Appendix H - About The Size Of Things

Last Update: 8 May 2007 – Needs Completing

The universe is very large. Galaxies are very large too. Stars are most often far larger than planets, and planets can be very big. What determines just how big these astronomical bodies need to be, compared with the human scale? Similarly, atoms and their nuclei are extremely small compared with the human scale. The order of magnitude of the sizes and masses of all these things follows from the values of the universal constants. In this Section we show how.

§H.1 The Observable Universe

The size of the universe is not constant, of course. The radius of the observable universe is just dictated by its age, $R_u = fct_u$, where, as we explain in the Cosmology Tutorial Chapter 5C (Cosmic geometry Part 2), f ~ 3.46 in our universe at the present epoch. At present t_u = 13.7 Byrs = 4.3 x 10^{17} s.

A time scale, t_e, can be defined using the electron mass, i.e., $t_e = \hbar / m_e c^2 = 1.28\text{x}10^{-21}\text{s}$.

Dirac noted that the current age of the universe, expressed in units of te, is of the same order as the large number defined by the reciprocal of the gravitational ‘fine structure’ constant. Specifically, to a good approximation we have,

$$\frac{t_u}{t_e} \approx \frac{2}{\alpha_G} \approx 3.4\text{x}10^{38} \qquad \text{(H.1)}$$

Of course, if G is constant, this relationship can only hold in the present epoch. (Dirac postulated that G might vary so as to ensure the above relationship holds at all times. But such large changes in G are easily discredited).

The radius of the observable universe is,

$$R_u \sim 4.5 \text{ x } 10^{26} \text{ m} = 4.7 \text{ x } 10^{10} \text{ lyrs} = 1.5 \text{ x } 10^{10} \text{ parsec} \qquad \text{(H.2)}$$

A convenient size scale is provided by the ‘size’ of the hydrogen atom ground state, a_0, given by,

$$a_0 = \frac{1}{\alpha} \cdot \frac{\hbar}{m_e c} = \frac{ct_e}{\alpha} = 0.53\text{x}10^{-10}\text{m} \qquad \text{(H.3)}$$

Using a_0 to normalise the radius of the universe, and also using (1.1), gives,

$$\frac{R_u}{a_0} = \alpha f \frac{t_u}{t_e} = 2f\frac{\alpha}{\alpha_G} = 10\frac{\alpha}{\alpha_G} = 0.85 \times 10^{37} \qquad \text{(H.4)}$$

(using f ~ 3.46).

The mass of the observable universe is a slightly slippery concept. The reason is the negative gravitational potential energy, together with the mass-energy equivalence principle, means that the total mass-energy of the universe is probably exactly zero. However, for the present purposes, we can ignore the gravitational potential energy and consider only the positive contributions to the mass-energy. Since mass is dominant in the present epoch, this amounts to the integral of the mass density over the whole observable universe. Knowing the volume, from (H.2), we therefore only need the average density. As discussed in Chapters 1, 2B and 5B,C of the Cosmology Tutorial, the mean matter density is within 1% of the critical density,

$$\rho_{crit} = \frac{3}{8\pi G t_u^2} = 9.6\text{x}10^{-27}\,\text{kg}/\text{m}^3 \qquad \text{(H.5)}$$

This excludes only the (negative) gravitational potential energy. Both dark matter and dark energy are included in the estimate of (H.5). In fact, they account for 96% of it.

Hence the mass of the observable universe is,

$$M_u \approx 4\times 10^{54}\,kg \qquad \text{(H.6)}$$

As we have discussed Chapters 5B,C of the Cosmology Tutorial, only ~4% of the mass of the universe is believed to consist of ordinary (baryonic) matter. Hence, the mass of ordinary matter in the observable universe is roughly,

$$M_u^{baryonic} \approx 1.5\times 10^{53}\,kg \qquad \text{(H.7)}$$

The total mass can be normalised by the proton mass and expressed in terms of the dimensionless universal constants as,

$$\frac{M_u}{M_p} = f^3\left(\frac{M_p}{m_e}\right)\frac{1}{\alpha_G^2} = 2\times 10^{81} \qquad \text{(1.8)}$$

§H.2 Superclusters of Galaxies

I need to add this in. Do supercluster sizes follow simply from Q? Are they just one angular degree in size?

§H.3 Galaxies

Why are galaxies typically of the order of 100kps in size, rather than, say, 1kps or 10,000kps? Why do they typically contain 10^{12} stars, rather than, say, a mere million or 10^{18}? In this Section we present a rationalisation of the size of galaxies due to Rees and Ostriker (197?), and Silk (1976), and others. However, the reader should be warned that this is not established theory. The formation of galaxies presents many challenges and is not a mature subject.

The argument considers the formation of a galaxy by the gravitational collapse of a gas cloud, initially occupying a much larger space. One should not suppose that, like a stone falling to earth, gravitational collapse of a gas cloud can occur unimpeded. The planets of our solar system are orbiting the sun quite happily without the whole system showing any immediate signs of collapsing. A comet, perhaps visiting our sun for the first time, will accelerate towards it. But in general, it will not crash into it[1] (has there ever been such an

[1] The crashing of comet Hale-Bopp into Jupiter on the ?? of ?? 20?? was notable for its rarity. Such an event has not previously been witnessed by humans – or, at least, not recorded.

event?). Rather, with its newly gained energy, it will whip around the sun and shoot away back into distant space, possibly never to visit us again. Similarly, if a cloud of gas particles were to initiate a collapse, they would start to move faster. In other words, they would become hotter.

This is simply a consequence of the conservation of energy. Recall that gravitational potential energy is negative. As the cloud contracts, the gravitational potential energy becomes larger in magnitude – but negative. In other words, the collapse causes energy to be released, and this can only appear (initially) as increased kinetic energy of the gas particles. The effect of this speeding up is to counter the collapse. One way of understanding this is by analogy with the example of the comet, above. Each particle of gas can behave in the same manner. Having gained additional kinetic energy by undergoing a small amount of contraction, the particle has enough energy…to go back where it came from! Another way of understanding it is in terms of the gas pressure. The reduced volume due to contraction, coupled with the increased temperature, causes the gas pressure to rise. Assuming that the pressure was previously such that the gas cloud was in equilibrium against gravity, it can be shown that the increased pressure must now be in excess of what is required for equilibrium. So we again conclude that the gas cloud gets pushed back whence it came.

The moral of this story is one of the most important facts in astronomy: gravitational collapse cannot happen unless energy is removed from the collapsing material. So long as the gas cloud retains its increased kinetic energy, it will frustrate the collapse process. Since this kinetic energy constitutes heat, removing it means there must be a cooling mechanism. The major problem in understanding both galaxy and star formation lies in elucidating what mechanisms provide this essential cooling.

Providing that some cooling mechanism exists, it does not matter crucially how rapid this cooling is. If the cooling is slow, then the collapse will be slow, that’s all. But the rapidity of the cooling can effect the type of collapse which occurs. If the cooling is slow and sedate, then the collapsing gas cloud will tend to retain its homogeneity as it collapses. On the other hand, if the cooling is sufficiently rapid, the collapse may become non-uniform and lead to fragmentation of the cloud into separate clouds, each one of which then proceeds to collapse towards its own centre. The time scale which discriminates between a ‘slow’ or a ‘fast’ cooling rate is the natural free-fall timescale of the cloud.

We imagine a homogeneous sphere of particles, initially at rest, which then proceed to accelerate towards their centre of mass under gravity. In this idealisation, the increasing velocity of the particles is contrived so as not to inhibit collapse, because it merely hastens the impending crunch at the centre. This problem in mechanics is easily solved to give the time for the crunch to occur, i.e., the free-fall time,

$$t_{ff} = \sqrt{\frac{3\pi}{32G\rho_0}} \tag{2.1}$$

where ρ_0 is the density of the gas prior to collapse.

Having determined t_{ff} as the time which discriminates between a fast and a slow cooling rate – what *is* the cooling rate? The assumption of Rees and Ostriker (197?), and Silk (1976), is that the dominant physical process causing cooling will be Bremsstrahlung.

This is the emission of gamma radiation due to free electrons being deflected by nuclei (which, in the primordial universe, means almost entirely hydrogen and helium nuclei). The cross section for Bremsstrahlung is derived in standard quantum field theory texts. It can be written as a product of the electron-nucleus elastic scattering cross-section (the famous Rutherford scattering formula) and a factor which accounts for the probability of gamma ray emission. In the non-relativistic limit, the Rutherford differential cross-section is,

$$\frac{d\sigma}{d\Omega} = \frac{(\alpha Z)^2}{16E_e^2 \sin^4(\theta/2)} \qquad (2.2)$$

where $E_e = mv^2/2$ is the non-relativistic electron kinetic energy, and v the electron velocity. [See, for example, Mandl and Shaw (1993), Equ.(8.95)]. Note that the total cross-section is not defined, since integration of Equ.(2.2) diverges near the forward direction.

In the case of 'soft' Bremsstrahlung, in which the gamma energies are small compared with E_e, and for non-relativistic electrons, Bjorken & Drell (1964), Equ.(7.64), gives an approximate expression for the factor which accounts for the probability that the electron-nucleus collision will emit a gamma ray. As would be expected, this is proportional to the fine structure constant, α, but also depends upon the electrons velocity, v,

$$\frac{8}{3\pi}\alpha v^2 \cdot \sin^2(\theta/2) \cdot \log\left(\frac{E_\gamma^{max}}{E_\gamma^{min}}\right) \qquad (2.3)$$

where $E_\gamma^{max}, E_\gamma^{min}$ are the maximum and minimum gamma ray energies. If a minimum gamma energy of zero were inserted in this expression it would be divergent. This is the well known "infrared catastrophe", which we will not discuss further but for which see Mandl and Shaw (1993) or Bjorken & Drell (1964). Fortunately, this divergent term will cancel in our treatment.

The product of (2.2) and (2.3) provides the Bremsstrahlung cross-section in the non-relativistic limit. Restricting attention to hydrogen we have,

$$\frac{d\sigma}{d\Omega} = \frac{1}{6\pi} \cdot \frac{\alpha^3 v^2}{E_e^2 \sin^2(\theta/2)} \cdot \log\left(\frac{E_\gamma^{max}}{E_\gamma^{min}}\right) \qquad (2.2)$$

Although the total cross-section is not strictly defined, we can integrate Equ.(2.2) up to some small angle, θ_{min}, to provide an effective cross-section for our purposes. Thus,

$$\Im = \int \frac{d\Omega}{\sin^2(\theta/2)} = 2\pi \int_{-1}^{1-\delta} \frac{-2dc}{1-c} = -4\pi \log\left(\frac{\delta}{2}\right) = -8\pi \log\left(\frac{\theta_{min}}{2}\right) \qquad (2.3)$$

Due to the logarithmic dependence of (2.3), we may choose an extremely small θ_{min} without incurring too large a penalty on the size of $\Im$. Thus, we have an effective cross-section,

$$\frac{d\sigma}{d\Omega} = \frac{\Im}{6\pi} \cdot \frac{\alpha^3 v^2}{E_e^2} \cdot \log\left(\frac{E_\gamma^{max}}{E_\gamma^{min}}\right) \tag{2.4}$$

This is the cross-section for hydrogen to produce Bremsstrahlung when subject to a flux of electrons. (We have taken the usual liberties with the units which we'll fix later). The electron flux can be expressed as the product of the free electron number density and their typical velocity, $\rho_e^N v$. (Strictly there should be an integration over the thermal distribution of electron velocities). Thus, the total number of Bremsstrahlung photons emitted per hydrogen atom per second is,

$$\text{Photons per second} = \rho_e^N v\sigma = \frac{\Im}{6\pi} \cdot \rho_e^N \frac{\alpha^3 v^3}{E_e^2} \cdot \log\left(\frac{E_\gamma^{max}}{E_\gamma^{min}}\right) \tag{2.5}$$

We want to find the cooling rate. This is the rate of loss of energy. Consequently we need to know the average energy of the emitted photons. The quantum field theoretic expressions derived in Mandl and Shaw (1993) or Bjorken & Drell (1964) show that the Bremsstrahlung photon spectrum is proportional to $1/E_\gamma$. It follows that the average photon energy is given by,

$$\langle E_\gamma \rangle = \frac{\int_{E_\gamma^{min}}^{E_\gamma^{max}} E_\gamma \cdot \frac{1}{E_\gamma} \cdot dE_\gamma}{\int_{E_\gamma^{min}}^{E_\gamma^{max}} \frac{1}{E_\gamma} \cdot dE_\gamma} = \frac{E_\gamma^{max} - E_\gamma^{min}}{\log\left(E_\gamma^{max} / E_\gamma^{min}\right)} \tag{2.6}$$

Thus, to find the total energy emitted per second as photons, we need to multiply (2.5) by (2.6), and we find that the divergent logarithmic term cancels. Thus,

$$\text{Bremsstrahlung Power} = \frac{\Im}{6\pi} \cdot \rho_e^N \frac{\alpha^3 v^3}{E_e^2} \cdot \left(E_\gamma^{max} - E_\gamma^{min}\right) \approx \frac{\Im}{6\pi} \cdot \rho_e^N \frac{\alpha^3 v^3}{E_e} \tag{2.7}$$

where we have replaced the minimum photon energy with zero, and the maximum photon energy with the electron's kinetic energy.

The cooling timescale, τ, may be defined as the time which is required for all the electron's kinetic energy to be lost by Bremsstrahlung. Hence, we can equate the Bremsstrahlung power with E_e/τ. This gives,

$$\tau = \frac{6\pi}{\Im} \cdot \frac{E_e^2}{\rho_e^N \alpha^3 v^3} = \frac{3\pi}{2\Im} \cdot \frac{m_e^2 v}{\rho_e^N \alpha^3} \tag{2.8a}$$

Inserting factors of c and $\hbar$ to get the dimensions right gives,

$$\tau = \frac{3\pi}{2\Im} \cdot \frac{v}{c^2 \rho_e^N \alpha^3} \left(\frac{m_e c}{\hbar}\right)^2 \tag{2.8b}$$

Choosing θ_{min} to be 0.01°, 0.1°, 1° or 5° gives $\Im = 235$, 177, 119 and 79 respectively. Actually, the lowest energy photons will be correlated with the lowest scattering angles, so we should not be concerned about including very small angle. Hence, choosing θ_{min} around 1° suggests that the effective value of $3\pi / 2\Im$ is ~0.04.

Finally, we can express the typical electron velocity in terms of the temperature, i.e.,

$$\frac{v}{c} \approx \sqrt{\frac{3kT}{mc^2}} \tag{2.9}$$

which gives,

$$\tau \approx 0.07 \frac{1}{c\rho_e^N \alpha^3} \left(\frac{m_e c}{\hbar} \right)^2 \sqrt{\frac{kT}{m_e c^2}} \tag{2.10a}$$

This can also be written as,

$$\tau \approx 0.07 \frac{1}{c\rho_e^N \alpha^5 a_0^2} \sqrt{\frac{kT}{m_e c^2}} \tag{2.10b}$$

Before proceeding with our object of rationalising the size of galaxies, it is of interest to evaluate this cooling timescale for representative gas cloud conditions. Stahler & Palla (2004), Table 2.2, give some typical conditions of giant galactic gas clouds. We shall ignore the neutral gas clouds, since these do not have free electrons to cause Bremsstrahlung. Stahler & Palla distinguish 'warm' and 'hot' ionised gas clouds as having the following conditions:-

Gas Cloud	ρ_N **(m^{-3})**	**T (K)**
Ionised, Warm	0.3 x 10^6	8,000
Ionised, Hot	3 x 10^3	500,000

We assume a high level of ionisation, so that the electron density is essentially the same as the hydrogen density. We find using Equ.(2.10b) that the Bremsstrahlung cooling timescale for these conditions is ~0.5 Myrs and ~400 Myrs respectively. These seem to be sensible timescales for stellar (or galactic) evolution.

We can also consider the cooling timescale associated with the universal conditions which prevail immediately after the freeze-out of hydrogen recombination at ~460,000 years. The temperature is then 2574 K, and the hydrogen number density is 1.8 x 10^8 m^{-3}. We have shown in Chapter 9b of the Tutorial that the remnant free electron density after freeze-out is ~7 x 10^{-5} times the hydrogen density, i.e. ~1.3 x 10^4 m^{-3}. Thus, the cooling timescale turns out to be ~6 Myrs. Since this is a long time compared with the doubling-time for universal expansion at this epoch (i.e. ~0.5 Myrs), we can conclude that gravitational collapse cannot happen so early in the universe's life.

We now return to our estimate of the galactic size scale. The argument is that the gravitational collapse will avoid fragmenting into smaller clouds only if the cooling timescale is longer than the free-fall timescale. But a galaxy, being a collection of stars, must have formed in a manner which permitted fragmentation, since this fragmentation is the process which leads to the formation of the individual stars. Using Equs.(2.1) and (2.10b), the required fragmentation can occur only if the cooling timescale falls below the free-fall timescale,

$$0.07\frac{1}{c\rho_e^N\alpha^5 a_0^2}\sqrt{\frac{kT}{m_e c^2}} \le \sqrt{\frac{3\pi}{32G\rho}} \qquad (2.11)$$

But the temperature can be estimated from the loss of gravitational potential, as,

$$kT \approx \frac{GMM_p}{R} \qquad (2.12)$$

Where M is the mass of the whole gas cloud, of current radius R, and MP is the proton mass. The latter is relevant, rather than the mass of the electron, because, being so much heavier, it accounts for almost all the gravitational potential energy. (The protons subsequently share their larger kinetic energy with the electrons by collisions, i.e. they come into thermal equilibrium).

In Equ.(2.11), ρ is the mass density, whereas ρ_e^N is the electron number density. So long as the ionisation fraction is high, we can therefore put, approximately, $\rho = M_p\rho_e^N$. Substitution of this and (2.12) into (2.11) followed by a little algebraic simplification, leads to the condition for fragmentation being,

$$R \le 3.8\frac{\alpha^4}{\alpha_G}\sqrt{\frac{M_P}{m_e}}\cdot a_0 = 4.2\text{x}10^{21}\,\text{m} = 135\text{kpc} \qquad (2.13)$$

Thus, the prediction is that, until a gas cloud gets within a size of radius ~135 kpc, it cannot start to fragment. When it does start to fragment, the overall size of the region occupied by the material will cease to shrink, since the collapse will now occur preferentially within the myriad of fragments – each collapsing towards their own local centres.

This is a reasonable estimate of the observed size of galaxies. Our own Galaxy has

R ~ 20 kpc. Radio galaxies tend to be larger, e.g. Cygnus A has R ~ 100 kpc. Giant radio galaxies, such as DA240 and 3C236, can have radii of order 1,000 kpc. Our own Galaxy lives within a cluster, or Local Group, with a radius of several hundred kpc. Typical galaxy clusters, if there is such a thing as 'typical', have radii of about 1,500 kpc. The Virgo cluster has a radius of ~ 6,000 kpc. The largest identified structures in the universe are the superclusters, i.e. clusters of clusters of galaxies, and have sizes of the order of 50 Mpc or so. Consequently, there is rather an overlap of size scales between galaxies and clusters of galaxies, though the superclusters are on a larger scale altogether. However, Equ.(2.13) at least serves as a crude rationalisation of the order of magnitude of the size of galaxies – or perhaps the smaller clusters of galaxies.

Rees and Ostriker (197?), and Silk (1976), and others have gone on to argue that the assumption that Bremsstrahlung dominates the cooling is correct only when the typical thermal energy exceeds one Rydberg, i.e. when $kT > 0.2\alpha^2 m_e c^2 = 13.6\text{eV}$, i.e. for temperatures higher than about 10^5 K. (Hence applicable to the 'hot' interstellar gas clouds, but not the 'warm' ones, see above). Hence, imposing the condition,

$$kT = \frac{GMM_P}{R} > 0.2\alpha^2 m_e c^2 \qquad (2.14)$$

and inserting R from (2.13) gives a minimum galactic mass of,

$$M \geq 1.9 \frac{\alpha^5}{\alpha_G^2} \sqrt{\frac{M_P}{m_e}} \cdot M_P \approx 5 \times 10^{67} M_P = 4 \times 10^{10} M_\Theta \qquad (2.15)$$

As a lower bound to the mass of galaxies this is not too bad. Typical galaxies contain around 10^{11} stars and have masses of about $10^{12} M_\Theta$, most of which is dark matter. [NB: Carr & Rees (1979) derived the same algebraic expression as our Equ.(2.15) but appear to evaluate it incorrectly, their Equ.(44) stating that the lower mass limit is $10^{12} M_\Theta$].

§H.4 Stars

A cloud of galactic gas undergoing gravitational collapse does not necessarily form a star. To do so it must be massive enough so that the temperatures and pressures reached at its centre are sufficient to ignite nuclear fusion reactions. This is what makes a star a star. On the other hand, if the gas cloud is vastly bigger, we have seen in Section 2 that it might be large enough to form a galaxy rather than a single star. So there must be an upper bound to the mass of a single star. This is determined by the radiation pressure, which tends to blow the star apart. For smaller stars, gravity ensures that the star is stable. For larger stars, the radiation pressure becomes more significant and ultimately cannot be stabilised by gravity. We will show in Chapter ? that the lower and upper bounds to a star's mass arising from the requirement that nuclear fusion occurs in a stable system can be written,

$$\left(\frac{0.367\alpha}{\alpha_G}\right)^{3/2} \left(\frac{M_p}{m_e}\right)^{3/4} M_p < M_{star} < \left[\frac{7.78}{\alpha_G}\right]^{3/2} M_p \qquad (3.1)$$

These lower and upper bounds are:-

$9 \times 10^{55}\ M_p = 1.4 \times 10^{29}$ kg $= 0.07 M_☼$ and $5 \times 10^{58}\ M_p = 8 \times 10^{31}$ kg $= 40 M_☼$ (3.2)

A typical star therefore contains about $\alpha_G^{-3/2} \sim 10^{57}$ nucleons, and no stable star differs from this by a factor of more than ~40 either way. Thus, stars come in only a rather narrow range of masses.

The radius of a star is not uniquely defined by its mass, since, towards the end of its life a star may expand greatly to become red giant, and later on a dwarf star of some sort. However, confining attention to the main sequence, when the star is stably burning its inventory of hydrogen, its density will lie in the range,

$$\left(\frac{\alpha^2 M_p c}{183.0\hbar}\right)^3 M_p < \rho < 0.00933\left[\alpha_G \frac{m_e c}{\hbar}\right]^3 \frac{M_{star}^2}{M_p} \qquad (3.3)$$

Unfortunately, this covers an enormous range, from 4 kg/m^3 to ~10^{11} kg/m^3. The latter applies only at the centre of the most massive stars. For many stars, an average density typical of terrestrial solids provides a rough approximation. Since the typical atomic size is ~a_0, this density is,

$$\rho_{solid} \sim \frac{M_p}{4a_0^3} \qquad (3.4)$$

So, a typical star with $\alpha_G^{-3/2} \sim 10^{57}$ nucleons has a size of roughly,

$$r_{star} \sim \alpha_G^{-1/2}.a_0 \qquad (3.5)$$

§H.5 Planets and Asteroids

A planet is essentially an accumulation of material which is not large enough to be a star, in other words, not large enough to ignite nuclear fusion. The largest planets tend to be composed of gas. Smaller planets tend to be solid. The smallest planets would have difficulty retaining an atmosphere due to their low gravity, and for this reason the smaller planets will tend to be solid. (Is this correct? They could retain their gas if they were cool enough. Why can't giant planets be solid?). The gaseous planets will naturally be spherical, or, more exactly, slightly oblate ellipsoids, since this is the shape dictated by gravity. What about solid planets? This provides the distinction between planets and asteroids. Planets are (virtually) spherical, whereas asteroids can take irregular, potato shapes. The distinction is actually a matter of size, since an irregular shape is insupportable in a large enough mass. Gravity will crush and distort even a solid into a spherical shape, leaving aside those minor perturbations known as mountains. The question, then, is how high can a mountain be before it is crushed by gravity? If the maximum mountain height is a substantial fraction of the body's dimensions, then we have an asteroid rather than a planet. To address this question requires a knowledge of the strength of solids. Can the strength of solids be deduced in terms of the universal constants? Well,yes and no.

Initially, assume the compressive strength of the rock comprising the planet is known, call it σ_u. Imagine for simplicity a mountain of rectangular section and of height h. The compressive stress at its base is simply $\sigma = h\rho g$, where ρ is the density of the rock and g is the acceleration due to gravity on this planet or asteroid. The latter can be written in terms of the radius of the planet, R, as $g = 4\pi GR\rho / 3$. So, the condition that the mountain avoids being crushed down to a smaller height by gravity is,

$$h < h^{MAX} = \frac{\sigma_u}{\rho g} = \frac{3\sigma_u}{4\pi GR\rho^2} \qquad (4.1)$$

In the case of earth, for which g = 9.81m/s^2, using an ultimate compressive strength of σ_u = 1000 MPa = 10^9 N/m^2, and the density of iron (7800 kg/m^3), gives a limiting mountain height of 13,000 m. Mount Everest, at roughly 9,000m, complies with this limit. It appears that the maximum supportable height is not dramatically higher – though the density we have used is probably too large, and the volume would be a factor of 2 or

3 smaller since mountains are not rectangular blocks. On the other hand, our assumed compressive strength is rather generous.

A working definition of an asteroid might be that it can support a mountain whose height is 10% of the planets radius. Equ.(4.1) then yields the upper limit to the size of an asteroid, or, equivalently, the lower limit to the size of a planet, namely,

$$R^{upper}_{asteroid} = R^{lower}_{planet} = \sqrt{\frac{30\sigma_u}{4\pi G\rho^2}} \tag{4.2}$$

Hence, again using $\sigma_u = 1000$ MPa = 10^9 N/m^2, and the density of iron (7800 kg/m^3) gives the dividing line between planets and asteroids to be a radius of ~800 km. This compares with the radius of Pluto (2,300 km) and the radius of the largest asteroid, Ceres, which is 465 km. Hence, these do indeed conform to our working definition[2].

To express (4.2) in terms of the universal constants alone, we need to find expressions for the typical density of solids and their strength. The former is easy. The mass scale of atoms is set by the proton mass, M_p, and the size scale is set by the ground state size of hydrogen, a_0, given by Equ.(1.3). Hence, the density of solids is expected to of order,

$$\rho_{solids} \approx \frac{3M_p}{4\pi}\left(\frac{\alpha m_e c}{\hbar}\right)^3 = 2680 \text{kg}/\text{m}^3 \tag{4.3}$$

[NB: The mean density of planet earth is 5511 kg/m^3]. Equ.(4.3) involves universal constants alone. To attempt to express the strength of solids in a similar manner is rather more challenging. The energy scale is set by the electron binding energy, again taking hydrogen as an exemplar, i.e., the energy scale is the Rydberg,

$$\text{energy} \sim 0.5\alpha^2 m_e c^2 = 13.6\text{eV} = 2.2\text{x}10^{-18}\text{J} \tag{4.4a}$$

Actually, the energy levels of a single electron orbiting a nucleus with charge Z are given by $E_n = Z^2\alpha^2 m_e c^2 / 2n^2$, where n is the principal quantum number. However, for the outermost electron, the other electrons which lie between it and the nucleus will shield it from all but the one excess charge. Hence, for the outermost electron in an atom we can crudely approximate the binding energy by replacing Z by 1, giving $E_n = \alpha^2 m_e c^2 / 2n^2$. The elements of most interest, comprising our planetary rocks and metals, will have outermost orbitals with n of 3 and 4, including aluminium, silicon, calcium, iron and nickel, for example. Thus, we shall assume n = 4 for our energy scale, giving a refinement to Equ.(4.4a) as follows,

$$\text{energy} \sim \frac{\alpha^2 m_e c^2}{32} = 0.85\text{eV} = 1.4\text{x}10^{-19}\text{J} \tag{4.4b}$$

Inter-atomic forces act over a distance comparable with their size. Atoms do not differ greatly in size, the heavier atoms being of similar size to the lighter atoms. This is by virtue of the inner electrons being held closer to the nucleus by the larger charge of the

[2] Unfortunately, the recent ruling of ???? does not agree. They seem determined to downgrade Pluto to a planetessimal.

heavier nuclei. Hence, the size scale of all atoms is roughly that of the hydrogen atom, a_0. There is, however, a slight effect of mass, and so we shall assume a size scale of $\sim 2a_0$ for the solid elements of interest, e.g. iron.

A typical force scale obtains when the energy scale of Equ.(4.4b) is divided by the size scale, $2a_0$, giving,

$$\text{force} \sim \frac{\alpha^3 m_e^2 c^3}{64\hbar} = 1.3\text{x}10^{-9}\,\text{N} \qquad (4.5)$$

Dividing through by the atomic cross-section, of order $4a_0^2 \sim 10^{-20}\,\text{m}^2$ gives a quantity with the dimensions of stress,

$$\text{Youngs} \sim \frac{\alpha^5 m_e^4 c^5}{256\hbar^3} \sim 10^{11}\,\text{N/m}^2 = 10^5\,\text{MPa} \qquad (4.6)$$

This is the stress required to displace an atom a distance equal to its radius – so the strain is of order unity. Consequently, this stress is a rough estimate of the expected magnitude of the elastic modulus of the material (Young’s modulus). In fact, it is quite a good estimate, structural steels having a Young’s modulus at room temperature of ~2 x 10^5 MPa, with most other solids being rather less stiff.

However, the manner in which Equ.(4.6) was derived suggests that it could equally be interpreted as the strength of the material. In truth, the strength of solids is far less than their elastic modulus, typically by a factor of between 100 and 1000, and possibly even greater. Thus, the strongest steels have ultimate tensile strengths (UTS) of ~1000 MPa, and less than half this for most structural steels. Rocks, of course, will be substantially weaker. The reasons for the shortfalls in the strengths of solids, compared with this simplistic theoretical strength, are well known. It is due to the inherent flaws in solid materials: dislocations, inclusions, grain boundaries, microcracks, multiple phases, etc. In particular, the movement of dislocations in ductile metals typically occurs when applied strains reach $\sim 10^{-3}$, and hence at stresses which are around 10^{-3} of the elastic modulus, thus defining the yield strength. In principle it is possible to determine the strength of a solid material from first principles in terms of the universal constants. In practice, the complexities of real, anisotropic, inhomogeneous, multi-grained, multi-phase, impurity bearing materials full of dislocations, disclinations, inclusions, vacancies and cavities makes this exercise impracticable. Nevertheless, a factor of ~100 to ~1000 reduction in the theoretical strength is easily rationalised as a rough estimate. Shouldn’t I include it then? We shall refer to this factor as X_{fract}. Thus we have,

$$\text{Strength} \sim \frac{1}{X_{fract}} \cdot \frac{\alpha^5 m_e^4 c^5}{256\hbar^3} \sim 10^8 - 10^9\,\text{N/m}^2 = 10^2 - 10^3\,\text{MPa} \qquad (4.7)$$

Substituting (4.7) and (4.3) into (4.2) thus allows us to express the lower limit to the planetary size in terms of the universal constants (permitting ourselves the fiddle-factor X_{fract}),

$$R_{asteroid}^{upper} = R_{planet}^{lower} = \sqrt{\frac{5\pi}{96X_{fract}} \cdot \frac{\hbar^3}{G\alpha m_e^2 M_p^2 c}} \qquad (4.8)$$

Using X_{fract} approaching 1000 reproduces the dividing line between planets and asteroids derived above, i.e., a radius of ~800 km. Equ.(4.8) can be written in terms of the dimensionless gravitational 'fine structure' constant, α_G. When normalised by a_0 this gives,

$$\frac{R^{upper}_{asteroid}}{a_0} = \frac{R^{lower}_{planet}}{a_0} = \sqrt{\frac{5\pi}{96X_{fract}} \cdot \frac{\alpha}{\alpha_G}} = 0.013\sqrt{\frac{\alpha}{\alpha_G}} \approx 10^{16} \qquad (4.9)$$

where we have used X_{fract} = 1000 in the last expression.

The demarcation between planets and asteroids can also be expressed in terms of mass, using the density given by Equ.(4.3). We find,

$$\frac{M^{upper}_{asteroid}}{M_p} = \frac{M^{lower}_{planet}}{M_p} = \left(\frac{5\pi}{96X_{fract}} \cdot \frac{\alpha}{\alpha_G}\right)^{3/2} = 2x10^{-6}\left(\frac{\alpha}{\alpha_G}\right)^{3/2} \approx 3x10^{48} \qquad (4.10)$$

where we have again used X_{fract} = 1000 in the last expression. The smallest planets (largest asteroids) therefore have a mass of around 5 x 10^{21} kg. (The earth's mass is about 1000 times this, being about 10 times the linear dimension).

§H.6 Land-Based Life Forms

Fortunately, for this exercise, we do not need to resolve what constitutes 'life'. We are only concerned with the limitations imposed by mechanical strength or mechanical processes. There is a widespread fallacy that the maximum size of land animals is determined by their static strength: if an elephant were much bigger, its legs would snap. In a sense this is approximately true, but it does not refer to static strength. We have already seen that the limit on static strength is what defines the maximum height of mountains, which is of the order of 10^4 metres on earth. Why should an elephant not be far larger, then? The distinction between mountains and animals is that mountains do not generally move around. Energetic motion is far more demanding on the strength of an animal than mere static support against gravity, as anyone who has broken a bone playing sport will testify. In particular, the sudden arrest of rapid motion, such as in a collision, is probably the most structurally demanding event. Such impacts have the capacity to induce transient stresses which are far greater in magnitude than those due simply to self weight. Animals are poorly designed in the sense that they have the capacity, under their own musculature alone, to reach running speeds such that a collision with a solid object like a large rock will do them a great deal of harm. However, as a rule of thumb, we would expect an animal to be able to sustain a standing fall without sustaining damage. This is the criterion we shall use to define the upper bound size of land based life forms.

Approximating the animal to a cube of side h, in true physicist's fashion, its gravitational potential energy is 0.5mgh. If the fracture energy per unit area is G_{fract}, the animal can sustain a fall without breaking so long as $0.5mgh < h^2 G_{fract}$. Thus, the maximum height of a reasonably robust animal is,

$$h^{max} = \sqrt{\frac{2G_{fract}}{\rho g}} \qquad (5.1)$$

where ρ is the density of the animal. Using 1000 kg/m^3 and a fracture energy of 16,000 J/m^2 gives h^{max} = 1.8 metres. Given that larger animals like horses and elephants are in danager when they fall, this is a pretty good estimate.

The value assume for the fracture energy requires some discussion. Using the energy scale defined by Equ.(4.4b), and dividing by the area scale given by $(2a_0)^2$, we get an energy per unit area of,

$$G_{fract} \sim \frac{\text{energy}}{\text{area}} \sim \frac{\alpha^4 m_e^3 c^4}{128\hbar^2} = 12\,\text{J/m}^2 \tag{5.2a}$$

Engineers use a quantity called the "toughness", which is roughly $\sqrt{YG_{fract}}$, where Y is the material's Young's modulus. Using Y ~ 10^5 MPa together with G_{fract} ~ 12 J/m^2, gives a toughness of ~1 MPa√m. This is a good estimate for the toughness of the most brittle materials, such as ceramics. However, ductile materials have far higher toughnesses. Ductile steels have toughnesses in excess of 100 MPa√m, so that their fracture energy is around 10,000 times that derived from Equ.(5.2).

The reason why ductile materials have much larger toughness than the simplistic theoretical estimate implies is essentially the same as the reason why their tensile strength is so much lower (as we saw in Section 4). The tensile strength is lowered because large deformations are possible at much smaller stresses than implied by elastic behaviour, due to mediation by dislocations. The same phenomenon, allowing large deformations to occur, results in large amounts of absorbed energy prior to fracture. Thus, the same factor, X_{fract}, which we used to reduce the tensile strength of the material may now also be used to increase its fracture energy. This is, of course, an extreme simplification (no complaints from material scientists, please!). We approximate the fracture energy as,

$$G_{fract} \sim X_{fract} \cdot \frac{\alpha^4 m_e^3 c^4}{128\hbar^2} = 12 X_{fract}\ \text{J/m}^2 \sim 12{,}000\,\text{J/m}^2 \tag{5.2b}$$

using X_{fract} ~ 1000, as before. Hence we now get an estimate of fracture strength which reproduces a reasonable maximum animal size on earth, i.e. around a couple of metres.

Hence, using Equ.(5.1) and substituting for the fracture energy from (5.2b), the density from Equ.(4.3) and also $g = 4\pi G R\rho/3$, gives,

$$\frac{h^{max}}{a_0} = \sqrt{\frac{\pi X_{fract}}{48} \cdot \frac{\hbar^2}{GRM_p^2 m_e}} \tag{5.3a}$$

If we also substitute the minimum planetary mass from Equ.(4.8), and use X_{fract} = 1000, we get,

$$\frac{h^{max}}{a_0} = \left(\frac{\pi X_{fract}^3}{120}\right)^{1/4} \left(\frac{\alpha}{\alpha_G}\right)^{1/4} = 72\left(\frac{\alpha}{\alpha_G}\right)^{1/4} = 8\text{x}10^{10} \tag{5.3b}$$

Thus, for the minimum planetary size (~800 km radius), the maximum land animal size is predicted to be ~4 metres.

The maximum animal size is simply converted to a maximum animal mass using the density, Equ.(4.3), giving,

$$\frac{M^{max}}{M_p} = \frac{3}{4\pi}\left(\frac{\pi X_{fract}^3}{120}\right)^{3/4}\left(\frac{\alpha}{\alpha_G}\right)^{3/4} = 9x10^4\left(\frac{\alpha}{\alpha_G}\right)^{3/4} \sim 10^{32} \qquad (5.4)$$

which is ~170 tonnes. We can expect our 4 metre high animal on our very low gravity small planet to be very heavy. However, the mass will have been over-estimated by virtue of the animal being assumed to be cube-shaped. Realistic volumes for a given height might be only a few percent of this. Nevertheless, we expect the answer to be of the order of a few tonnes or tens of tons.

§H.7 Fluid-Floating Life Forms (Oceans, Gas Planets)

Also covers life forms which might 'swim' in gas giant planets. To be added.

§H.8 Molecular Basis of Life

i.e. proteins, DNA, etc. To be added

§H.9 Atoms

To be added

§H.10 Atomic Nuclei

To be added

§H.11 Various Black Holes

To be added

§H.11.1 Stellar Black Holes

To be added

§H.11.2 Galactic Black Holes

To be added

§H.11.3 Exploding Black Holes

To be added

Table 1: Size and Mass of Things Normalised by a_0 and M_p

Object	Size/a_0	Mass/M_p
Observable Universe	$2f\frac{\alpha}{\alpha_G} \sim 10^{37}$	$f^3\left(\frac{M_p}{m_e}\right)\frac{1}{\alpha_G^2} \sim 2x10^{81}$
Superclusters of galaxies		
Galaxies	$3.8\frac{\alpha^4}{\alpha_G}\sqrt{\frac{M_P}{m_e}} = 8x10^{31}$	$1.9\frac{\alpha^5}{\alpha_G^2}\sqrt{\frac{M_P}{m_e}} \approx 5x10^{67}$
Galactic Black Holes		
Stars (and upper limit for planets)	$\sim \alpha_G^{-1/2} \sim 10^{19}$	$\left(\frac{0.367\alpha}{\alpha_G}\right)^{3/2}\left(\frac{M_p}{m_e}\right)^{3/4} < \frac{M_{star}}{M_p} < \left[\frac{7.78}{\alpha_G}\right]^{3/2}$ i.e. $10^{56} < \frac{M_{star}}{M_p} < 5x10^{58}$
Lower Limit for Planets = Upper Limit for Asteroids	$\sqrt{\frac{5\pi}{96X_{fract}}\cdot\frac{\alpha}{\alpha_G}} = 0.013\sqrt{\frac{\alpha}{\alpha_G}} \sim 10^{16}$	$\left(\frac{5\pi}{96X_{fract}}\cdot\frac{\alpha}{\alpha_G}\right)^{3/2} = 2x10^{-6}\left(\frac{\alpha}{\alpha_G}\right)^{3/2} \approx 3x10^{48}$
Stellar Black Holes		
Exploding Black Holes		
Land-based Life	$\left(\frac{\pi X_{fract}^3}{120}\right)^{1/4}\left(\frac{\alpha}{\alpha_G}\right)^{1/4} = 72\left(\frac{\alpha}{\alpha_G}\right)^{1/4} = 8x10^{10}$	$\frac{3}{4\pi}\left(\frac{\pi X_{fract}^3}{120}\right)^{3/4}\left(\frac{\alpha}{\alpha_G}\right)^{3/4} = 9x10^4\left(\frac{\alpha}{\alpha_G}\right)^{3/4} \sim 10^{32}$
Life in oceans or within gas planets		
Molecular Basis of Life		
Atoms		
Atomic Nuclei		

Numerical values use: f =5; X_{fract} = 1000   Need to change to f = 3.46