid can be tricky.

So, first let's distinguish between **three** different kinds of distance measurement

- **proper distances**: r are "true" distances - the number of non-expanding rulers between objects
  r increases due to expansion according to the velocity-distance relation $v = dr/dt = H r$
  the **current** proper distance, $r(t_o)$, to an object is also called its **comoving distance**, r_o.
  the proper distances all change with the scale factor: $r = a(t)\, r(t_o) = a(t)\, r_o$

- **pseudo** (my term) **distances**: D are derived from certain measurements....
  eg luminosity (D_L) or angular diameter (D_A), **as if** space-time were static and Euclidean.
  they are not "real" distances so much as convenient **functions of distance**
  they are closely tied to observations and play a crucial role in establishing the world model

- **redshift**: z though not an explicit distance, it is our primary observable
  it is derived from spectra and gives the scale factor when the light set out: $a(t_e) = (1 + z)^{-1}$
  Given a world model, it is relatively easy to derive all the other distances using z.

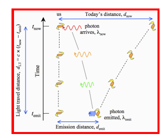
- This section deals mainly with **proper distances**: r
  The following two sections consider , among other things, the pseudo distances: D.

- Before we start, let's recall two useful/sensible units of time and distance for cosmology:

Hubble time:	$t_{H,o} = H_o^{-1} = 10.0\ h^{-1}$ Gyr $= 13.9\ h_{72}^{-1}$ Gyr
Hubble distance:	$r_{H,o} = c / H_o = c\, t_{H,o} = 13.9\ h_{72}^{-1}$ G lyr $= 4.26\ h_{72}^{-1}$ Gpc

they are units comparable to the current age and visible size of the Universe.

(a) Three Proper Distances



- Because the Universe is expanding, the proper distance to an object depends on **time**: $r(t)$
  In fact, there are **three** distances one might consider: [image]

  - The proper distance to the object when the light **set out**: $r(t_e)$ (e = emit)
    if the light has travelled for a long time, $r(t_e)$ could be quite small, but certainly less than:

  - The proper distance to the object when the light **arrives**: $r(t_o)$ (o = observed = now)
    recall, $r(t_o) = r_o$, is also called the **comoving distance**
    Since $t_e < t_o$ then $r(t_e) < r(t_o)$; or using the scale factor: $r(t_e) = a(t_e) \times r(t_o) = a(t_e) \times r_o$

  - The distance **light travelled** during its journey: $r_{lbt} = c(t_o - t_e)$ (lbt: look-back-time)
    expansion guarantees that r_{lbt} is **intermediate** in length between $r(t_e)$ and $r(t_o)$
    the light travel time, $(t_o - t_e)$, is also the **look-back time**, as invoked in "popular" statements:
    "That galaxy is 5 billion light years away, hence we see it as it was 5 billion years ago"

- What is $r(t_o)$ as a function of z? Here are two approaches:
  Imagine a photon at time t on its way to us.
  In interval dt it crosses c dt which then expands to become $c\, dt / a(t)$ at time $t_o = t_{now}$
  We can also use the various forms for dt given in [sec 6ci] to give several equivalent integrals:

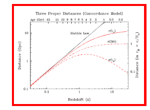
$$r(t_o) = c \int dt / a(t) \ \ (t_e \text{ to } t_o) = r_{H,o} \int da / a^2 E(a) \ \ (a \text{ to } 1) = r_{H,o} \int -dz / E(z) \ \ (z \text{ to } 0)$$

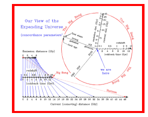
Alternatively, from the RW metric: photons move on radial null geodesics, so $d\psi = ds = 0$, giving:

$$c^2\, dt^2 = a^2(t)\, dr_o^2 \rightarrow c \int dt / a(t) = \int dr = r(t_o) \ \ (t_e \text{ to } t_o \text{ and } 0 \text{ to } r_o) \ \text{as before.}$$

- What about the lookback time? Again, using the alternatives [sec 6ci] for dt:

$$LBT = (t_o - t_e) = \int dt \ \ (t_e \text{ to } t_o) = t_{H,o} \int da / a E(a) \ \ (a \text{ to } 1); = t_{H,o} \int -dz / (1+z) E(z) \ \ (z \text{ to } 0)$$





Which are all straightforward to evaluate (numerically).

- Examples of $r(t_o)$, $r(t_e)$ and $r_{lbt} = c(t_o - t_e)$ are shown here: [images]
  A different presentation, using "The Astronomer's Universe", is shown here [image]

(i) Two Examples: z = 6 & 1000

- Let's do a couple of examples using the Einstein-de Sitter Universe ($\Omega_{m,o} = 1$, $\Omega_{v,o} = \Omega_{k,o} = 0$)
  In this case, $E(a) = a^{-3/2}$ and $E(z) = (1 + z)^{3/2}$

- For a z = 6 QSO: what are $r(t_o)$, $r(t_e)$, & $c\,(t_o - t_e)$
  There are several ways to treat this; let's use the z relations:

$$r(t_o) = r_{H,o} \int -(1 + z)^{-3/2}\, dz \;\; (z \text{ to } 0) = 2\, r_{H,o} [1 - 7^{-1/2}] = 1.24\, r_{H,o}$$

$$r(t_e) = r(t_o) / (1 + z) = 1.24\, r_{H,o} / 7 = 0.18\, r_{H,o}$$

$$(t_o - t_e) = t_{H,o} \int -(1 + z)^{-5/2}\, dz \;\; (z \text{ to } 0) = 2/3\, t_{H,o} [1 - 7^{-3/2}] = 0.63\, t_{H,o}$$

The QSO was $0.18\, r_{H,o}$ when the light set out; it is now $1.24\, r_{H,o}$; the light travelled for $0.63\, t_{H,o}$.

- Pushing further: consider the CMB at z = 1000, we have

$$r(t_o) = 2\, r_{H,o} [1 - 1000^{-1/2}] = 1.94\, r_{H,o}$$

$$r(t_e) = 1.94\, r_{H,o} / 1001 = 0.0019\, r_{H,o} \quad \text{or only} \sim 8 \text{ Mpc!}$$

We are seeing the CMB when it was closer than the Virgo cluster!
That is the main reason physical scales appear so large on the CMB [sec 8b].

(b) Horizons

- Current estimates suggest $\Omega_{tot} = 1.0$ → we are in a **spatially infinite** Universe.
  Can we see all of this universe? **No!**, for **three** (somewhat overlapping) reasons:
  - Right now, galaxies beyond $r_{H,o} = c/H_o = 4.26$ Gpc are receding **faster than light**
    As a result, light is actually getting **further** from us, not closer!
  - The universe is only ~ 14 Gyr old: light cannot travel infinitely far in that time
  - Our view of the universe is inevitably restricted to **events on our past light cone**
    this is a 3-d slice within a 4-d spacetime; like a sheet in space
    There is much in the Universe not on that sheet!

- On earth, our view is limited by the **horizon**
  Hence, some of these cosmic limitations are referred to as **horizons**

(i) The Hubble Sphere: r_H (where expansion velocity = c)

- Consider point 1: the velocity distance relation, v = Hr, defines a critical radius $r_H = c/H = c\, t_H$
  Beyond this distance, objects recede faster than light and wavefronts actually get further from us
  If the expansion remained constant, then ultimately we can/cannot see objects inside/outside r_H.

- Of course, the rate of expansion changes and so, therefore, does the size of the Hubble sphere.
  For any time: $c = H(a)\, r_H(a) = H_o\, E(a)\, r_H(a)$ giving:

$$r_H(a) = r_{H,o} / E(a) = r_{H,o} / E(z)$$

or, re-expressing $r_H(a)$ in its larger current **comoving** size, $r_{o,H}(a)$, we have:

$$r_{o,H}(a) = r_H(a) / a = r_{H,o} / a\, E(a) = r_{H,o} (1 + z) / E(z)$$

In physical coordinates, r_H increases from zero at t = a = 0 (z = ∞) and continues to increase.
In comoving coordinates, $r_{o,H}$ grows/shrinks in de/ac-celerating universes.

- For the concordance model, ignoring the early period of inflation, we find for **comoving** $r_{o,H}$:
  the Hubble sphere initially grew outwards: eg at z = 9 (a = 0.1) $r_{o,H} \approx 0.61\ r_{H,o}$
  $r_{o,H}$ then grew to a maximum size when $\Omega_m = \Omega_v$ at a = 0.72, when $r_{o,H} = 1.15\ r_{H,o} \approx 4.9$ Gpc
  Since then, $r_{o,H}$ has been shrinking and is currently at $r_{o,H}(a{=}1) = r_{H,o} = 4.26\ h_{72}^{-1}$ Gpc
  It is currently shrinking at a rate $dr_{o,H}/da \times da/dt = H_o\ dr_H/da = 0.18$ Mpc/Myr or about 0.6 c
  Objects currently at the Hubble sphere are not, of course, visible to us (nor will they ever be).

(ii) The Particle Horizon: $r_{p\text{-}hor}$ (furthest currently visible)

- Consider point 2 above, and ask: what is the furthest object **currently visible** in the sky?
  Its light has travelled non-stop **since the big bang**, so that $t_e = 0$ and LBT = $(t_o - t_e) = t_{age}$
  We want the **comoving** version of this distance, $r(t_o)$.
  From the above [sec 7a] relations, just take limits t (0 to t_o) or a (0 to 1) or z (∞ to 0):

$$r_{p\text{-}hor}(t_o) = c \int dt / a(t) \quad (0 \text{ to } t_o) = r_{H,o} \int da / [a^2E(a)] \quad (0 \text{ to } 1) = r_{H,o} \int -dz / E(z) \quad (\infty \text{ to } 0)$$

- For non-zero Ω_m or Ω_r, the integral is **finite**, and there is indeed a most distant object visible.
  For example, take the Einstein-de Sitter case (flat, matter only): $E(z) = (1 + z)^{3/2}$, giving:

$$r_{p\text{-}hor}(t_o) = r_{H,o} \int -(1 + z)^{-3/2}\, dz \quad (\infty \text{ to } 0) = 2\ r_{H,o}$$

  so, our three proper distances are: $r(t_e) = 0$; $r_{lbt} = c\ t_o = 2/3\ r_{H,o}$; $r(t_o) = r_{p\text{-}hor}(t_o) = 2\ r_{H,o}$

  Using the concordance values in E(z) gives $r_{p\text{-}hor} = 2.55\ r_{H,o} = 10.9\ h_{72}^{-1}$ Gpc (35.5 Gly)

- It seems puzzling that light cannot cross the "entire universe" when it has "no size" at the BB.
  But at t = 0, for non-zero Ω_m or Ω_r, **all** locations expand with **infinite speed** [sec 6cii]
  Hence, at t = 0 all points are **profoundly isolated**; even the **closest** location is outside the horizon.
  Only after some **deceleration** does light make progress and crosses comoving coordinate space.

  Only if the **initial** expansion is **zero** can light cross everywhere in the first instant.
  This occurs in two FRW cases:

  empty open (Milne; linear expansion) and
  pure vacuum (de Sitter; exponential expansion)

  In these cases, there is a period when the entire universe is causally connected
  If the conditions are right, it can establish coherent (global) properties during this time.
  As we shall see [sec 11] an early period of inflation may have done just this in our universe.

- The particle horizon is important in understanding **structure formation**.
  By defining the region **visible** at any time, it also defines the region in **causal contact**
  Even when optically opaque, gravity still travels at light speed and sets the sphere of influence.
  Thus, the particle horizon defines the largest possible self-interacting region.
  Everything on larger, "super-horizon" scales is fundamentally disconnected.

- So, how does the particle horizon size grow with time? Simply change the upper limit to t / a / z:
  In **comoving coordinates** we have for the particle horizon at any time:

$$r_{p\text{-}hor}(t) = \int c\, dt / a(t) \quad (0 \text{ to } t) = r_{H,o} \int da / a^2E(a) \quad (0 \text{ to } a) = r_{H,o} \int -dz / E(z) \quad (\infty \text{ to } z)$$

  Giving $2\ a^{1/2}\ r_{H,o}$ for Einstein-de Sitter; $a\ r_{H,o}$ for flat radiation; ∞ for flat vacuum (de Sitter)

- The time defined by $r_{p\text{-}hor} / c$ goes by another name, **conformal time**:

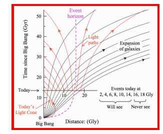
$$\eta(t) \equiv \int dt / a(t) \quad (0 \text{ to } t) \equiv \int da / a^2E(a) \quad (0 \text{ to } a) \equiv \int -dz / E(z) \quad (\infty \text{ to } z).$$

  This is a useful surrogate for time, in part because it gives the horizon size at a given t, a, z.

It is the time variable of choice for studying growth of perturbations.
Space-time diagrams (see below) can also look much simpler when plotted using conformal time.

(iii) The Event Horizon: $r_{e\text{-}hor}$ (furthest ultimately visible)

- If an event happens **now** in a distant galaxy (eg a supernova), when will we actually see it?
  If the answer is "never", then the galaxy is said to lie beyond our **event horizon** [image]
  Like the other horizons, it can change with time (and even be infinite).



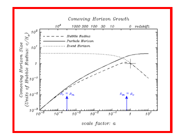
- In comoving coordinates, for any time, it is given by:

$$r_{e\text{-}hor}(t) = c\int dt / a(t) \quad (t \text{ to } \infty) = r_{H,o} \int da / a^2E(a) \quad (a \text{ to } \infty) = r_{H,o} \int -dz / E(z) \quad (z \text{ to } -1)$$

  Which is the same as $r_{p\text{-}hor}$ but with different (complementary) integration ranges
  Not surprisingly, if an FRW model has finite $r_{p\text{-}hor}$ it often has infinite $r_{e\text{-}hor}$, and visa-versa
  Einstein-de Sitter & flat radiation have $r_{e\text{-}hor} = \infty$; while flat vacuum (de Sitter) has $r_{H,o} /2a^2$
  Vacuum in the concordance model ensures we will never see remote parts of the Universe.
  Indeed, as time passes, we will see less and less as the event horizon shrinks (in comoving radius).

- The event horizon is important during vacuum driven inflation [sec 11]:
  A small causally connected region expands and is pushed outside the (shrinking) event horizon
  After inflation, we have a huge smooth super-horizon region, previously causally conntected
  As normal expansion resumes, the horizon moves out and the region **re-enters the horizon**.

  It also plays a crucial role generating and amplifying quantum fluctations.
  Ultimately, these fluctations provide the seeds for future galaxy formation [sec 11].



- The three horizons (r_H $r_{p\text{-}hor}$ and $r_{e\text{-}hor}$) for the concordance model are shown here [image].

(c) Space-Time Diagrams

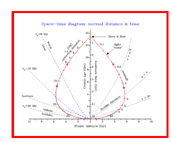
- In special relativity, **space-time diagrams** are often useful in clarifying a situation.
  They (usually) plot ct vertically and a spatial coordinate horizontally, so light rays move at 45°

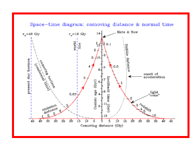
  Such diagrams are more complex on curved expanding space-times in cosmological GR:

  - locally, light cones are "tipped" to align with tilted world-lines.
  - globally, light rays can follow curved paths as they move across expanding coordinate grids.

  It is, however, often possible to clean up these diagrams & reestablish 45° light paths:

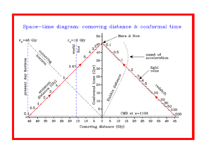
  - change the spatial axis to **comoving distance**, so objects have **vertical** world lines.
  - change the time axis from ct to $c\eta(t)$ = **conformal time**





- These [images] show space-time diagrams with light cones for the concordance model.
  There are three of them:
  - normal distance vs normal time -- light cone appears like an avocado seed
  - comoving distance vs normal time -- light cone appears like a spikey shield volcano
  - comoving distance vs conformal time -- light cone appears like a 45 degree cone.

  They are quite instructive: take a little time to figure them out.



(d) Energy within the Hubble Sphere

- Although we used **local** energy conservation to legitimately derive $\rho = \rho_o\, a^{-3(1+w)}$,
  discussion of **global** cosmological energy is much more tricky/impossible, not least because:
  the Universe is ~infinite in extent
  there is no global inertial frame
  expansion kinetic energy is ill-defined for $r > r_H$ where $v > c$
  gravitational potential energy is ill-defined with no zero-level at infinity.

  Ignoring these concerns (!), we proceed to estimate the total energy within a Hubble sphere: $r_H = c / H$.

Our aim is merely to suggest that the Universe's total energy might actually be **ZERO**

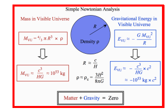
- Let's push the simple Newtonian analysis [sec 6a v] a little further:
  for a critically expanding sphere of radius R, density $\rho = 3\,H^2 / 8\pi G$ and velocity law $v = H\,r$:

  $$PE = -(3/5)\, GM^2/R = -(3/20)\, R^5 H^4 / G$$

  $$KE = ½ \int 4\pi r^2 dr\, \rho\, (H\,r)^2 = +(3/20)\, R^5 H^4 / G$$

  and we recover the Newtonian result: $E_{tot} = PE + KE = E_\infty = 0$ for all R.

- Relativistically, however, we mustn't forget the (positive) **rest mass energy**: [image]



  $$RE = M\,c^2 = (4/3)\pi R^3 \rho\, c^2 = (1/2)\, R^3 H^2 c^2 / G = (1/2)\, R^3 r_H^2 H^4 / G$$

  Hence: $|PE| / RE \approx (R / r_H)^2$ so that on small scales rest mass utterly dominates cosmic energy.

  However, when $R \rightarrow r_H$ the -ve gravitational energy grows to match the +ve rest mass energy.
  e.g. for $R \sim r_{H,o}$ we have $RE \sim |PE| \sim r_{H,o}^5\, H_o^4 / G = c^5 / H_o\, G \approx 10^{76}$ erg $\equiv 10^{22}\, M_\odot \equiv 10^{11}\, M_{gal}$
  This crudely illustrates how on cosmic scales the **total energy**, including rest mass, could be **zero**.
  In GR the condition of flat geometry is equivalent to our Newtonian "zero energy".

- A more professional analysis of inflation does indeed suggest the Universe might have zero net energy.

  Everything may have arisen from **Nothing**

  This is surely one of the most astonishing of Nature's possible properties.
  → All matter in the Universe is borrowed against the negative gravitational energy of space-time.

- Inflation is the mechanism that "splits nothing" into positive energy and equal negative binding energy.
  That is how the Universe can be made from nothing!

Next Prev Top

(8) Observables vs Redshift: A Toolkit

- Objects in the Universe have **intrinsic** properties: luminosity, size, velocity, space density
  From far far away, we witness **observed** properties: flux, angle, proper motion, etc.
  The relationships between these two are straightforward in a **static Euclidean space**
  However, the **real** Universe has an expanding possibly curved space-time:
  → how does this affect the simple Euclidean relationships?

- Well, things get a bit more complex.
  As usual, we need to keep our eye on **two** new aspects:
  - The fact that the geometry might be **curved**
  - The fact that during light's journey from object to us, the Universe has **expanded**.

  With these in mind we can derive close analogs to the Euclidean relationships.
  Indeed, one usually expresses them in **Euclidean form**, using a **pseudo-distance**, D, in place of r.
  The most famous of these are luminosity distance, D_L, and angular diameter distance, D_A
  I'll try to keep the convention of labelling pseudo-distances with capital D.

(a) Luminosity Distance

So far, we've only used the radial part of the RW-metric. Now we need to use its **angular** part.

- In a static Euclidean space, brightness f (W m^{-2}) is derived by spreading luminosity L (W) over a spherical shell with radius equal to the observer-object distance d (m): $f = L / 4\pi d^2$

- Moving to the RW-metric, we need to make **four** modifications:

- the appropriate distance is $r(t_o) = r_o = r_{H,o} \int dz / E(z)$ (z to 0), the **current** proper distance
- the surface area is reduced (increased) relative to $4\pi d^2$ for closed (open) geometries.
- the energy of each photon is **reduced** by a factor $(1 + z)$
- the rate at which photons arrive is also **reduced** by a factor $(1 + z)$

▪ Consider points 1 & 2; the RW-metric for intervals on the comoving sphere is ($dt = dr = 0$; $a = 1$):

$$ds^2 = R_o^2 S_k^2(r_o/R_o) (d\theta^2 + \sin^2\theta \, d\phi^2)$$

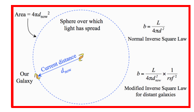
Integrating over θ and ϕ for the total spherical shell area, we get:

$$A(r) = 4\pi R_o^2 S_k^2(r_o/R_o) = 4\pi D(r_o)^2 \quad [\text{ where } D(r_o) \equiv R_o S_k(r_o/R_o)]$$

$D(r_o)$ is a **comoving distance measure** and is our first pseudo-distance:
think of $a \times D$ as giving the **correct** $d\Omega$ if we placed physical area dA at proper distance $a \times r_o$

$$dA(r,t) = [a(t) \times D(r_o)]^2 d\Omega$$

it is smaller (larger) than the proper distance, $a \times r_o$, for a closed (open) geometry.
for flat ($k = 0$) geometry, $D(r_o) = r_o$ and spherical shells have Euclidean area: $A = 4\pi D^2$ [image]



▪ For flux, we set $a = 1$ and include the redshift factors (points 3 & 4 above):

$f = L / 4\pi D^2 \times 1 / (1 + z)^2 = L / 4\pi D_L^2 \quad [\text{ where } D_L \equiv (1 + z) D]$

D_L is the **luminosity distance** and is our second pseudo distance
it gives the **correct** (bolometric) luminosity using the Euclidean formula.

▪ Making the equations more explicit:

$D(r_o) = R_o S_k(r_o/R_o)$	effective angular comoving distance
$r_o(z) = r_{H,o} \int -dz / E(z)$ (from z to 0)	true comoving proper distance
$S_k(x) = \sin(x)$ x $\sinh(x)$ ($k = +1$ 0 -1)	corrects for curvature
$R_o = r_{H,o} / \lvert \Omega_{k,o} \rvert^{1/2}$	the curvature radius [sec 6b iii]
$\Omega_{k,o} = 1 - \Omega_{t,o}$	the curvature parameter

Notice that for $k = 0$, $D = r_o$ and apart from the $(1 + z)^{-2}$ term, we recover the Euclidean relation.
Likewise, for $r_o << R_o$ and $z << 1$, we recover the full Euclidean relation.
Here is a figure showing $D_L(z)$ for several world models [image]

▪ There is an important detail we've ignored: the above relation works for **bolometric** fluxes.
In practice observations are usually in a spectral **band**, which introduces **two** additional effects:
- the bandwidth is stretched (compressed) in wavelength (frequency) by a factor $(1 + z)$
- the rest frame spectral region is bluer than the bandpass, by a facor $(1 + z)$; (**K correction**)

which lead to the following modifed relations:

$$f_\lambda = L_\lambda / 4\pi D_L^2 \times [L_\lambda(\lambda(1 + z)^{-1}) / L_\lambda(\lambda)] \times (1 + z)^{-1} \quad \text{(units: erg/s/cm}^2\text{/A)}$$

$$f_\nu = L_\nu / 4\pi D_L^2 \times [L_\nu(\nu(1 + z)) / L_\nu(\nu)] \times (1 + z) \quad \text{(units: erg/s/cm}^2\text{/Hz)}$$

$$\lambda f_\lambda = \lambda(1 + z)^{-1} L_\lambda(\lambda(1 + z)^{-1}) / 4\pi D_L^2 \quad \text{(units: erg/s/cm}^2\text{)}$$

$$\nu f_\nu = \nu(1 + z) L_\nu(\nu(1 + z)) / 4\pi D_L^2 \quad \text{(units: erg/s/cm}^2\text{)}$$

Usually, continuum fluxes require these relations, while emission lines are bolometric.

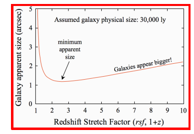
(b) Angular Diameter Distance

- In a static Euclidean space, an object at distance d of size ds subtends and angle $d\psi = ds / d$
  In moving to the RW-metric, we have **two** modifications
  - the object was **closer** when the light set out
  - curvature affects the linear distance swept out by an angle

- All rays we see have been travelling on **radial** null geodesics ($d\theta = d\phi = dt = 0$)
  for a proper transverse length, ds, which subtends angle $d\psi$, the RW-metric gives:

  $$ds \;=\; a(t_e)\, R_o\, S_k(r_o/R_o)\, d\psi \;=\; a(t_e)\, D(r_o)\, d\psi \;=\; D(r_o) / (1 + z)\;\; d\psi$$

  note $a(t_e)$ is included since the light ray geodesics start at emission time, t_e [image]



- Thus, we have for the angular diameter:

  $$d\psi = ds / a(t_e)\, D = ds / D_A \quad [\text{ where } D_A \equiv D / (1 + z) \;=\; D_L / (1 + z)^2 \,]$$

  D_A is the **angular diameter distance** and is our third pseudo distance
  it gives the **correct** angular diameter using the Euclidean formula.

- $D_A(z)$ curves are shown here [image BR 7.4]: they are **famous** for turning over
  → place galaxies further away: they first look smaller, then stay the same, then look **bigger**!
  the reason they look bigger is **not** due to curvature, but their **proximity** when the light set out.

- The above analysis applies to an object of specific physical size, ds.
  How does this change for a (large) comoving size dS which **expands with the universe**?
  For example, how big would a 100 Mpc SDSS supercluster appear at z = 0.2, 5, 1000?

  A comoving (ie current) size dS was **smaller** at redshift z: $ds = dS/(1 + z)$.
  Using this in the relation $d\psi = ds / D_A$ we get:

  $$d\psi \;=\; dS/(1 + z) / [D/(1 + z)] \;=\; dS/D \;=\; dS/D_{EA} \quad [\text{ where } D_{EA} \equiv D \,]$$

  $D_{EA} = D$ is the angular diameter distance for an object expanding with the Hubble flow.

  Let's do our example of 100 Mpc at z = 0.2, 5, 1000, choosing Einstein-de Sitter (flat, matter):
  for this, we have $D = r_o = r_{H,o} \int dz / (1 + z)^{3/2} = 2\, d_{H,o}\, [1 - (1 + z)^{-½}] \;=\; [0.17, 1.18, 1.94] \times r_{H,o}$
  Using $r_{H,o} = 4.26$ Gpc, we have $d\psi = 7.9^o,\; 1.14^o,\; 0.69^o$ for z = 0.2, 5, 1000.
  These angular sizes **don't** get bigger at high z, because our object was smaller back then.
  Notice that our supercluster is ~8° in the SDSS, it is still ~1° at z = 5 and ~0.7° on the CMB

(c) Proper Motion Distance

- Consider an object with **transverse velocity**, v_t
  the object takes time dt' to travel $ds = v_t\, dt'$, but we witness this as time $dt = dt'(1 + z)$.
  Combining the angular diameter and time dilation relations, the **observed proper motion** is:

  $$d\psi/dt \;=\; ds/D_A / dt'\,(1 + z) \;=\; v_t / D \;=\; v_t / D_M \quad [\text{ where } D_M \equiv D \,]$$

  $D_M = D$ is the **proper motion distance**, and is seen to be our original pseudo distance, D
  it gives the **correct** transverse velocity from a proper motion, assuming the Euclidean relation
  This is the appropriate distance to use when measuring projected jet speeds in radio galaxies.

  (note, David Hogg's classic "cheat sheet" uses D_M, D_C and D_H for my D, r_o and $r_{H,o}$)

(d) Surface Brightness

- In a static Euclidean space, surface brightness is famously independent of distance:

  $$SB \;=\; f / (d\psi)^2 \;=\; L/(4\pi d^2) / (s/d)^2 \;=\; L / 4\pi s^2 \quad (\text{independent of d})$$

Moving to the RW-metric: f **decreases** $\propto (1+z)^{-2}$, and $(d\psi)^2$ **increases** $\propto (1+z)^2$

$$SB = f/(d\psi)^2 = L/(4\pi D_L^2)/(s/D_A)^2 = L/4\pi s^2 (1+z)^{-4}$$

This is the (almost) equally famous $(1+z)^{-4}$ rapid drop in surface brightness with redshift.
note it is independent of curvature and, of course, is Euclidean at low-z.

- The relations for continuum fluxes are: $SB_\nu \propto (1+z)^{-3}$, and $SB_\lambda \propto (1+z)^{-5}$
- the steep drop in surface brightness with redshift is a mixed blessing:
  it severely hinders observation of high redshift galaxies, however
  without it we would all be dead.
  $(1+z)^{-4}$ takes the utterly lethal black body radiation at recombination and renders it harmless.

(e) Cosmic Volumes

- As we look out to a given redshift, the volume witnessed increases.
  Let's first consider the **comoving** volume, ie the volume it will expand to today.
  How does this depend on redshift?

  The proper comoving area at comoving distance r_o is $A(r_o) = d\Omega\, D^2$, where $D = R_o S_k(r_o/R_o)$.
  The proper comoving volume of a shell of depth dr_o is $dV_C = A(r_o)\, dr_o = d\Omega\, D^2\, dr_o$
  Now, since $r_o = r_{H,o} \int dz/E(z)$, then we have $dr_o = r_{H,o}\, dz/E(z)$ which gives:

  $dV_C = d\Omega\, r_{H,o}\, D^2\, dz/E(z)$, which can be integrated between redshifts:

$$V_C(z_1 \text{ to } z_2) = d\Omega\, r_{H,o} \int -D(z)^2\, dz/E(z) \quad (\text{from } z_1 \text{ to } z_2)$$

  Notice that D(z) is itself an integral, since $D = R_o S_k(r_o/R_o)$ and $r_o = r_{H,o} \int -dz/E(z)$ (z to 0)

  The total ($d\Omega = 4\pi$) comoving volume out to redshift z turns out to be:

  $$V_C(z) = 2\pi\, r_{H,o}^3/\Omega_{k,o} \left[D/r_{H,o}(1 + \Omega_{k,o} D^2/r_{H,o}^2)^{1/2} - |\Omega_{k,o}|^{-1/2} AS_k(\Omega_{k,o}^{1/2} D/r_{H,o}) \right]$$

  where $AS_k(x) = \arcsin(x)$ & $\text{arcsinh}(x)$ for k = +1 & -1, and $V_C(z) = 4/3\, \pi D^3$ for k = 0

- Let's now consider the **actual** (not comoving) volume, which is much less as z increases
  Repeating the analysis, but keeping non-comoving quantities, we have $A(r,a) = d\Omega\, a^2 D^2$
  dr_o becomes $a\, dr_o = a\, r_{H,o}\, dz/E(z)$ giving $dV = d\Omega\, r_{H,o}\, D^2 (1+z)^{-3} dz/E(z) = dV_C (1+z)^{-3}$
  which can also be integrated to yield a significantly smaller volume.

(f) Summary of Relations

- Let's gather many of these relations together.
  Imagine observing an object at redshift z, with (bolometric) luminosity L, and physical diameter S

  First some auxiliary functions & definitions we'll need:

  $$E(z) = [\Omega_{m,o}(1+z)^3 + \Omega_{r,o}(1+z)^4 + \Omega_{v,o} + \Omega_{k,o}(1+z)^2]^{1/2}$$

  $$\Omega_{k,o} = 1 - (\Omega_{m,o} + \Omega_{r,o} + \Omega_{v,o}) = 1 - \Omega_{t,o}$$

  $$D = R_o S_k(r_o/R_o) \quad \text{with} \quad S_k(x) = \sin(x)\ ;\ x\ ;\ \sinh(x) \quad \text{for } k = +1\ ;\ 0\ ;\ -1$$

  $$R_o = r_{H,o}/|\Omega_{k,o}|^{1/2} \quad \text{with} \quad r_{H,o} = c/H_o = c\, t_{H,o}$$

  $$dr_o = c\, dt/a = r_{H,o}\, dz/E(z) = r_{H,o}\, da/a^2E(a)$$

- Proper distance at time of observing:

$r(t_o) = r_o = c \int dt/a \quad [t_e \text{ to } t_o] = r_{H,o} \int -dz/E(z) \quad [z \text{ to } 0] = r_{H,o} \int da/a^2E(a) \quad [a \text{ to } 1]$

- Proper distance at time of emission:

  $r(t_e) = a(t_e)\, r(t_o) = r(t_o) / (1 + z)$

- Look-back time (LBT):

  $(t_o - t_e) = \int dt \quad [t_e \text{ to } t_o] = t_{H,o} \int -dz/(1 + z)E(z) \quad [z \text{ to } 0] = t_{H,o} \int da/a\,E(a) \quad [a \text{ to } 1]$

- The physical and comoving Hubble spheres at redshift z:

  $r_H(z) = c / H(z) = r_{H,o} / E(z)$

  $r_{o,H}(z) = c / aH(z) = r_{H,o} (1 + z) / E(z)$

- The comoving particle horizon at time / redshift / scale factor t / z / a:

  $r_{p\text{-}hor} = c \int dt/a \quad [0 \text{ to } t] = r_{H,o} \int -dz/E(z) \quad [\infty \text{ to } z] = r_{H,o} \int da/a^2E(a) \quad [0 \text{ to } a]$

- The comoving event horizon at time / redshift / scale factor t / z / a:

  $r_{e\text{-}hor} = c \int dt/a \quad [t \text{ to } \infty] = r_{H,o} \int -dz/E(z) \quad [z \text{ to } -1] = r_{H,o} \int da/a^2E(a) \quad [a \text{ to } \infty]$

- Observe a bolometric flux ($W\ m^{-2}$):

  $f = L / 4\pi D_L^2 \qquad D_L = D(1 + z)$ is the luminosity distance

- Observe an angular diameter (radians):

  $\psi = S / D_A \qquad D_A = D/(1 + z)$ is the angular diameter distance

  $\psi_c = S_c / D$ for an object expanding with the universe, with **comoving** size S_c

- Bolometric surface brightness, SB ($W\ m^{-2}\ sr^{-1}$):

  $SB = L / S^2 \times (1 + z)^{-4}$

- For transverse velocity v_t ($km\ s^{-1}$) we observe a proper motion, PM (radians s^{-1}):

  $PM = v_t / D_M \qquad D_M = D$ is the proper motion distance.

- Comoving proper volume, dV_c (m^3) in a shell from z to z + dz in solid angle $d\Omega$:

  $dV_c = d\Omega\, D^2\, dr_o = d\Omega\, D^2\, r_{H,o}\, dz/E(z)$

  $dV = dV_c \times (1 + z)^{-3}$ is the **physical** (non-comoving) volume in the same shell

(g) Useful Charts

- I have made a number of charts and graphs which show concordance model cosmic properties
  These should help you to get a feel for the relations given above.
  In particular, they all use **linear time** since this is closest to our experience.
  To cover everything, we need three epochs:

  - Full history : Parameters and Events and Graphs
  - The first Gyr : Parameters and Events
  - The first 500 kyr : Parameters and Events

Next Prev Top

(9) Problems With the Standard Model

- The **Standard Hot Big Bang** model described so far seems remarkably cogent.
  Recall, its basic assumptions are:
  - The Cosmological Principle -- large scale homogeneity & isotropy, yielding the RW metric
  - General Relativity -- relates contents to dynamics and geometry
  - Five components with equations of state -- radiation; neutrinos; baryonic matter; dark matter; dark energy
  - An initial perturbation spectrum close to "scale invariant" -- $P(k) \propto k$.
  - Adiabatic cooling from a hot early phase
  - The Standard Model of particle physics -- QED, QCD, etc

  This framwork manages to account for an extremely wide range of observations,
  It also provides several independent estimates of its basic parameters.
  Apart from the unknown nature of dark matter & energy, the framework seems quite robust.

- However, there exist a number of **problems** with this standard picture.
  "Problems", here, does **not** refer to an error or mismatch to data.
  Rather, they are deeper concerns about the seemingly unlikely nature of the **initial conditions**
  → why was the expansion launched exactly the way it needed to be to yield our current universe?

- This section summarizes these cosmic peculiarities (6 of them)
  The section following shows how an early burst of inflation can explain how they all come about.

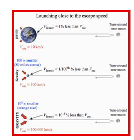
(a) The Flatness Problem

- The Universe is close to a rather special condition, which can be expressed in three equivalent ways:
  - The universe's geometry is within a few % of flat (zero curvature).
  - Its density is within a few % of the critical density.
  - Its expansion speed is within a few % of the (Newtonian) escape speed.

- As we'll see, the standard model rapidly evolves **away** from this special state, not towards it.
  Having expanded over **many** factors of 10, its early state must have been exceedingly close to the special case.
  In the absence of an explanation for this fine tuning, we seem to have a "problem".

(i) Newtonian Escape Speed

- Return to our Newtonian sphere of rocks expanding close to the escape speed [see sec 6 a iv].
  Recast this as launching a rock upwards from the Earth's surface [image].
  - At 1% below the escape speed it just reaches the moon before turning around
  - If the Earth were 100× smaller, you need to launch at $1/100^{th}$ % below v_{esc} to just reach the moon.
    this is equivalent to going to earlier times with a = 1/100, or z = 101, or t = 17.4 Myr.
  - If the Earth were 10^8× smaller, you need to launch at 10^{-8} % below v_{esc} to just reach the moon.
    this is 1 cm/s slower than 100,000 km/s, and matches cosmic expansion at t = 40 mins.



- Let's redo this example analytically:
  If our rock starts at r_{in} moving at v_{in} and turns around high up at r_{turn}, then we have for v_{in} and v_{esc}:
  $\frac{1}{2} v_{in}^2 - GM/r_{in} = E_\infty = -GM/r_{turn}$ and $\frac{1}{2} v_{esc}^2 = GM/r_{in}$

  Hence, $\frac{1}{2} v_{in}^2 = GM\,(1/r_{in} - 1/r_{turn})$ giving $(v_{in}/v_{esc})^2 = (1/r_{in} - 1/r_{turn}) / (1/r_{in}) \approx 1 - r_{in}/r_{turn}$

  So, in order to reach r_{turn}, it must start within Δv of v_{esc} where: $\Delta v / v_{esc} \approx \frac{1}{2}\, r_{in}/r_{turn}$
  As r_{in} gets smaller and smaller, the initial velocity must get closer and closer to v_{esc}.

- This is the Newtonian analog of the "flatness problem":
  → the early Universe is in a very deep and steep gravitational pit.
  → in order "get up high" (to become so "BIG"), it must have been launched **incredibly close to v_{esc}**.

(ii) Curvature/Omega Evolution

- This topic is usually cast in terms of **curvature**, which is of course related to **density**.

Measurements suggest the current total density, $\Omega_{t,o}$, is within 5% of unity (equivalently, $\Omega_{k,o} < 0.05$).
Now, although an **exactly** flat universe is **always** flat, what if the current Universe is in fact **slightly curved**?
Let's derive how the curvature **changes** as the Universe evolves -- does it get more or less curved?

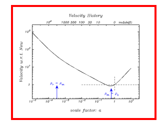
- The **current** total density parameter is $\Omega_{t,o} = \rho_{t,o} / \rho_{c,o} = 8\pi G\rho_{t,o} / 3H_o^2$
  and its deviation from 1 specifies the spatial curvature: $\Omega_{t,o} - 1 = -\Omega_{k,o} = kc^2/R_o^2H_o^2$ [see sec 6b iii]
  At epochs other than the present: $\Omega_t(a) = 8\pi G\rho_t / 3H^2$, where H and ρ_t are now both functions of a.
  Let's substitute: $H(a) = H_oE(a)$ and $\rho_t / \rho_{c,o} = \Omega_{m,o}a^{-3} + \Omega_{r,o}a^{-4} + \Omega_{v,o} = E^2(a) - \Omega_{k,o}a^{-2}$

  $$\Omega_t(a) = 8\pi G\rho_{c,o} (\rho_t / \rho_{c,o}) / 3H_o^2E^2(a) = [E^2(a) - \Omega_{k,o}a^{-2}] / E^2(a) = 1 - \Omega_{k,o} / a^2E^2(a)$$

  which is the relation we need.

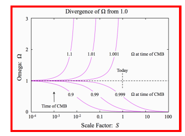
- More specifically, since $a^2E^2(a) = \Omega_{m,o}a^{-1} + \Omega_{r,o}a^{-2} + \Omega_{v,o}a^2 + \Omega_{k,o}$
  then for any currently non-flat universe ($\Omega_{k,o} \neq 0$) we find **going backwards in time** (ie $a \to 0$),
  the geometry was **more flat**: $\Omega_t \to 1$, $\Omega_k \to 0$, as long as either matter or radiation dominates.

- There is a more transparent way to express this, using: $a^2E^2(a) = a^2H^2(a) / H_o^2 = [v(a) / v_o]^2$
  This is the (normalized) velocity history discussed in sec 6c iii, with its [image] shown here.
  Summarizing:



  $$\Omega_t(a) = 1 - \Omega_{k,o} / a^2E^2(a) = 1 - \Omega_{k,o} / [v(a) / v_o]^2$$

  giving a simple rule [image]. :



  - **Decelerating** expansion (matter/radiation) makes the Universe **less flat**
  - **Accelerating** expansion (vacuum) makes the Universe **more flat**

  Hence, for the standard model, the Universe was **flatter in the past**, and will get **flatter in the future**

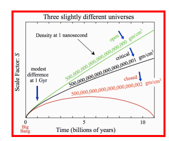
- Let's look at the past more quantitatively. In the radiation era we have [sec 6c ii] :

  $$a^2E(a)^2 \approx \Omega_{r,o} / a^2 \quad \text{and} \quad t = t_{H,o} \int da / a\,E(a)\ (0 \text{ to } a) \approx \tfrac{1}{2}\, t_{H,o}\, \Omega_{r,o}^{-1/2}\, a^2 \approx 2.33 \times 10^{19}\, h_{72}^{-1}\, a^2 \ \text{sec}.$$

  so the curvature becomes:

  $$\Omega_k(a) \approx \Omega_{k,o}\, a^2 / \Omega_{r,o} \approx 10^4\, a^2\, \Omega_{k,o} \approx 5.0 \times 10^{-16}\, t_{sec}\, h_{72}^{-1}\, \Omega_{k,o}$$

  At the time of the CMB ($a \sim 10^{-3}$) the curvature Ω_k is only ~1% of the current value, $\Omega_{k,o}$;
  at He synthesis ($a \sim 10^{-9}$) $\Omega_k \approx 10^{-14}\,\Omega_{k,o}$; and at the GUT era ($a \sim 10^{-28}$) $\Omega_k \approx 10^{-52}\,\Omega_{k,o}$! [image].
  Clearly, the fact that we are within 5% of flat today implies the early universe must have been **extremely** flat.



  Of course, inflation's early acceleration can generate just this kind of extreme flatness.

- One final potentially confusing point: although the Universe was **flatter** in the past (ie $\Omega_k \to 0$ & $\Omega_t \to 1$),
  its **curvature radius**, $R = a\,R_o$, was **smaller** in the past, and indeed $R \to 0$ as $a \to 0$!
  How do we understand this apparent contradiction? Easy: R and Ω_k depend differently on a.
  Recall [sec 6biii] the definitions of Ω_k and $\Omega_{k,o}$ which use R and R_o:

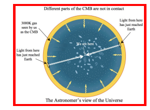
  $$\Omega_k = -kc^2 / R^2H^2 = -kc^2 / [a^2R_o^2\, H_o^2E^2(a)] = \Omega_{k,o} / a^2E^2(a)$$

  So R^2H^2 follows $a^2E^2(a)$, and although $R \to 0$ as $a \to 0$, $R \times H$ **increases** and drives Ω_k to 0.
  Of course, $a \times H(a) \propto v(a)$, the velocity history (our previous result), and we have $v \to \infty$ as $a \to 0$.

(b) The Horizon Problem

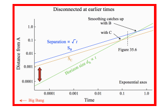
In the standard cosmology, it is puzzling why the Universe is uniform on large scales

In other words: Why does it obey the Cosmological Principle?

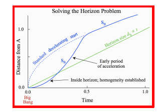


For example: the CMB is (almost) identical on opposite sides of the sky [image].
But the light from each side has only just reached us, in the middle
Hence the two sides have not yet been in contact -- so why are they the same?

Smoothing requires **communication**: things must be in contact or have a common origin to be be similar
You might think: Everything was "together/touching" at the big bang, so what's the problem?
However, in the standard cosmology, with $a \propto t^{1/2}$ we have $v \propto t^{-1/2}$ and the initial expansion ($a = 0$) is **infinitely fast**.
These [images] show how all points, no matter how close, are causally disconnected at the big bang.



In practice, smoothing can only occur at light speed: → we only expect things to be similar **within a horizon distance**.
At 400 kyr the horizon is roughly 400 kly across, which is ~1 degree on the CMB [image]
→ we only expect uniformity within 1 degree regions, which is clearly not the case at all.



The fundamental origin of the horizon problem is that the standard cosmology begins with infinite expansion and **decelerates**.
Inflation solves the problem by introducing **accelerating** expansion [image].

(c) The Monopole Problem

(d) The Structure Problem

The standard hot big bang theory has no account for the origin of structures.
They are part of the initial conditions

Two possible sources of fluctuations:
statistical fluctuations of particle numbers
quantum fluctuations in all things.

However, in a decelerating expansion, it transpires that such fluctuations **decrease** with expansion.
Basically, if the Universe is born smooth, it stays smooth, even given it's particulate/quantum nature.

Again, we need an accelerating expansion to amplify these seed fluctuations.

(e) The Existance Problem

- Look around you -- the objects, the trees, the earth, moon, sun and stars.
  Where did all that matter come from?
  As physicists, we are triply perplexed:
  (a) more than most people, we appreciate just how MUCH matter exists in the universe;
  (b) we know that matter is concentrated energy: multiply by c^2 and its a HUGE amount of energy;
  (c) we have deep respect for conservation of energy -- again: where did it all come from?

- From sec 7d, we get half the answer... appearances are misleading:
  → integrated over a Hubble sphere, the positive mass-energy is balanced by an equal negative gravitational energy
  The total energy is ZERO! In a sense, the Universe "SUMS TO NOTHING!".
  Your matter (and everything around you) is "on loan", borrowed against a huge intergalactic gravitational debt.

- Our puzzle doesn't vanish, however -- it changes to something a little more tractable:
  → **what mechanism** can start with nothing, and create arbitrary amounts of matter and gravity?
  You guessed it... inflation can.....

(f) The Expansion Problem

Why is the universe expanding?
GR tells us that any form of matter/radiation should tend to collapse
The standard hot big bang simply takes this expansion as an initial condition.
Notice that the flatness problem not only demands an initially expanding state, but the launching speed must have been both infinitely fast, and infinitely finely tuned to the (infinite) escape speed.
Thus, there is a double puzzle.

Solution: relax the condition that radiation (or matter) dominated the early universe
Instead, suggest that vacuum energy dominated
This ensures (a) the natural evolution is one of **expansion**
and (b) since it is accelerating expansion, it's launching speed will be exactly the escape speed.

Next Prev Top

(10) Inflation & Its Solutions

- As promised, we now turn to examine inflation and its implications for solving the above problems.
"Inflation" here refers to any framework that includes an early period of **accelerated expansion**.
It does **not** replace or contradict the standard hot big bang model, who's virtues all remain.
Rather, inflation is inserted at the beginning and provides a way to **launch** the standard model.

- As you are probably aware, it adds the following new features:
  - it ensures that the observable universe was once well inside a horizon,
  this allows time to establish homogeneity and isotropy, which we now observe on large scales.
  - it launched the expansion at exactly the "escape velocity", allowing the universe to become **large**.
  -- equivalently, it drove the geometry exceedingly close to flat (Euclidian)
  -- equivalently, it drove the total density exceedingly close to the critical density.
  - it amplified quantum fluctuations into classical perturbations, allowing subsequent structure formation.
  - after inflation, "reheating" provides an exceedingly hot relativistically dominated early phase
  - it dilutes to ~zero any relic particles left over from the earliest times (eg magnetic monopoles).
  - it miraculously converts **nothing** into arbitrarily large amounts of **something**
  actually, exactly equal amounts of **positive** (mass) energy and **negative** gravitational field energy.

- This field is very sophisticated, and most is beyond our (read my) expertise.
So, we'll approach this at a suitably unsophisticated level.
We'll start with the implications of acceleration, and then look at what might be driving the acceleration.

(a) Virtues of Accelerated Expansion

(i) Preliminaries

- Fortunately, we have already encountered accelerated expansion in the context of **dark energy** [see Topics 4di, 4f, 4h]
We simply need a dominant component with equation of state parameter: $w < -1/3$ ie $p < -\rho c^2/3$.
This follows directly from the Friedmann acceleration equation: $d^2a/dt^2 = -(4\pi G/3)\, a\, (\rho + 3p/c^2)$
for $w < -1/3$, d^2a/dt^2 is positive, gravity is repulsive, and expansion **accelerates**
(review [sec 4h] for an intuitive description of why this happens.)

- For a flat geometry we have [sec 6c iii] $\rho = \rho_o\, a^{-3(1+w)}$ and the Friedmann energy equation gives:

  $$da/dt = (8\pi G\rho/3)^{1/2}\, a = (8\pi G\rho_o/3)^{1/2}\, a^{-(1+3w)/2} = H_o\, a^{-(1+3w)/2}$$

  which has solutions:

  $w > -1$: $a(t) = [(3+3w)/2 \,.\, t/t_H]^{2/(3+3w)}$ power law expansion; index > 1 (acceleration) for $-1 < w < -1/3$.
  $w = -1$: $a(t) = a(t_o)\, e^{(t - t_o)/t_H}$ pure exponential expansion, with e-folding time $t_H = H^{-1}$.
  $w < -1$: $a(t) = [1 + (3+3w)/2 \,.\, (t - t_{now})/t_H]^{2/(3+3w)}$ big rip ($a \to \infty$) at $t = t_{now} + 2t_H / |3+3w|$.

  The last two solutions have no clear big bang, ie $a \to 0$ only as $t \to -\infty$

- Inflation is often assumed to be a vacuum energy, with $w \approx -1$, giving approximately exponential growth.
In this case, the number of expansion e-foldings, N, is given by (recall $e^N = 10^{N/2.3}$)

  $N = \ln (a_{end} / a_{start}) = (t_{end} - t_{start}) / t_H$ where
  $t_H = (3 / 8\pi G\rho)^{1/2} = 1.34\, \rho_6^{-1/2}$ sec (with ρ_6 in units of 10^6 gm cm^{-3})

(ii) Expansion Factor Required

(iii) Convergence to Flatness

(iv) Shrinking Event Horizon

(v) Dilution of Relics

(b) Dynamics of Scalar Fields

(c) Creating Perturbations

(d) Observational Tests of Inflation

Next Prev Top

(11) The First Three Minutes

Some useful figures: [images]

On the limits of extragalactic astronomy, so this will be brief

Next Prev Top

(12) Recombination & the CMB

Some useful figures: [images]

Some useful figures: [images]

Home Main Index Toolbox