review article

The anthropic principle and the structure of the physical world

B. J. Carr* & M. J. Rees

Institute of Astronomy, Madingley Road, Cambridge, UK

The basic features of galaxies, stars, planets and the everyday world are essentially determined by a few microphysical constants and by the effects of gravitation. Many interrelations between different scales that at first sight seem surprising are straightforward consequences of simple physical arguments. But several aspects of our Universe—some of which seem to be prerequisites for the evolution of any form of life—depend rather delicately on apparent 'coincidences' among the physical constants.

THE structure of the physical world is manifested on many different scales, ranging from the Universe on the largest scale, down through galaxies, stars and planets, to living creatures, cells and atoms. Only objects such as quarks and leptons may be devoid of further substructure. Each level of structure requires for its description and explanation a different branch of physical theory, so it is not always appreciated how intimately they are related. We will show here that most natural scales are determined (to an order of magnitude) by just a few physical constants. In particular, the mass scale and length scale (in units of the proton mass m_P and the Bohr radius a_0) of all structures down to the atom can be expressed in terms of the electromagnetic fine structure constant, $\alpha = e^2/\hbar c$, the gravitational fine structure constant, $\alpha_G = G m_P^2/\hbar c$ and the electron-to-proton mass ratio, m_e/m_p. The quantity m_e/m_p is related to α due to a coincidence in nuclear physics. These dependences are indicated explicitly in Fig. 1.

There are several amusing relationships between the different scales. For example, the size of a planet is the geometric mean of the size of the Universe and the size of an atom; the mass of man is the geometric mean of the mass of a planet and the mass of a proton. Such relationships, as well as the basic dependences on α and α_G from which they derive, might be regarded as coincidences if one did not appreciate that they can be deduced from known physical theory.

However, one of the scales in Fig. 1, that associated with the Universe, cannot be explained directly from known physics: it is apparently a coincidence that the present age of the Universe is of the order of α_G^{-1} times the electron timescale $\hbar/m_e c^2$. This led Dirac[1] to conjecture that G decreases with time as t^{-1}, so that the two timescales are always comparable. A more metaphysical explanation by Dicke[2] is that conditions are propitious for the existence of observers only when $t \simeq \alpha_G^{-1} \hbar/m_e c^2$, so that this 'coincidence' should be of no surprise. This line of argument, which is discussed later, appeals to the 'anthropic' principle[3]. Later we shall also mention other features of the material world that seem sensitive to apparent coincidences among physical constants.

To describe structures on scales smaller than the atom one needs extra constants: in particular, the weak and strong interaction coupling constants g_W and f. It might be assumed that these are independent of α and α_G but there are some 'anthropic' interconnections between these numbers. For example, the condition that neutrinos can blow off the envelope of a star in its supernova phase will be shown to be, roughly speaking, $\alpha_G \sim \alpha_W^4$ where $\alpha_W = g m_e^2 c/\hbar^3$ is the weak fine structure constant. Supernovae are essential if the heavy elements which are (presumably) necessary for life are to spread from their production sites throughout space. The same relationship between α_G and α_W explains why the cosmological helium production is ~25% by mass. If α_W were slightly smaller or larger, the helium production would be either 100% (in which case there would never be any water, perhaps another prerequisite for life) or 0% (in which case stellar evolution would be rather different). Finally there are coincidences between f and α and the elementary particle mass ratios which may be necessary for chemistry. There may be enough independent anthropic constraints to pin down the order of magnitude of α and α_G, and also that of α_W and f. Such considerations do not provide a real physical explanation, but they may indicate why these fundamental ratios are found to have their measured values.

One other important parameter in the Universe is the entropy per baryon or, equivalently, the photon-to-baryon ratio $\mathcal{S}$, which is of the order of 10^8. The value of $\mathcal{S}$ has no explanation within the conventional hot big bang theory but it is also associated with a coincidence: if $\mathcal{S} \sim 10^8$, the matter and radiation densities are comparable at the time they thermally decouple. Although the anthropic principle does not seem to require a specific value for $\mathcal{S}$, it does require an upper limit on $\mathcal{S}$ of order $\alpha_G^{-1/4}$; otherwise, galaxies and stars would be unable to form through gravitational condensation. In fact, there are several schemes which suggest how an entropy parameter of the order of $\alpha_G^{-1/4}$ could be generated naturally.

Scales of structure in nature

We now present some simple arguments for the mass and length scales shown in Fig. 1. These are mostly straightforward consequences of simple physics. Many have been given before, but it is useful to bring them together. All our discussions will be based on order of magnitude arguments, so equations will be of the '$\sim$' rather than '$=$' kind, factors like π being neglected.

*Present address: California Institute of Technology, Pasadena, California 91109.

0028–0836/79/0278—0605$01.00

Protons: the size of a proton is taken to be the Compton wavelength associated with its rest mass:

$$r_p \sim \frac{\hbar}{m_p c} \sim 10^{-13} \quad \text{cm} \tag{1}$$

This is roughly the range of the strong interaction. Associated with r_p is the timescale $t_p = r_p/c \sim 10^{-23}$ s, the typical lifetime of a strong interaction resonance. By analogy with equation (1), the size or localisability of an electron is sometimes referred to as

$$r_e \sim \frac{\hbar}{m_e c} \sim 10^{-10} \quad \text{cm} \tag{2}$$

and we can define a corresponding electron timescale $t_e = r_e/c \sim 10^{-20}$ s.

Atoms: for a hydrogen atom the radius of the lowest energy electron orbit is

$$a_o \sim \frac{\hbar^2}{m_e e^2} \sim 10^{-8} \quad \text{cm} \sim 1 \text{ Bohr} \tag{3}$$

and the associated energy is

$$E_o \sim \frac{e^4 m_e}{\hbar^2} \sim 10 \quad \text{eV} \sim 1 \text{ Rydberg} \tag{4}$$

As one considers atoms of greater atomic number Z, the binding energy increases roughly as Z^2 and the radius of the inner shell electron orbit decreases roughly as Z^{-1}. However, the radius of the outer shell electron orbit remains of the order of a_o, so all atoms have about the same size. If one introduces the fine structure constant

$$\alpha \equiv \frac{e^2}{\hbar c} \approx \frac{1}{137} \tag{5}$$

one can write a_o and E_o as

$$a_o \sim \alpha^{-2}\left(\frac{e^2}{m_e c^2}\right) \qquad E_o \sim \alpha^2 m_e c^2 \tag{6}$$

The quantity $(e^2/m_e c^2)$ is the 'classical' radius of the electron $\sim 10^{-13}$ cm. Atomic masses span a range from 1 to roughly α^{-1} times m_p. The nuclei of atoms with atomic number exceeding α^{-1} have so much electrostatic binding energy that they are unstable to electron–positron pair production. In practice, the instability to fissioning comes in at a somewhat smaller atomic number.

The Planck scales: the only quantities of dimensions mass and length which can be constructed from G, $\hbar$ and c are the Planck scales:

$$M_{Pl} \sim \left(\frac{G}{\hbar c}\right)^{-1/2} \sim 10^{-5} \quad \text{g} \qquad R_{Pl} \sim \left(\frac{G}{\hbar c^3}\right)^{1/2} \sim 10^{-33} \quad \text{cm} \tag{7}$$

By introducing the gravitational fine structure constant,

$$\alpha_G \equiv \frac{G m_p^2}{\hbar c} \approx 5 \times 10^{-39} \tag{8}$$

these scales can be expressed as

$$M_{Pl} \sim \alpha_G^{-1/2} m_p \qquad R_{Pl} \sim \alpha_G^{1/2} r_p \tag{9}$$

So M_{Pl} is much larger than m_p but R_{Pl} is much smaller than r_p. The Planck length is the scale on which quantum gravitational fluctuations in the metric become of the order of unity, so the concept of space breaks down at such small scales. M_{Pl} can be interpreted as the mass of a black hole of radius R_{Pl}. Space may be thought of as being filled with virtual black holes of this size. Such 'instantons' have an important role in quantum gravity theory.

Black holes: the radius of a spherically symmetrical black hole of mass M is

$$R = \frac{2GM}{c^2} \sim \alpha_G \left(\frac{M}{m_p}\right) r_p \tag{10}$$

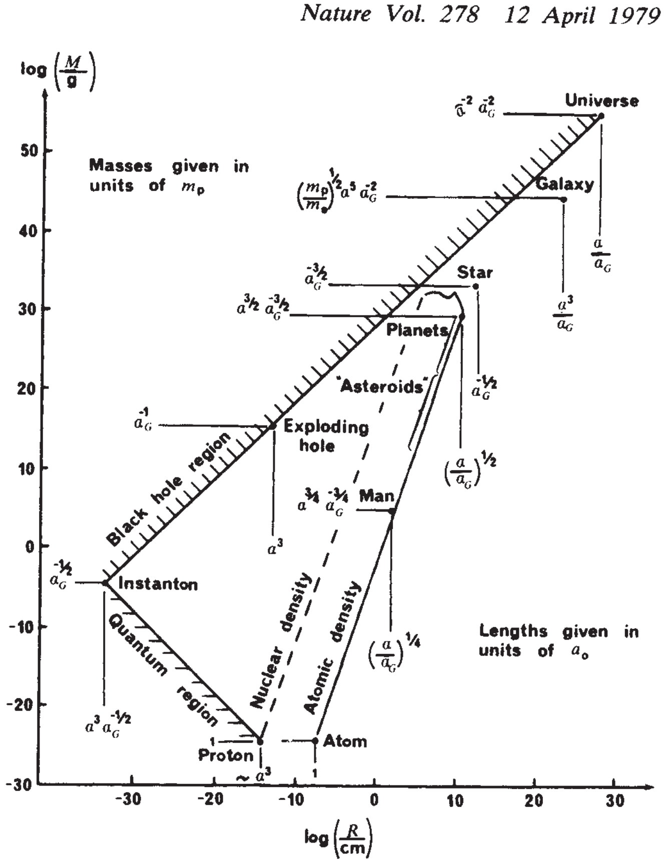

Fig. 1 The mass and length scales of various natural structures expressed in terms of the electromagnetic and gravitational fine structure constants, α and α_G. Some of these scales also depend on the electron-to-proton mass ratio, but this we have eliminated using $m_e/m_p \sim 10\alpha^2$. The asteroid scale also depend on the molecular weight A of rocky material. All these scales can be deduced directly from known physics except for the mass and length scale of the Universe, which depends on the age of the Universe being α_G^{-1} times the electron timescale $\hbar/m_e c^2$. Also shown are the atomic density line, the nuclear density line, the black hole density line and the 'quantum line' corresponding to the Compton wavelength. Most characteristic scales depend on simple powers of α_G; the wide span of so many orders of magnitudes is a consequence of the huge numerical value of α_G^{-1}, which reflects the weakness of gravity on the microscopic scale.

This is the radius of the event horizon, the region from within which nothing can ever escape (at least, classically). Black holes larger than 1 $M_\odot$ may form from the collapse of stars or dense star clusters. Smaller holes require much greater compression for their formation than could arise in the present epoch, but they might have been produced in the first instants after the big bang when the required compression could have occurred naturally. Such 'primordial' black holes could have any mass down to the Planck mass. In fact, Hawking[4] has shown that small black holes are not black at all; because of quantum effects they emit particles like a black body of temperature given by

$$k\theta \sim \frac{\hbar c^3}{GM} \sim \alpha_G^{-1}\left(\frac{M}{m_p}\right)^{-1} m_p c^2 \tag{11}$$

This means that a hole of mass M will evaporate completely in a time

$$t_{evap} \sim \alpha_G^2 \left(\frac{M}{m_p}\right)^3 t_p (N(\theta))^{-1} \tag{12}$$

$N(\theta)$ is the number of species contributing to the thermal radiation: for $k\theta \lesssim m_e c^2$ these include only photons, neutrinos and gravitons; but at higher temperatures other species may contribute. The evaporation terminates in a violent explosion. For a solar mass hole, this quantum radiance is negligible: θ is only 10^{-7} K and $t_{evap} \sim 10^{64}$ yr. But for small holes it is very important. A Planck mass hole has a temperature of 10^{32} K and only survives for a time $\sim R_{Pl}/c \sim 10^{-43}$ s. Those holes which are terminating their evaporation in the present epoch are particularly interesting. As the age of the Universe is $t_o \sim \alpha_G^{-1} t_e$, the mass of such holes would be

$$M_h \sim \alpha_G^{-2/3} \left(\frac{t_o}{t_p}\right)^{1/3} m_p \sim \alpha_G^{-1} m_p \sim 10^{15} \text{ g} \qquad (13)$$

and, from equation (10), their radius would be r_p. (The corresponding temperature is ~10 Mev: low enough to eliminate any uncertainty in $N(\theta)$ due to species of exotic heavy particles.)

Stars: Jordan[5] first noted that the Sun has a mass of about $\alpha_G^{-3/2}$ times the proton mass; most stars have a mass in the range 0.1–10 times this. Jordan constructed an elaborate cosmological theory to explain this 'coincidence', but we will now see why stars are expected to lie in this mass range[6,7]. The virial theorem implies that the gravitational binding energy of a star must be of the order of its internal energy. Its internal energy comprises the kinetic energy per particle (radiation pressure being assumed negligible for the moment) and the degeneracy energy per particle. The degeneracy energy will be associated primarily with the Fermi-momentum of the free electrons, $p \sim \hbar/d$, where d is their average separation. Provided the electrons are non-relativistic, the degeneracy energy is $p^2/2m_e$, so the virial theorem implies

$$kT + \frac{\hbar^2}{2m_e d^2} \sim \frac{GMm_p}{R} \sim \left(\frac{N}{N_o}\right)^{2/3} \frac{\hbar c}{d} \qquad (14)$$

Here N is the number of protons in the star, $N_o \equiv \alpha_G^{-3/2}$, and $R \sim N^{1/3} d$ is its radius. As a cloud collapses under gravity, equation (14) implies that, for large d, T increases as d^{-1}. For small d, however, T will reach a maximum

$$kT_{max} \sim \left(\frac{N}{N_o}\right)^{4/3} m_e c^2 \qquad (15)$$

and then decrease, reaching zero when d is

$$d_{min} \sim \left(\frac{N}{N_o}\right)^{-2/3} r_e \qquad (16)$$

A collapsing cloud becomes a star only if T_{max} is high enough for nuclear reactions to occur, that is $kT_{max} > qm_e c^2$ where q depends on the strong and electromagnetic interaction constants and is $\sim 10^{-2}$. From equation (15), one therefore needs $N > 0.1\, N_o$. Once a star has ignited, further collapse will be postponed until it has burnt all its nuclear fuel. An upper limit to the mass of a star derives from the requirement that it should not be radiation–pressure dominated. Such a star would be unstable to pulsations which would probably result in its disruption. Using the virial theorem (that is, equation (14) with the degeneracy term assumed negligible) to relate a star's temperature T to its mass $M \sim Nm_p$ and radius R, the ratio of radiation pressure to matter pressure can be shown to be

$$\frac{P_{rad}}{P_{mat}} \sim \frac{aT^4R^3}{NkT} \sim \left(\frac{N}{N_o}\right)^2 \qquad (17)$$

so the upper limit to the mass of a star is also $\sim N_o m_p$. A more careful calculation shows that there is an extra numerical factor of the order of 10, so one expects all main-sequence stars to lie in the range $0.1 \leqslant N/N_o \leqslant 10$ observed.

In other words, only assemblages of 10^{56}–10^{58} particles can turn into stable main sequence stars with H-burning cores. Less massive bodies held together by their own gravity can be supported by electron 'exclusion principle' forces at lower temperatures (they would not get hot enough to undergo nuclear fusion unless squeezed by an external pressure); heavier bodies are fragile and unstable owing to radiation pressure effects.

The M–R relationship for main-sequence stars is shown in Fig. 2, and can be deduced from the virial theorem and a knowledge of T, the central temperature. T adjusts itself so that the nuclear energy generation rate balances the luminosity, the radiant energy content divided by the photon leakage time,

$$L \simeq acT^4R^4/\kappa M \qquad (18)$$

where κ is the opacity. The appropriate κ decreases as M increases (electron scattering being the dominant opacity for upper main-sequence stars) but the energy generation increases so steeply with T that M/R depends only weakly on M.

White dwarfs and neutron stars: when a star has burnt all its nuclear fuel, it will continue to collapse according to equation (14) and, providing it is not too large, it will end up as a cold electron-degeneracy supported 'white dwarf' with the radius $R \propto M^{-1/3}$ indicated by equation (16). However, when M gets so large that $kT_{max} \sim m_e c^2$, the electrons will end up relativistic, with the degeneracy energy being pc rather than $p^2/2m_e$. Thus the degeneracy term in equation (14) acts as d^{-1} instead of d^{-2}, and consequently there is no $T = 0$ equilibrium state. From equation (15) this happens if M exceeds the mass

$$M_c \sim \alpha_G^{-3/2} m_P \sim 1\, M_\odot \qquad (19)$$

which characterises stars in general. A more precise expression for this critical value of M (Chandrasekhar mass) is $5.6\,\mu^2 M_\odot$ where μ is the number of free electrons per nucleon[8]. For a star which has burned all the way up to iron, $\mu \approx 1/2$ and the limiting Chandrasekhar mass, taking into account the onset of inverse β-decay when electrons get relativistic, is $\sim 1.25\, M_\odot$. Note that the white dwarf line in Fig. 2 meets the main-sequence line on the nuclear ignition ($T \sim 10^7$ K) isotherm. Stars bigger than M_c will collapse beyond the white dwarf density but may manage to shed some of their mass in a supernova explosion. The remanant core will comprise mainly neutrons, the electrons having been squeezed onto the protons through $p + e^- \rightarrow n + \nu$, and this core,

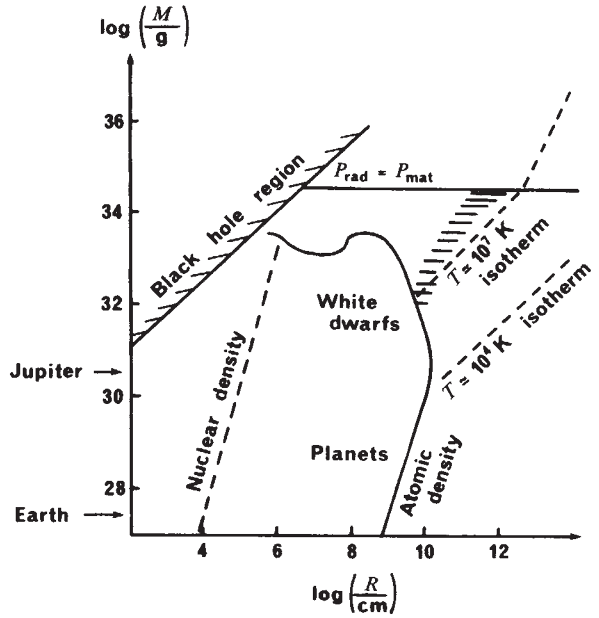

Fig. 2 An enlargement of part of Fig. 1. All solid planets lie on the atomic density line and all main sequence stars lie in the narrow mass range 0.1–10 times $\alpha_G^{-3/2} m_p$ (shaded). Regions with mass below this range would never get hot enough to ignite their nuclear fuel; regions with mass above this range would be radiation pressure dominated and consequently unstable. General relativistic instabilities are important for masses $\gg M_c$ (not shown). The precise form of the main-sequence band depends on the opacity and the nuclear reaction rate. All stars smaller than the Chandrasekhar mass ($\sim 1.4\, M_\odot$) must eventually end up on the electron-degeneracy-supported white dwarf line. Stars bigger than this either shed some of their mass in a supernova explosion and end up as a neutron star or collapse to a black hole. Objects on the dotted line which bridges the white dwarf and nuclear density lines are unstable. Also shown are the ($T \sim 10^7$ K) nuclear ignition isotherm and the ($T \sim 10^4$ K) ionisation isotherm.

if small enough, may be supported by the neutron-degeneracy pressure. The limiting mass for a neutron star is more difficult to calculate than that for a white dwarf because of strong interaction effects and because, from equation (15), with $m_e \to m_p$, the particles which dominate the neutron star's mass are relativistic. The maximum mass is still of the order of $1\,M_\odot$, however, and a neutron star bigger than this must collapse to a black hole. The maximum mass lies close to the intercept of the black hole line, given by equation (10), and the nuclear density line $\rho \sim m_p/r_p^3$. The intricacies of the line which bridges the white dwarf and neutron star regimes in Fig. 2 reflect the effects of gradual neutronisation and strong interactions. Stars on this bridge would be unstable and so are not of physical interest.

The above order-of-magnitude arguments show why the effects of radiation pressure, and relativistic degeneracy both become important for masses $\gtrsim \alpha_G^{-3/2}\, m_p$. (Note also that general relativity is unimportant for white dwarfs because the binding energy per unit mass is only $\sim(m_e/m_p)$ of c^2 at the Chandrasekhar limit.)

Planets and asteroids: the above analysis applies only to objects where the thermal energy kT and/or the degeneracy energy per electron exceeds the ionisation energy $\sim\alpha^2 m_e c^2$. If T is too small, the object will be solid or liquid rather than gaseous, in which case its structure is determined by the balance of degeneracy and chemical bonds rather than degeneracy and gravity. Since the degeneracy and electrostatic binding energy per particle are $\hbar^2 n^{2/3} m_e^{-1}$ and $e^2 n^{1/3}$, the density of any solid or liquid is of the order of the atomic density

$$\rho_o \sim e^6 m_e^3 m_p \hbar^{-6} \sim \frac{m_p}{a_o^3} \sim 1 \quad \mathrm{g\,cm^{-3}} \tag{20}$$

From equation (16), the atomic density line meets the white dwarf line ($d_{min} \sim a_o$) at a mass and radius

$$M \sim \left(\frac{\alpha}{\alpha_G}\right)^{3/2} m_p \sim \alpha^{3/2} M_c \sim 10^{30}\ \mathrm{g}$$

$$R \sim \left(\frac{\alpha}{\alpha_G}\right)^{1/2} a_o \sim 10^{10} \quad \mathrm{cm} \tag{21}$$

This point in Fig. 2 also lies on the $kT \sim \alpha^2 m_e c^2$ isotherm. Equation (21) characterises the maximum size of a planet and is of the order of the size of Jupiter. The central pressure of a self-gravitating body more massive than this can be balanced by degeneracy pressure only if the electrons are crushed closer together than in an ordinary solid. A lower limit for the size of a (solid) planet can be specified if the planet is reasonably round—that it be bigger than its own mountains. The maximum size of a mountain can be found from the following argument[7]. A mountain of height h must not provide so much pressure on the planet's surface that it liquifies its base. The condition for this is easily shown to be

$$h < h_{max} \sim \frac{E_{liq}}{A m_p\, g} \tag{22}$$

where A is the molecular weight of the planetary material, g is the surface gravity and E_{liq} is the liquefaction energy per molecule. The binding energy of a solid is about 0.1 Rydberg per molecule and E_{liq} is about a tenth of this (as only the directionality of the bonds is broken in liquefaction), so

$$E_{liq} \sim \beta\alpha^2 m_e c^2 \quad (\beta \sim 10^{-2}) \tag{23}$$

For the Earth, $g \sim 10^3\ \mathrm{cm\,s^{-2}}$ and equation (22) implies a maximum height of the order of 10 km. In general one has

$$g \sim \frac{GM}{R^2} \sim G\left(\frac{M}{m_p}\right)^{1/3} A^{2/3}(ya_o)^{-2}\, m_p \tag{24}$$

where ya_o is the molecular size of the planetary material. (A solid terrestrial-type planet is mostly SiO_2 and iron, for which $A \sim 60$ and $y \sim 4$.) Equations (22–24) thus imply

$$h_{max} \sim (\beta y^2)\left(\frac{\alpha}{\alpha_G}\right)\left(\frac{M}{m_p}\right)^{-1/3} A^{-5/3} \tag{25}$$

This is less than the radius of the planet, $R \sim ya_o(M/Am_p)^{1/3}$, provided M and R exceed

$$M \sim (y\beta)^{3/2} A^{-2}\left(\frac{\alpha}{\alpha_G}\right)^{3/2} m_p, \qquad R \sim A^{-1}\left(\frac{\alpha}{\alpha_G}\right)^{1/2} a_o (y^3\beta)^{1/2} \tag{26}$$

which is thus the maximum size of an irregularly shaped 'asteroid'. The value of A is constrained by the requirement that solid planets must be made of material whose atomic number is sufficiently high that it is not vaporised by the high temperatures attained during the formation of the Solar System.

'Habitable' planets: the mass range of habitable planets can be narrowed down still further if an atmosphere and an appropriate surface temperature are required[9]. The optimum temperature for organisms is $T \simeq \varepsilon/k$, where $\varepsilon \simeq 10^{-3}$ Rydberg is the typical energy released in relevant reactions involving complex molecules: were T much larger, the molecules would be disrupted; were it much smaller, biological processes would require an exponentially long timescale. If an atmosphere composed of gas heavier than hydrogen were necessary for life, then the thermal velocity of hydrogen at this temperature should be slightly greater than the escape velocity from the planet, which implies

$$\frac{GMm_p}{R} \sim \varepsilon\alpha^2 m_e c^2 \tag{27}$$

With the relationship $M/R^3 \sim Am_p/y^3 a_o^3 \sim m_p/a_o^3$, this implies that a life-supporting planet must have a mass

$$M \sim \varepsilon^{3/2}\left(\frac{\alpha}{\alpha_G}\right)^{3/2} m_p \tag{28}$$

Only a small fraction of planets in the mass range permitted by equations (21) and (26) would satisfy this criterion.

Press[9] has used this line of argument to set an upper limit to the size of living creatures on the basis that they must not 'break' when they fall: the energy released by a body of mass M_m in falling from height h_m on a planet whose surface gravity is given by equation (28) must be insufficient to break the molecular bonds on an area h_m^2. This yields an upper limit

$$M_m h_m\left(\frac{GM}{R^2}\right) \leqslant \left(\frac{M_m}{m_p}\right)^{2/3} \varepsilon\alpha^2 m_e c^2 \tag{29}$$

Substituting for M with equation (28) and assuming atomic density, this implies

$$h_m \sim \varepsilon^{1/4}\left(\frac{\alpha}{\alpha_G}\right)^{1/4} a_o \sim \left(\frac{\varepsilon}{0.001}\right)^{1/4} \quad \mathrm{cm} \tag{30}$$

with only a weak dependence on ε. If one assumes that the surface stresses are distributed along faults as wide as a polymer, the estimate for h_m is increased by a factor of 100. (In fact, equation (30) is more appropriate as the estimate for the size of a water droplet, which can be derived from a similar argument[7].) Allowing for the factor of 100, the mass associated with h_m is

$$M_m \sim 10^3\left(\frac{\alpha}{\alpha_G}\right)^{3/4} m_p \sim 10^5 \quad \mathrm{g} \tag{31}$$

which is roughly the mass of a man. We plot this mass in Fig. 1, although its value is very uncertain (and is in any case irrelevant to swimming organisms); note its dependence on $\alpha_G^{-3/4}$.

The Universe: in the simplest Friedmann cosmological model, the age of the Universe t_o is of the order of H_o^{-1} where H_o is the Hubble parameter. (This relation fails only if the Universe is closed and near its maximum expansion.) Since $H_o \sim 50\ \mathrm{km\,s^{-1}Mpc^{-1}}$, this implies $t_o \sim 10^{10}$ yr, a conclusion which is supported by several independent arguments. The associated horizon size (the distance travelled by light since the beginning of the Universe) satisfies the approximate relationship

$$ct_o \sim \alpha_G^{-1}\left(\frac{\hbar}{m_e c}\right) \sim \left(\frac{\alpha}{\alpha_G}\right) a_o \tag{32}$$

In other words the ratio of the size of the observable Universe to the size of an atom is comparable with the ratio of the electrical (or nuclear) and gravitational forces between elementary particles. There is no explanation for this well known coincidence within conventional physics, but Dirac[1] has conjectured that α_G is always given by

$$\alpha_G \sim \frac{\hbar}{m_e c^2 t} \sim \left(\frac{t}{t_e}\right)^{-1} \tag{33}$$

Assuming that $\hbar$, c, and m_e are constant in time, this requires that G decreases as t^{-1}, so Dirac invokes equation (32) as the basis for a new cosmology. Such a variation of G may be inconsistent with observation.

An alternative 'anthropic' explanation for equation (32) was first suggested by Dicke[2]. As life presumably requires elements heavier than hydrogen and helium and such elements can only be produced and spread throughout the Universe by supernovae, Dicke argued that any 'cognisable' Universe would be one in which some stars have already completed their main-sequence evolution. The luminosity of a star whose opacity is dominated by electron scattering (as applies for large stars) is $\sim(P_{rad}/p)L_E$; (P_{rad}/p) is given in terms of M by equation (17) and L_E is the Eddington luminosity, $4\pi\, GMm_pc/\sigma_T$, where $\sigma_T \simeq \alpha^2 r_e^2$ is the Thomson cross-section. A characteristic timescale[10] is that over which an object of luminosity L_E would radiate away its entire rest mass:

$$t_s = \frac{c\sigma_T}{4\pi G m_p} = \left(\frac{\alpha^2}{\alpha_G}\right)\left(\frac{m_p}{m_e}\right)^2 t_p \tag{34}$$

and this timescale involves the 'large number' α_G^{-1} explicitly. If $\eta \leq 10^{-2}$ is the fraction of a star's rest mass that can be released through nuclear burning, the main sequence lifetime is thus

$$t_{MS} \sim \eta\left(\frac{P_{rad}}{p}\right)^{-1} t_s \simeq \left\{\eta\alpha^2\left(\frac{m_p}{m_e}\right)^2\right\}\alpha_G^{-1}\left(\frac{M}{M_c}\right)^{-2} t_p \tag{35}$$

The quantity in braces is of the order of unity; as stars must have a mass of the order of M_c, living observers could exist only when

$$t_o > t_{MS} \sim \alpha_G^{-1} t_p \tag{36}$$

However, t_o cannot be much bigger than t_{MS} or else most of the Universe would have been processed into white dwarfs, neutron stars or black holes. Therefore observers are most likely to exist at an epoch $t_o \sim t_{MS}$, and they will find equation (32) automatically fulfilled.

In fact, t_o exceeds the value of t_{MS} in equation (36) by a factor $m_p/m_e \sim 10^3$. However, t_{MS} is sensitive to the value of M and equation (36) is really appropriate only for an upper main-sequence star. Dicke's argument would therefore be more convincing if one could show that the first stars would be of low mass. A possible reason might come from the general argument[11,12] that the mass scale at which a collapsing cloud must stop fragmenting is

$$M_{frag} \sim x^{-1/2}\left(\frac{kT}{m_p c^2}\right)^{1/4} M_c \sim 10^{-2}\, x^{-1/2} M_c \tag{37}$$

where T is the cloud's temperature (expected to be $\leq 10^4$ °K) and x is the ratio of its luminosity to that of a black body with this temperature. M_{frag} has only a weak dependence on T and is at least an order of magnitude smaller than M_C, thus increasing t_{MS} to a value more comparable to t_o. Alternatively, one might argue that the first stars do have $t_{MS} \ll t_o$ but that some of the elements vital for life must be generated through the s-processes associated with later-forming less massive stars.

The total mass associated with the observable Universe (the mass within the horizon volume) is $\sim\rho_o c^3 t_o^3$, where ρ_o is the present matter density, given by the Friedmann equation:

$$\rho_o = \frac{3H_o^2}{8\pi G} + \frac{Kc^2}{16\pi G} \tag{38}$$

Here K is the scalar curvature of the Universe. Providing the K term is smaller than the others, we deduce that the mass of the Universe is

$$M_u \sim c^3 t_o^3 G^{-1} H_o^2 \sim \frac{c^3 t_o}{G} \sim \alpha_G^{-2}\left(\frac{m_p}{m_e}\right) m_p \tag{39}$$

The fact that the number of protons in the Universe is of the order of α_G^{-2}, another cosmological coincidence, is thus explained providing one can justify neglecting the K term in equation (38). It has been argued that K must always be zero by appealing to Mach's principle[13] but, apart from this, there may be anthropic reasons for expecting that the K term is small. If K is negative, galaxies could not have condensed out from the general expansion unless $(-K)$ were less than $G\rho/c^2$ at their formation epoch. Otherwise, their gravitational binding energy would not have been large enough to halt their expansion. $(-K)$ may, in fact, exceed $G\rho/c^2$ now, but not by a large factor. If K is positive, it must be $< G\rho_o/c^2$, otherwise the whole Universe would have recollapsed before $t \simeq t_{MS}$. The simple relationship between M_u and t_o given by equation (39), with equation (36), explains why the Universe must be as big and diffuse as it is to last long enough to give rise to life. Relationships (32) and (39) also mean that the Universe has an optical depth of the order of unity to electron scattering.

Galaxies: it is not certain how galaxies form, so any estimate of their scale is very model dependent. The scale indicated in Fig. 1 derives from the following argument[14–16]. We assume that galaxies originate from overdense regions in the gaseous primordial material, and that they have a mass M and radius R_B when they become bound. After binding, protogalaxies may virialise at a radius $\sim R_B/2$. Thereafter, they will deflate on a cooling timescale, with a virial temperature

$$T \sim \frac{GMm_p}{kR} \tag{40}$$

Providing kT exceeds one Rydberg the dominant cooling mechanism is bremsstrahlung and the associated cooling timescale is

$$t_{cool} \sim \frac{1}{n\alpha\sigma_T c}\left(\frac{kT}{m_e c^2}\right)^{1/2} \simeq \frac{m_e^2 c^3}{\alpha e^4 n}\left(\frac{kT}{m_e c^2}\right)^{1/2} \tag{41}$$

The free-fall timescale is

$$t_{ff} \sim \left(\frac{GM}{R^3}\right)^{-1/2} \tag{42}$$

and this exceeds t_{cool} when R falls below a mass-independent value

$$R_g \sim \alpha^4 \alpha_G^{-1}\left(\frac{m_p}{m_e}\right)^{1/2} a_o \tag{43}$$

which, from a more precise calculation, is 75 kpc. Until a massive cloud gets within this radius it will contract quasistatically and cannot fragment into stars.

This argument applies only if the mass is so high that $kT_{virial} \gtrsim \alpha^2 m_e c^2$ at the 'magic radius' R_g: that is,

$$M \gtrsim M_g = \alpha_g^{-2}\alpha^5\left(\frac{m_p}{m_e}\right)^{1/2} \simeq 10^{12} M_\odot \tag{44}$$

Gas clouds of mass $< M_g$ cool more efficiently owing to recombination, and can never be pressure supported. Thus, M_g is a characteristic maximum galactic mass. Primordial clouds of mass $> M_g$ are inhibited from fragmentation and may remain as hot pressure-supported clouds. This type of argument can be elaborated and made more realistic[14–16]; but one still obtains a mass $\sim M_g$ above which any fluctuations are likely to remain amorphous and gaseous, and which may thus relate to the mass-scale of galaxies. M_g and R_g may thus characterise the mass and radius of a galaxy. Of all the scale estimates in Fig. 1, this is the least certain—the properties of galaxies may be a consequence of irregularities imprinted in the Universe by processes at early epochs.

This completes the justification of Fig. 1 except that we may eliminate the dependences on m_e/m_p using the relationship

$$m_e/m_p \sim 10\alpha^2 \quad (45)$$

which results from a coincidence in the nuclear physics (see equation (58)). From Fig. 1 some amusing interconnections can be deduced. Regarding the length scales, for example; man ~ (planet × atom)$^{1/2}$; planet ~ (Universe × atom)$^{1/2}$. Regarding the mass-scales: Planck ~ (exploding hole × proton)$^{1/2}$; exploding black hole ~ (Universe × proton)$^{1/2}$; man ~ (planet × proton)$^{1/2}$. These relationships cannot of course be regarded as coincidences as they can be expected *a priori*.

Other cosmological coincidences

Another dimensionless constant which characterises our Universe is the photon-to-baryon ratio $\mathcal{S} \sim 10^8 \Omega^- g1$ (where Ω is the baryon density in units of the critical density). According to the hot big bang model the background radiation dominated the density of the Universe until a time

$$t_{eq} \sim \mathcal{S}^2 \alpha_G^{-1/2} t_p \sim 10^{11}\, \Omega^{-2}\, \mathrm{s} \quad (46)$$

and thermally decoupled from the matter when T fell below $\sim 0.1\alpha^2 m_e c^2/k$ at a later time

$$t_{dec} \sim \mathcal{S}^{1/2} \alpha_G^{-1/2} \alpha^{-3} \left(\frac{m_p}{m_e}\right)^{3/2} t_p \sim 10\ \mathrm{g}1^3\, \Omega^{-1/2}\, \mathrm{s} \quad (47)$$

(These equations derive from the fact that the radiation density and temperature fall like R^{-4} and R^{-1} respectively. Here R is the length scale of the Universe, which is $\propto t^{1/2}$ for $t < t_{eq}$ and $\propto t^{2/3}$ for $t > t_{eq}$.) The observation that t_{eq} and t_{dec} are within an order of magnitude of each other (for $\Omega \sim 0.1$) corresponds to the coincidence

$$\mathcal{S} \sim 10\alpha^{-2} \left(\frac{m_p}{m_e}\right) \sim \alpha^{-4} \quad (48)$$

Now, the formation of galaxies cannot occur until, first, T has fallen below the decoupling temperature, and second, the radiation density has fallen below the matter density. Thus, the anthropic principle requires that t_o exceeds both t_{eq} and t_{dec}, these conditions being satisfied if

$$\mathcal{S} < \alpha_G^{-1/4} \left(\frac{m_p}{m_e}\right)^{1/2} \sim \alpha_G^{-1/4} \alpha^{-1} \sim 10^{11} \quad (49)$$

A lower limit on $\mathcal{S}$ is obtained if one requires that the Jeans mass in the period t_{eq} to t_{dec},

$$M_J \sim G^{-3/2} p^{3/2} \rho^{-2} c^3 \sim G^{-1} t_{eq} c^3 \sim \alpha_G^{-3/2} \mathcal{S}^2 m_p \quad (50)$$

exceed the galactic mass indicated by equation (44). Such a condition would have to be satisfied if one identified M_J with the mass of a supercluster of galaxies[12] and it implies

$$\mathcal{S} > \alpha_G^{-1/4} \alpha^{9/2} \left(\frac{m_p}{m_e}\right)^{5/4} \sim \alpha_G^{-1/4} \alpha^2 \sim 10^6 \quad (51)$$

Equations (49) and (51) constrain $\mathcal{S}$ to lie fairly close to its observed value.

The rough equality between t_{eq} and t_{dec} has prompted suggestions that the background radiation may have been generated in some way during the early history of the Universe. For example, the radiation may have been produced by a first generation of pregalactic objects[17]. Such objects have a characteristic lifetime $\sim t_{MS}$ given by equation (35), so after t_{MS} the radiation density would be expected to be ηF times the matter density, where F is the fraction of the Universe which goes into the stars and η is the fraction of each star's rest mass released through nuclear energy generation and supernova outbursts. The value of t_{eq} associated with the generated $\mathcal{S}$ is thus $(\eta F)^{3/2} t_{MS}$ and equations (35) and (46) imply, for objects with $L \simeq L_E$, that

$$\mathcal{S} \sim \alpha_G^{-1/4} \left\{ \alpha \left(\frac{m_p}{m_e}\right) F^{3/4} \eta^{5/4} \right\} \quad (52)$$

If the term in braces is of the order of unity, one would expect $\mathcal{S} \sim \alpha_G^{-1/4}$ as observed. The same relationship between $\mathcal{S}$ and α_G would apply if the radiation were generated by black holes accreting at the Eddington limit[18]. Another possibility[19,20] is that a photon–baryon ratio of the order of

$$\mathcal{S} \sim \left(\frac{G m_W^2}{\hbar c}\right)^{1/4} \quad (53)$$

where m_W is the W-boson mass, could be generated by CP violating processes in primordial black hole evaporation, or by the large fluctuations which might arise at the epoch of weak/electromagnetic interaction symmetry breaking[18]. This is just the relation $\mathcal{S} \sim \alpha_G^{-1/4}$ with m_p replaced by m_W. One could probably conceive of several cosmological scenarios in which a value of $\mathcal{S}$ of the order of $(Gm^2/\hbar c)^{-1/4}$ would be generated, where m is the mass of some elementary particle. The important point is that the coincidences involving $\mathcal{S}$ can be explained naturally without recourse to anthropic considerations. (The processes of primordial nucleosynthesis, apart from deuterium production, are insensitive to $\mathcal{S}$ provided that $\mathcal{S} \gtrsim 10^3$).

Dicke's anthropic interpretation[2] of equation (32) was advanced before the microwave background was discovered. One could perhaps replace (or strengthen) the argument leading to equation (36) by the more general statement—independent of considerations of stellar physics—that observers require $kT \leqslant 0.1\alpha^2 m_e c^2$ and the possibility of thermodynamic disequilibrium, and therefore could survey the Universe only when $t_o \gtrsim t_{dec}$. From equation (46), if $\mathcal{S} \sim \alpha_G^{-1/4}$ the condition $t_o > t_{eq}$ is the same as equation (36).

Dicke's argument helps us to understand why the 'coincidence' of equation (32) prevails—why a similar large number arises in two contexts that might at first sight seem unrelated. It does not tell us why α_G^{-1} has its particular immense value. The task of deriving such a large pure number from basic theory might seem a daunting one. However, Zeldovich[21] points out that a complete quantum theory of gravity, where tunnelling and topology changes play a key role, is likely to involve exponentials of pure numbers, so that a factor $\sim 10^{40}$ may readily arise. Although no such number has been predicted, a relationship between α and α_G may be forthcoming from quantum field theory. It has been suggested[22,23] that all space integrals in quantum electrodynamics should be cut off at the Planck length (given by equation (7)), thereby reducing otherwise divergent integrals to finite functions of the parameter ($\alpha \log \alpha_G$). Various arguments suggest that a self-consistent electrodynamics is possible only if this parameter has some specific value of the order of unity, that is, one requires

$$\alpha^{-1} \sim \log \alpha_G^{-1} \quad (54)$$

This argument may not be convincing, but empirically relation (54) is satisfied. Figure 1 mainly shows that all objects where gravity is important have masses exceeding m_p by some simple power of α_G^{-1}. As the scaling is so simple, could one envisage a hypothetical universe in which all microphysical laws were unchanged, but G was (say) a million times stronger? Planetary and stellar masses ($\propto \alpha_G^{-3/2}$) would then be lowered by 10^9; but hydrogen-burning main-sequence stars would still exist albeit with lifetimes ($t_{MS} \propto \alpha_G^{-1}$ according to equation (35)) of only ~100 yr rather than $\sim 10^8$ yr. Moreover, Dicke's argument would still apply: a hypothetical observer looking at the

Universe when $t_o \simeq t_{MS}$ would still find equation (32) fulfilled, although the number of particles in his universe would be 10^{12} times lower than in ours. Are there anthropic arguments against the 'cognisability' of the kind of small-scale speeded-up universe that corresponds to a much smaller α_G^{-1}?

One might argue that the gradual evolutionary emergence of complex organisms demands a large ratio between cosmic and microphysical timescales, and requires also that organised structures can become large (masses $\ggg m_p$) before gravity overwhelms chemical forces. This favours a very large α_G^{-1}; moreover, as $M_u/M_c \simeq \alpha_G^{-1}$, a universe where α_G^{-1} is large contains more independent sites where evolution might occur. But there is no basis here for being at all quantitative. There are, however, some more specific (though somewhat contrived) anthropic arguments that pin down α_G.

One connection between α and α_G, given by Carter[3,24], is related to the existence of convective stars. If radiation transport is unable to maintain a star's surface temperature T_S above the ionisation temperature $\sim 0.1\alpha^2 m_e c^2/k$, a convective outer layer develops and this supplements the heat transport so that it does. The value of T_S which can be maintained by radiative transport, assuming (from the Coulomb penetration condition) a central temperature $T_C \sim 10^{-2}\alpha^2 m_p a^2/k$, is

$$T_S \sim \left(\frac{L}{aR^2}\right)^{1/4} \sim (\text{optical depth through star})^{-1/4}.\ T_C \quad (55)$$

T_S exceeds $0.1\alpha^2 m_e c^2/k$ provided M exceeds

$$M_* \sim M_c \alpha_G^{-1/2}\alpha^6\left(\frac{m_e}{m_p}\right)^2 \sim M_c \alpha_G^{-1/2}\alpha^{10} \quad (56)$$

using relation (45). The mass M_* thus divides the (convective) red dwarfs from the (radiative) blue giants. This lies in the range around M_c in which main-sequence stars exist only because

$$\alpha_G \sim \alpha^{20} \quad (57)$$

Were G (and hence α_G) slightly larger, all stars would be blue giants; if it were slightly smaller, all stars would be red dwarfs. Carter ascribes anthropic significance to relation (57) on the basis that the formation of planetary systems may be associated with convective stars. (This is supported by the observation that red dwarfs have much less angular momentum than blue giants and a loss of angular momentum may be a consequence of planet formation, but this is not a compelling argument. Alternatively, convective stars may lose their angular momentum through a wind, and this wind, by blowing away the gaseous envelope of planets close to the star, facilitates the formation of solid planets with non-hydrogen atmospheres.) Carter infers that no planets, and hence no life, would form if α_G were much larger than α^{20}. If α_G were much smaller, all stars would be chemically homogeneous due to convective mixing and one would not get the 'onion-skin' shell structure which characterises pre-supernova models. Carter's convective condition[12,14] is also the condition that the number of stars in a galaxy be the same as the number of galaxies in the Universe.

The constants pertaining to nuclear physics are the strong (scalar) coupling constant $f^2 \simeq 15$ and the mass ratios $m_e/m_N \simeq 1/1837$, $m_\pi/m_N \simeq 1/7$ and $\Delta_N/m_N \simeq 1/730$ where $\Delta_N = (m_N - m_p)$. (Many more mass ratios are associated with high-energy physics but it is not clear which of them are fundamental.) The important features of nuclear physics depend[24] on the following four coincidences:

$$f^2 \simeq 2m_N/m_\pi \ (a) \qquad \Delta_N/m_e \simeq 2 \ (b) \qquad \alpha \simeq \Delta_N/m_\pi \ (c) \qquad f \simeq 1/3\alpha^{1/2} \ (d) \quad (58)$$

Equation (58a) implies that strong interactions are only marginally strong enough to bind nucleons into nuclei. If f were slightly weaker, only hydrogen could exist; if it were slightly stronger nuclei of almost unlimited size might exist. (58b) implies that neutrons are unstable to β-decay in isolation but not in the presence of relativistic degenerate electrons. (58c) implies that the electrostatic energy in light nuclei $\sim \alpha m_\pi$ is comparable to the neutron–proton mass difference. (58d) implies that the electrostatic energy is small compared with nuclear binding energy in light nuclei but comparable to it for nuclei with $Z \sim (f\alpha)^{-3} \sim 30$, so such large nuclei are unstable to electromagnetic disruption. If the relations indicated by equation (58) were not satisfied, elements vital to life would not exist, so one might also ascribe anthropic significance to these relations. Note that the combination of (58a), (b), (c) and (d) implies equation (45). Kahn[27] noted that $m_e \ll m_p$ may be a prerequisite for complex chemistry, as this ensures that the ions are located to a precision $\sim (m_e/m_p)^{1/4}$ times their mean spacing.

There may be other anthropic aspects to the value of f. For example, elements heavier than helium can form in stars through the triple-α reaction ($He^4 + He^4 \rightarrow Be^8$, $Be^8 + He^4 \rightarrow C^{12}$), and it is apparently accidental that this process yields a substantial amount of carbon[25,26]. ^8Be is unstable (otherwise the 'helium flash' in giants would lead to a catastrophic explosion); but the reaction proceeds to ^{12}C because this latter nucleus happens to have an energy level just above the sum of the energies of ^8Be and ^4He. There is, however, no similar favourably placed resonance in ^{16}O; otherwise almost all the carbon would be transmuted into oxygen. If f were even merely a few per cent larger, there would be 100% cosmological helium production because He^2 would be bound, and deuterium could form by $p + p \rightarrow {}^2He + \gamma$, ${}^2He \rightarrow D + e^+ + \nu$, even if there were no frozen-out neutrons. A small increase in f would also affect the binding energies of heavier nuclei.

Few features of the physical world seem to depend on the actual value of the weak interaction coupling constant. There are, however, two processes, both crucial to nucleosynthesis, which are sensitive to neutrino cross-sections and thus imply a coincidence of 'anthropic' relevance.

The first of these is connected with the production of helium through cosmological nucleosynthesis. The helium abundance Y is essentially determined by the temperature T_F at which the neutron–proton ratio freezes out. This occurs when the weak interactions ($p + e^- \rightarrow n + \nu$, $p + \bar{\nu} \rightarrow n + e^+$), proceeding at a rate $\propto g_W^2 T^5$, where $g_W \sim 10^{-49}$ erg cm^3 is the weak coupling constant, become slower than the cosmological expansion rate $\sim (GaT^4/c^2)^{1/2}$ at a temperature

$$kT_F \sim G^{1/6} g_W^{-2/3} \hbar^{11/6} c^{-5/6} \quad (59)$$

As virtually all the frozen-out neutrons burn into helium, the resultant helium abundance is

$$Y \sim \frac{2n_N/n_p}{n_N/n_p + 1} \quad \text{where} \quad \frac{n_N}{n_p} \sim \exp\left(\frac{-\Delta_N c^2}{kT_F}\right) \quad (60)$$

The fact that $Y \sim 25\%$ rather than 0% or 100% results only because $kT_F \sim \Delta_N c^2 \sim m_e c^2$ and, from equation (59), this derives from the 'coincidence'

$$\left(\frac{Gm_e^2}{\hbar c}\right)^{1/4} \sim \left(\frac{g_W m_e^2 c}{\hbar^3}\right) \quad (61)$$

The dimensionless number on the right may be regarded as the weak fine structure constant ($\alpha_W \sim 10^{-11}$) and the number on the left as the quarter power of the electron gravitational fine structure constant. In the Weinberg–Salam unified theory of weak and electromagnetic interactions[28] α_W is related to α by

$$\alpha_W \sim \alpha\left(\frac{m_e}{m_W}\right)^2 \quad (62)$$

So equation (61) combined with equation (45) may be written as

$$\alpha_G \sim \left(\frac{m_e}{m_W}\right)^8 \left(\frac{\alpha^2 m_p}{m_e}\right)^2 \sim \left(\frac{m_e}{m_W}\right)^8 \quad (63)$$

It is unclear to what extent these coincidences can be interpreted anthropically. Relation (63), for example, may reflect some deep-rooted connection between gravity and the weak and electromagnetic interactions. However, life could probably not exist if Y were 100% (as would be the case if α_W were slightly smaller), as there would be no water. Also, the lifetime of a helium star is less than that of a hydrogen star and might not be long enough to permit the evolution of life. It is not obvious that 0% helium production (as would arise if α_W were slightly larger) would be incompatible with life.

Equation (61) may, however, be associated with another anthropic condition (which would limit α_W in both directions)—the existence of supernovae. What actually ejects the stellar envelope in a supernova explosion is still uncertain, but it may be the outward surge of neutrinos generated by the high temperatures in the collapsing core[30]. For this model to work, one requires the timescale on which neutrinos interact with nuclei in the envelope to be comparable to the dynamical timescale. If it were much longer, the envelope would be essentially transparent to the neutrinos; if it were shorter, the neutrinos would be trapped in the core, and could not escape to deposit their momentum in the less tightly bound surrounding layers. The two timescales are comparable if

$$c^{-3}\hbar^{-4} g_W^2 n (kT)^2 \sim (Gnm_p)^{1/2} \quad (64)$$

where n is the nucleon number density. For the neutrinos to be produced at all (by $e^+ + e^- \rightarrow \nu + \bar{\nu}$) one requires kT to be of the order of $m_e c^2$. One also expects the density at the bounce to be of the order of the nucleon degeneracy density which, for $M \sim M_c$, is of the order of the nuclear density: $n \sim (\hbar/m_p c)^{-3}$. Putting these values for n and T into equation (64) gives relation (61) apart from a $(m_e/m_p)^{1/2}$ factor. Hence, the condition that stars go through a supernova phase is essentially the same as the condition that there be an interesting amount of cosmological helium production.

Conclusion

The possibility of life as we know it evolving in the Universe depends on the values of a few basic physical constants—and is in some respects remarkably sensitive to their numerical values. Indeed, the various anthropic relations quoted above in principle determine the order-of-magnitude of most of the fundamental constants of physics. Equations (54) and (57) specify α and α_G; equations (61) and (58*d*) specify α_W and f; and equations (9), (45), (58) and (63) determine m_p, m_e, m_π, m_n and m_W in terms of the Planck mass. From a physical point of view, the anthropic 'explanation' of the various coincidences in nature is unsatisfactory, in three respects. First, it is entirely *post hoc*: it has not yet been used to predict any feature of the Universe (although some people have used it to rule out various cosmological models). Second, the principle is based on what may be an unduly anthropocentric concept of an observer. The arguments invoked here assume that life requires elements heavier than hydrogen and helium, water, galaxies, and special types of stars and planets. It is conceivable that some form of intelligence could exist without all of these features—thermodynamic disequilibrium is perhaps the only prerequisite that we can demand with real conviction. Third, the anthropic principle does not explain the exact values of the various coupling constants and mass-ratios, only their order of magnitudes. With enough anthropic conditions, one may be able to be more precise about the constants of nature, but the present situation is unsatisfactory.

On the other hand, nature does exhibit remarkable coincidences and these do warrant some explanation. Apart from Dirac's suggestion[1] (which only considers the age of the Universe coincidence), the anthropic explanation is the only candidate and the discovery of every extra anthropic coincidence increases the *post hoc* evidence for it. The concept would be more palatable if it could be given a more physical foundation. Such a foundation may already exist in the Everett 'many worlds' interpretation of quantum mechanics, according to which, at each observation, the Universe branches into a number of parallel universes, each corresponding to a possible outcome of the observation[30]. The Everett picture is entirely consistent with conventional quantum mechanics; it merely bestows on it a more philosophically satisfying interpretation. There may already be room for the anthropic principle in this picture. Wheeler[31] envisages an infinite ensemble of universes, all with different coupling constants and so on. Most are 'stillborn', in that the prevailing physical laws do not allow anything interesting to happen in them; only those which start off with the right constants can ever become 'aware of themselves'. One would have achieved something if one could show that any cognisable universe had to possess some features in common with our Universe. Such an ensemble of universes could exist in the same sort of space as the Everett picture invokes. Alternatively, an observer may be required to 'collapse' the wave function[32]. These arguments go a little way towards giving the anthropic principle the status of a physical theory but only a little: it may never aspire to being much more than a philosophical curiosity. One day, we may have a more physical explanation for some of the relationships discussed here that now seem genuine coincidences. For example, the coincidence $\alpha_G \sim (m_e/m_W)^8$, which is essential for nucleogenesis, may eventually be subsumed as a consequence of some presently unformulated unified physical theory. However, even if all apparently anthropic coincidences could be explained in this way, it would still be remarkable that the relationships dictated by physical theory happened also to be those propitious for life.

We thank John Barrow, Dick Bond, Bill DeCampli, Max Delbrück, Carol Readhead, Vernon Sandberg, Frank Tipler and Steve Webber for helpful discussions. B.J.C. is grateful to the Lindemann Trust and the NSF for financial support.

1. Dirac, P. A. M. Nature **139**, 323 (1937); *Proc. R. Soc.* **165A**, 199 (1938).
2. Dicke, R. H. *Nature* **192**, 440 (1961).
3. Carter, B. in *Confrontation of Cosmological Theories with Observation* (ed. Longair, M. S.) 291 (Reidel, Dordrecht, 1974).
4. Hawking, S. W. *Nature* **248**, 30 (1974).
5. Jordan, P. *Die Herkunft der Sterne* (Stuttgart, 1974).
6. Weisskopf, V. F. *Science* **187**, 605 (1975).
7. Dyson, F. J. in *Aspects of Quantum Theory* (eds Salam, A. & Wigner, E. P.) 213 (Cambridge University Press, 1972).
8. Chandrasekhar, S. *Mon. Not. R. astr. Soc.* **95**, 207 (1935).
9. Press, W. H. preprint (1973).
10. Salpeter, E. E. *Astrophys. J.* **140**, 796 (1964).
11. Rees, M. J. *Mon. Not. R. astr. Soc.* **176**, 483 (1976).
12. Silk, J. I. *Nature* **265**, 710 (1977).
13. Sciama, D. W. *Mon. Not. R. astr. Soc.* **113**, 34 (1953).
14. Rees, M. J. & Ostriker, J. P. *Mon. Not. R. astr. Soc.* **179**, 541 (1977).
15. Silk, J. I. *Astrophys. J.* **211**, 638 (1977).
16. White, S. D. M. & Rees, M. J. *Mon. Not. R. astr. Soc.* **183**, 341 (1978).
17. Rees, M. J. *Nature* **275**, 35 (1978).
18. Carr, B. J. *Mon. Not. R. astr. Soc.* **181**, 293 (1977).
19. Zeldovich, Y. B. *J.E.T.P. Lett.* **24**, 25 (1977).
20. Liang, E. preprint (1977).
21. Zeldovich, Y. B. *Sov. Phys. Uspecki* **20**, 945 (1978).
22. Isham, C. J., Salam, A. & Strathdee, J. *Phys. Rev.* **D3**, 1805 (1971).
23. DeWitt, B. S. *Phys. Rev. Lett.* **13**, 114 (1964).
24. Carter, B. *Large Numbers in Astrophysics and Cosmology* Paper presented at Clifford Centennial Meet., Princeton (1970).
25. Salpeter, E. E. *Astrophys. J.* **115**, 326 (1952); *Phys. Rev.* **107**, 516 (1967).
26. Hoyle, F. *Astrophys. J. Suppl.* **1**, 121 (1954).
27. Kahn, F. D. in *Emerging Universe* (eds Jacobs, K. C. & Saslaw, W. C.) 71 (Virginia University Press, 1972).
28. Weinberg, S. *Phys. Rev. Lett.* **19**, 1264 (1967): Salam, A. in *Elementary Particle Theory* (ed. Svartholm, N.) (Almquist and Wiksell, Stockholm, 1968).
29. Bruenn, S. W., Arnett, W. D. & Schamm, D. N. *Astrophys. J.* **213**, 213 (1977).
30. Everett, H. *Rev. mod. Phys.* **29**, 454 (1957).
31. Wheeler, J. A. in *Gravitation* Ch. 44 (1971).
32. Patton, G. M. & Wheeler, J. A. in *Quantum Gravity* (eds Isham, C. J., Penrose, R. & Sciama, D. W.) Ch. 9 (Oxford University Press, 1975).