arXiv:astro-ph/0511440 v1 15 Nov 2005

Varying Constants

John D. Barrow
DAMTP
Centre for Mathematical Sciences
Cambridge University
Cambridge CB3 0WA
UK

May 22, 2006

Abstract

We review properties of theories for the variation of the gravitation and fine structure 'constants'. We highlight some general features of the cosmological models that exist in these theories with reference to recent quasar data that are consistent with time-variation in the fine structure 'constant' since a redshift of 3.5. The behaviour of a simple class of varying-alpha cosmologies is outlined in the light of all the observational constraints.

1 Introduction

There are several reasons why the possibility of varying constants should be taken seriously [1]. First, we know that the best candidates for unification of the forces of nature in a quantum gravitational environment only seem to exist in finite form if there are many more dimensions of space than the three that we are familiar with. This means that the true constants of nature are defined in higher dimensions and the three-dimensional shadows we observe are no longer fundamental and do not need to be constant. Any slow change in the scale of the extra dimensions would be revealed by measurable changes in our three-dimensional 'constants'. Second, we appreciate that some apparent constant might be determined partially or completely by spontaneous symmetry-breaking processes in the very early universe. This introduces an irreducibly random element into the values of those constants. They may be different in different parts of the universe. The most dramatic manifestation of this process is provided by the chaotic and eternal inflationary universe scenarios where both the number and the strength of forces in the universe at low energy can fall out differently in different regions. Third, any outcome of a theory of quantum gravity will be intrinsically probabilistic. It is often imagined that the probability distributions for observables will be very sharply peaked but this may not be the case for all possibilities. Thus, the value of the gravitation 'constant', G, or its time derivative, $\dot{G}$, might be predicted to be spatial random variables. Fourth, a non-uniqueness of the vacuum state for the universe would allow other numerical combinations of the constants to have occurred in different places. String theory indicates that there is a huge 'landscape' ($> 10^{500}$) of possible vacuum states that the universe can find itself residing in as it expand and cools. Each will have different constants and associated forces and symmetries. It is sobering to remember that at present we have no idea why any of the constants of Nature take the numerical values they do and we have never successfully predicted the value of any dimensionless constant in advance of its measurement. Fifth, the observational limits on possible variations are often very weak (although they can be made to sound strong by judicious parametrisations). For example, the cosmological limits on varying G tell us only that $\dot{G}/G \leq 10^{-2}H_0$, where H_0 is the present Hubble rate. However, the last reason to consider varying constants is currently the most compelling. For the first time there is a body of detailed astronomical evidence for the time variation of a traditional constant. The observational programme of Webb et al [2, 3] has completed detailed analyses of three separate quasar absorption line data sets taken at Keck and finds persistent evidence consistent with the fine structure constant, α, having been *smaller* in the past, at $z = 1 - 3.5$. The shift in the value of α for all the data sets is given provisionally by $\Delta\alpha/\alpha = (-0.57 \pm 0.10) \times 10^{-5}$. This result is currently the subject of detailed analysis and reanalysis by the observers in order to search for possible systematic biases in the astrophysical environment or in the laboratory determinations of the spectral lines.

The first investigations of time-varying constants were those made by Lord Kelvin and others interested in possible time-variation of the speed of light at the end of the nineteenth century. In 1935 Milne devised a theory of gravity, of a form that we would now term 'bimetric', in which there were two times – one (t) for atomic phenomena, one (τ) for gravitational phenomena – linked by $\tau = \log(t/t_0)$. Milne [4] required that the 'mass of the universe' (what we would now call the mass inside the particle horizon $M \approx c^3G^{-1}t$) be constant. This required $G \propto t$. Interestingly, in 1937 the biologist J.B.S. Haldane took a strong interest in this theory and wrote several papers [5] exploring its consequences for the evolution of life. The argued that biochemical activation energies might appear constant on the t timescale yet increase on the τ timescale, giving rise to a non-uniformity in the evolutionary process. Also at this time there was widespread familiarity with the mysterious 'large numbers' $O(10^{40})$ and $O(10^{80})$ through the work of Eddington (although they had first been noticed by Weyl [6] – see ref. [7] and [1] for the history). These two ingredients were merged by Dirac in 1937 in a famous development (supposedly written on his honeymoon) that proposed that these large numbers (10^{40}) were actually equal, up to small dimensionless factors. Thus, if we form $N \sim c^3t/Gm_{\rm n} \sim 10^{80}$, the number of nucleons in the visible universe, and equate it to the square of $N_1 \sim e^2/Gm_{\rm n}^2 \sim 10^{40}$, the ratio of the electrostatic and gravitational forces between two protons then we are led to conclude that one of the constants, $e, G, c, h, m_{\rm n}$ must vary with time. Dirac [8] chose $G \propto t^{-1}$ to carry the time variation. Unfortunately, this hypothesis did not survive very long. Edward Teller [9] pointed out that such a steep increase in G to the past led to huge increases in the Earth's surface temperature in the past. The luminosity of the sun varies as $L \propto G^7$ and the radius of the Earth's orbit as $R \propto G^{-1}$ so the Earth's surface temperature $T_{\oplus}$ varies as $(L/R^2)^{1/4} \propto G^{9/4} \propto t^{-9/4}$ and would exceed the boiling point of water in the pre-Cambrian era. Life would be eliminated. Gamow subsequently suggested that the time variation needed to reconcile the large number coincidences be carried by e rather than G, but again this strong variation was soon shown to be in conflict with geophysical and radioactive decay data. This chapter was brought to an end by Dicke [10] who pointed out that the $N \sim N_1^2$ large number coincidence was just the statement that t, the present age of the universe when our observations are being made, is of order the main-sequence stellar lifetime, $t_{\rm ms} \sim (Gm_{\rm n}^2/hc)^{-1} h/m_{\rm n}c^2 \sim 10^{10}$ yrs, and therefore inevitable for observers made out of chemical elements heavier than hydrogen and helium. Dirac never accepted this anthropic explanation for the large number coincidences (believing that 'observers' would be present in the universe long after the stars had died) but curiously can be found making exactly the same type of anthropic argument to defend his own varying G theory by highly improbable arguments (that the Sun accretes material periodically during its orbit of the galaxy and this extra material cancels out the effects of overheating in the past) in correspondence with Gamow in 1967 (see [1] for fuller details).

Dirac's proposal acted as a stimulus to theorists, like Jordan, Brans and Dicke [11], to develop rigorous theories which included the time variation of G self-consistently by modelling it as arising from the space-time variation of some scalar field $\phi(\mathbf{x}, t)$ whose motion both conserved energy and momentum and created its own gravitational field variations. In this respect the geometric structure of Einstein's equations provides a highly constrained environment to introduce variations of 'constants'. Whereas in Newtonian gravity we are at liberty to introduce a time-varying $G(t)$ into the law of gravity by

$$F = -\frac{G(t)Mm}{r^2} \tag{1}$$

This creates a non-conservative dynamical system but can be solved fairly straightforwardly [12]. However, this strategy of simply 'writing in' the variation of G by merely replacing G by $G(t)$ in the equations that hold when G is a constant fails in general relativity. If we were to imagine the Einstein equations would generalise to (G_{ab} is the Einstein tensor)

$$G_{ab} = \frac{8\pi G(t)}{c^4} T_{ab} \tag{2}$$

then taking a covariant divergence and using $\nabla^a G_{ab} = 0$, together with energy-momentum conservation ($\nabla^a T_{ab} = 0$) requires that $\nabla G \equiv 0$ and no variations are possible in eq. (2). Brans-Dicke theory is a familiar example of how the addition of an extra piece to T_{ab} together with the dynamics of a $G(\phi)$ fields makes a varying G theory possible. Despite the simplicity of this lesson in the context of a varying G theory it was not taken on board when considering the variations of other non-gravitational constants and the literature is full of limits on their possible variation which have been derived by considering a theory in which the time-variation is just written into the equations which hold when the constant does not vary. These 'limits' are clearly invalid but they will play an important role in guiding us towards the areas where a full theory will find the strongest rigorous bounds. Recently, the interest in the possibility that α varies in time has led to the first extensive exploration of simple self-consistent theories in which a variations occur through the variation of some scalar field.

2 A Simple Varying-Alpha Theory

We are going to consider some of the cosmological consequences of a simple theory with time varying α. Such a theory was first formulated by Bekenstein [13] as a generalisation of Maxwell's equations but ignoring the consequences for the gravitational field equations. Recently, Magueijo, Sandvik and myself have completed this theory [14, 15, 16, 17, 18] to include the coupling to the gravitational sector and analysed its general cosmological consequences. This theory considers only a variation of the electromagnetic coupling and so far ignores any unification with the strong and electroweak interactions. Extensions to include the weak interaction via a generalised Weinberg-Salam theory have also been developed recently, see refs. [19, 20].

Our aim in studying this theory is to build up understanding of the effects of the expansion on varying α and to identify features that might carry over into more general theories in which all the unified interactions vary [21, 22, 23]. The constraint imposed on varying α by the need to bring about unification at high energy is likely to be significant but the complexities of analysing the simultaneous variation of all the constants involved in the supersymmetric version of the standard model are considerable. At the most basic level we recognise that any time variation in the fine structure could be carried by either or both of the electromagnetic or weak couplings above the electroweak scale.

The idea that the charge on the electron, or the fine structure constant, might vary in cosmological time was proposed in 1948 by Teller, [9], who suggested that $\alpha \propto (\ln t)^{-1}$ was implied by Dirac's proposal that $G \propto t^{-1}$ and the numerical coincidence that $\alpha^{-1} \sim \ln(hc/Gm_{\mathrm{p}})$, where m_{p} is the proton mass. Later, in 1967, Gamow [24] suggested $\alpha \propto t$ as an alternative to Dirac's time-variation of the gravitation constant, G, as a solution of the large numbers coincidences problem and in 1963 Stanyukovich had also considered varying α, [25], in this context. However, this power-law variation in the recent geological past was soon ruled out by other evidence [26].

There are a number of possible theories allowing for the variation of the fine structure constant, α. In the simplest cases one takes c and $\hbar$ to be constants and attributes variations in α to changes in e or the permittivity of free space (see [27] for a discussion of the meaning of this choice). This is done by letting e take on the value of a real scalar field which varies in space and time (for more complicated cases, resorting to complex fields undergoing spontaneous symmetry breaking, see the case of fast tracks discussed in [28]). Thus $e_0 \to e = e_0\epsilon(x^\mu)$, where ϵ is a dimensionless scalar field and e_0 is a constant denoting the present value of e. This operation implies that some well established assumptions, like charge conservation, must give way [29]. Nevertheless, the principles of local gauge invariance and causality are maintained, as is the scale invariance of the ϵ field (under a suitable choice of dynamics). In addition there is no conflict with local Lorentz invariance or covariance.

With this set up in mind, the dynamics of our theory is then constructed as follows. Since e is the electromagnetic coupling, the ϵ field couples to the gauge field as ϵA_μ in the Lagrangian and the gauge transformation which leaves the action invariant is $\epsilon A_\mu \to \epsilon A_\mu + \chi_{,\mu}$, rather than the usual $A_\mu \to A_\mu + \chi_{,\mu}$. The gauge-invariant electromagnetic field tensor is therefore

$$F_{\mu\nu} = \frac{1}{\epsilon}\left((\epsilon A_\nu)_{,\mu} - (\epsilon A_\mu)_{,\nu}\right), \tag{3}$$

which reduces to the usual form when ϵ is constant. The electromagnetic part of the action is still

$$S_{\rm em} = -\int d^4x\sqrt{-g}F^{\mu\nu}F_{\mu\nu}. \tag{4}$$

and the dynamics of the ϵ field are controlled by the kinetic term

$$S_\epsilon = -\frac{1}{2}\frac{\hbar}{l^2}\int d^4x\sqrt{-g}\frac{\epsilon_{,\mu}\epsilon^{,\mu}}{\epsilon^2}, \tag{5}$$

as in dilaton theories. Here, l is the characteristic length scale of the theory, introduced for dimensional reasons. This constant length scale gives the scale down to which the electric field around a point charge is accurately Coulombic. The corresponding energy scale, $\hbar c/l$, has to lie between a few tens of MeV and Planck scale, $\sim 10^{19}$GeV to avoid conflict with experiment.

Our generalisation of the scalar theory proposed by Bekenstein [13] described in refs. [15, 16, 17, 18] includes the gravitational effects of ψ and gives the field equations:

$$G_{\mu\nu} = 8\pi G\left(T^{\rm matter}_{\mu\nu} + T^{\psi}_{\mu\nu} + T^{\rm em}_{\mu\nu}e^{-2\psi}\right). \tag{6}$$

The stress tensor of the ψ field is derived from the lagrangian $\mathcal{L}_\psi = -\frac{\omega}{2}\partial_\mu\psi\partial^\mu\psi$ and the ψ field obeys the equation of motion

$$\Box\psi = \frac{2}{\omega}e^{-2\psi}\mathcal{L}_{\rm em} \tag{7}$$

where we have defined the coupling constant $\omega = (c)/l^2$. This constant is of order ~ 1 if, as in [14], the energy scale is similar to Planck scale. It is clear that $\mathcal{L}_{\rm em}$ vanishes for a sea of pure radiation since then $\mathcal{L}_{\rm em} = (E^2 - B^2)/2 =$

0. We therefore expect the variation in α to be driven by electrostatic and magnetostatic energy-components rather than electromagnetic radiation.

In order to make quantitative predictions we need to know how much of the non-relativistic matter contributes to the RHS of Eqn. (7). This is parametrised by $\zeta \equiv \mathcal{L}_{\rm em}/\rho$, where ρ is the energy density, and for baryonic matter $\mathcal{L}_{\rm em} = E^2/2$. For protons and neutrons $\zeta_{\rm p}$ and $\zeta_{\rm n}$ can be *estimated* from the electromagnetic corrections to the nucleon mass, 0.63 MeV and -0.13 MeV, respectively [30]. This correction contains the $E^2/2$ contribution (always positive), but also terms of the form $j_\mu a^\mu$ (where j_μ is the quarks' current) and so cannot be used directly. Hence we take a guiding value $\zeta_{\rm p} \approx \zeta_{\rm n} \sim 10^{-4}$. Furthermore the cosmological value of ζ (denoted $\zeta_{\rm m}$) has to be weighted by the fraction of matter that is non-baryonic. Hence, $\zeta_{\rm m}$ depends strongly on the nature of the dark matter and can take both positive and negative values depending on which of Coulomb-energy or magnetostatic energy dominates the dark matter of the Universe. It could be that $\zeta_{\rm CDM} \approx -1$ (superconducting cosmic strings, for which $\mathcal{L}_{\rm em} \approx -B^2/2$), or $\zeta_{\rm CDM} \ll 1$ (neutrinos). BBN predicts an approximate value for the baryon density of $\Omega_{\rm B} \approx 0.03$ (where $\Omega_{\rm B}$ is the density of matter in units of the critical density $3H^2/8\pi G$) with a Hubble parameter of $H = 60$ Kms^{-1} Mpc^{-1}, implying $\Omega_{\rm CDM} \approx 0.3$. Thus depending on the nature of the dark matter $\zeta_{\rm m}$ can be virtually anything between -1 and $+1$. The uncertainties in the underlying quark physics and especially the constituents of the dark matter make it difficult to impose more certain bounds on $\zeta_{\rm m}$.

We should not confuse this theory with other similar variations. Bekenstein's theory does not take into account the stress energy tensor of the dielectric field in Einstein's equations. Dilaton theories predict a global coupling between the scalar and all other matter fields (not just the electromagnetically charged material) [31, 32, 33, 34, 35]. As a result they predict variations in other constants of nature, and also a different cosmological dynamics.

2.1 The cosmological equations

Assuming a homogeneous and isotropic Friedmann metric with expansion scale factor $a(t)$ and curvature parameter k in eqn. (6), we obtain the field equations ($c \equiv 1$)

$$\begin{aligned} \left(\frac{\dot a}{a}\right)^2 &= \frac{8\pi G}{3}\left(\rho_{\rm m}\left(1+\zeta_{\rm m}\exp\left[-2\psi\right]\right)+\rho_{\rm r}\exp\left[-2\psi\right]+\frac{\omega}{2}\dot\psi^2\right) \\ &\quad -\frac{k}{a^2}+\frac{\Lambda}{3}, \end{aligned} \tag{8}$$

where Λ is the cosmological constant. For the scalar field we have the propagation equation,

$$\ddot\psi + 3H\dot\psi = -\frac{2}{\omega}\exp\left[-2\psi\right]\zeta_{\rm m}\rho_{\rm m}, \tag{9}$$

where $H \equiv \dot a/a$ is the Hubble expansion rate. We can rewrite this more simply as

$$(\dot{\psi a^3}) = N \exp[-2\psi] \tag{10}$$

where N is a positive constant defined by

$$N = -\frac{2\zeta_{\rm m}\rho_{\rm m}a^3}{\omega} \tag{11}$$

Note that the sign of the evolution of ψ is dependent on the sign of $\zeta_{\rm m}$. Since the observational data is consistent with a *smaller* value of α in the past, we will in this paper confine our study to *negative* values of $\zeta_{\rm m}$, in line with our recent discussion in Refs. [14, 15, 16, 17, 18]. The conservation equations for the non-interacting radiation and matter densities are

$$\begin{aligned} \dot{\rho_{\rm m}} + 3H\rho_{\rm m} &= 0 && (12) \\ \dot{\rho_{\rm r}} + 4H\rho_r &= 2\dot{\psi}\rho_r. && (13) \end{aligned}$$

and so $\rho_{\rm m} \propto a^{-3}$ and $\rho_{\rm r}\, e^{-2\psi} \propto a^{-4}$, respectively. If additional non-interacting perfect fluids satisfying equation of state $p = (\gamma - 1)\rho$ are added to the universe then they contribute density terms $\rho \propto a^{-3\gamma}$ to the RHS of eq.(8) as usual. This theory enables the cosmological consequences of varying e, to be analysed self-consistently rather than by changing the constant value of e in the standard theory to another constant value, as in the original proposals made in response to the large numbers coincidences.

We have been unable to solve these equations in general except for a few special cases. However, as with the Friedmann equation of general relativity, it is possible to determine the overall pattern of cosmological evolution in the presence of matter, radiation, curvature, and positive cosmological constant by matched approximations. We shall consider the form of the solutions to these equations when the universe is successively dominated by the kinetic energy of the scalar field ψ, pressure-free matter, radiation, negative spatial curvature, and positive cosmological constant. Our analytic expressions are checked by numerical solutions of (8) and (9).

There are a number of conclusions that can be drawn from the study of the simple BSBM models with $\zeta_{\rm m} < 0$. These models give a good fit to the varying α implied by the QSO data of refs. [2, 3]. There is just a single parameter to fit and this is given by the choice

$$-\frac{\zeta_{\rm m}}{\omega} = (2 \pm 1) \times 10^{-4} \tag{14}$$

The simple solutions predict a slow (logarithmic) time increase during the dust era of $k = 0$ Friedmann universes. The cosmological constant turns off the time-variation of α at the redshift when the universe begins to accelerate ($z \sim 0.7$) and so there is no conflict between the α variation seen in quasars at $z \sim 1 - 3.5$ and the limits on possible variation of α deduced from the operation of the Oklo natural reactor [36] (even assuming that the cosmological variation applies unchanged to the terrestrial environment). The reactor operated 1.8

billion years ago at a redshift of only $z \sim 0.1$ when no significant variations were occurring in α. The slow logarithmic increase in α also means that we would not expect to have seen any effect yet in the anisotropy of the microwave backgrounds [37, 38]: the value of α at the last scattering redshift, $z = 1000$, is only 0.005% lower than its value today. Similarly, the essentially constant evolution of α predicted during the radiation era leads us to expect no measurable effects on the products of Big Bang nucleosynthesis (BBN) [39] because α was only 0.007% smaller at BBN than it is today. This does not rule out the possibility that unification effects in a more general theory might require variations in weak and strong couplings, or their contributions to the neutron-proton mass difference, which might produce observable differences in the light element productions and new constraints on varying α at $z \sim 10^9 - 10^{10}$. By contrast, varying-alpha cosmologies with $\zeta > 0$ lead to bad consequences unless the scalar field driving the alpha variations is a 'ghost' field, with negatively coupled kinetic energy, in which case there are interesting cosmological consequences, [44]. The fine structure falls rapidly at late times and the variation is such that it even comes to dominate the Friedmann equation for the cosmological dynamics. We regard this as a signal that such models are astrophysically ruled out and perhaps also mathematically badly behaved.

We should also mention that theories in which α varies will in general lead to violations of the weak equivalence principle (WEP). This is because the α variation is carried by a field like ψ and this couples differently to different nuclei because they contain different numbers of electrically charged particles (protons). The theory discussed here has the interesting consequence of leading to a relative acceleration of order 10^{-13} [40] if the free coupling parameter is fixed to the value given in eq. (14) using a best fit of the theories cosmological model to the QSO observations of refs. [2, 3]. Other predictions of such WEP violations have also been made in refs. [41, 30, 42, 43]. The observational upper bound on this parameter from direct experiment is just an order of magnitude larger, at 10^{-12}, and limits from the motion of the Moon are of similar order, but space-based tests planned for the STEP mission are expected to achieve a sensitivity of order 10^{-18} and will provide a completely independent check on theories of time-varying e and α.This is an exciting prospect for the future.

2.2 The nature of the Friedmann solutions

Let us present the predicted cosmological evolution of α in the BSBM theory, that we summarised above, in a little more detail. During the radiation era the expansion scale factor of the universe increases as $a(t) \sim t^{1/2}$ and α is essentially constant in universes with an entropy per baryon and present value of α like our own. It increases in the dust era, where $a(t) \sim t^{2/3}$. The increase in α however, is very slow with a late-time solution for ψ proportional to $\frac{1}{2}\log(2N\log(t))$, and so

$$\alpha \sim 2N \log t \tag{15}$$

This slow increase continues until the expansion becomes dominated by negative curvature, $a(t) \sim t$, or by a cosmological vacuum energy, $a(t) \sim \exp[\Lambda t/3]$. Thereafter α asymptotes rapidly to a constant. If we set the cosmological constant equal to zero and $k = 0$ then, during the dust era, α would continue to increase indefinitely. The effect of the expansion is very significant at all times. If we were to turn it off and set $a(t)$ constant then we could solve the ψ equation to give the following exponentially growing evolution for α, [45]:

$$\alpha = \exp[2\psi] = A^{-2}\cosh^2[AN^{1/2}(t+t_0)]; \; A \text{ constant.} \tag{16}$$

From these results it is evident that non-zero curvature or cosmological constant brings to an end the increase in the value of α that occurs during the dust-dominated era. Hence, if the spatial curvature and Λ are both too *small* it is possible for the fine structure constant to grow too large for biologically important atoms and nuclei to exist in the universe. There will be a time in the future when α reaches too large a value for life to emerge or persist. The closer a universe is to flatness or the closer Λ is to zero so the longer the monotonic increase in α will continue, and the more likely it becomes that life will be extinguished. Conversely, a non-zero positive Λ or a non-zero negative curvature will stop the increase of α earlier and allow life to persist for longer. If life can survive into the curvature or Λ-dominated phases of the universe's history then it will not be threatened by the steady cosmological increase in α unless the universe collapses back to high density.

There have been several studies, following Carter, [46] and Tryon [47], of the need for life-supporting universes to expand close to the 'flat' Einstein de Sitter trajectory for long periods of time. This ensures that the universe cannot collapse back to high density before galaxies, stars, and biochemical elements can form by gravitational instability, or expand too fast for stars and galaxies to form by gravitational instability [48, 7]. Likewise, it was pointed out by Barrow and Tipler, [7] that there are similar anthropic restrictions on the magnitude of any cosmological constant, Λ. If it is too large in magnitude it will either precipitate premature collapse back to high density (if $\Lambda < 0$) or prevent the gravitational condensation of any stars and galaxies (if $\Lambda > 0$). Thus, we can provide good anthropic reasons why we can expect to live in an old universe that is neither too far from flatness nor dominated by a much stronger cosmological constant than observed ($|\Lambda| \leq 10\,|\Lambda_{\text{obs}}|$).

Inflationary universe models provide a possible theoretical explanation for proximity to flatness but no explanation for the smallness of the cosmological constant. Varying speed of light theories [49, 27, 50, 51, 52] offer possible explanations for proximity to flatness and smallness of a classical cosmological constant (but not necessarily for one induced by vacuum corrections in the early universe). We have shown that if we enlarge our cosmological theory to accommodate variations in some traditional constants then it appears to be anthropically disadvantageous for a universe to lie too close to flatness or for the cosmological constant to be too close to zero. This conclusion arises because of the coupling between time-variations in constants like α and the curvature or

Λ, which control the expansion of the universe. The onset of a period of Λ or curvature domination has the property of dynamically stabilising the constants, thereby creating favourable conditions for the emergence of structures. This point has been missed in previous studies because they have never combined the issues of Λ and flatness and the issue of the values of constants. By coupling these two types of anthropic considerations we find that too small a value of Λ or the spatial curvature can be as poisonous for life as too much. Universes like those described above, with increasing $\alpha(t)$, lead inexorably to an epoch where α is too large for the existence of atoms, molecules, and stars to be possible [16].

Surprisingly, there has been almost no consideration of habitability in cosmologies with time-varying constants since Haldane's discussions [5] of the biological consequences of Milne's bimetric theory of gravity. Since then, attention has focussed upon the consequences of universes in which the constants are different but still constant. Those cosmologies with varying constants that have been studied have not considered the effects of curvature or Λ domination on the variation of constants and have generally considered power-law variation to hold for all times. The examples described here show that this restriction has prevented a full appreciation of the coupling between the expansion dynamics of the universe and the values of the constants that define the course of local physical processes within it. Our discussion of a theory with varying α shows for the first time a possible reason why the 3-curvature of universes and the value of any cosmological constant may need to be bounded *below* in order that the universe permit atomic life to exist for a significant period. Previous anthropic arguments [7] have shown that the spatial curvature of the universe and the value of the cosmological constant must be bounded *above* in order for life-supporting environments (stars) to develop. We note that the lower bounds discussed here are more fundamental than these upper bounds because they derive from changes in α which have direct consequences for biochemistry whereas the upper bounds just constrain the formation of astrophysical environments by gravitational instability. Taken together, these arguments suggest that within an ensemble of all possible worlds where α and G are time variables, there might only be a finite interval of *non-zero* values of the curvature and cosmological constant contributions to the dynamics that both allow galaxies and stars to form and their biochemical products to persist.

3 The Observational Evidence

New precision studies of relativistic fine structure in the absorption lines of dust clouds around quasars by Webb et al., [2, 3], have led to widespread theoretical interest in the question of whether the fine structure constant, $\alpha_{\rm em} = e^2/\hbar c$, has varied in time and, if so, how to accommodate such a variation by a minimal perturbation of existing theories of electromagnetism. These astronomical studies have proved to be more sensitive than laboratory probes of the constancy of the fine structure 'constant', which currently give bounds on the time variation of $\dot{\alpha}_{\rm em}/\alpha_{\rm em} \equiv -0.4 \pm 16 \times 10^{-16}$ yr^{-1}, [53], $|\dot{\alpha}_{\rm em}/\alpha_{\rm em}| < 1.2 \times 10^{-15}$ yr^{-1}, [54], $\dot{\alpha}_{\rm em}/\alpha_{\rm em} \equiv -0.9 \pm 2.9 \times 10^{-16}$ yr^{-1}, [55] by comparing atomic clock standards based on different sensitive hyperfine transition frequencies, and $\dot{\alpha}_{\rm em}/\alpha_{\rm em} \equiv -0.3 \pm 2.0 \times 10^{-15}$ yr^{-1} from comparing two standards derived from 1S-2S transitions in atomic hydrogen after an interval of 2.8 years [56]. The quasar data analysed in refs. [2, 3] consists of three separate samples of Keck-Hires observations which combine to give a data set of 128 objects at redshifts $0.5 < z < 3$. The many-multiplet technique finds that their absorption spectra are consistent with a shift in the value of the fine structure constant between these redshifts and the present of $\Delta\alpha_{\rm em}/\alpha_{\rm em} \equiv [\alpha_{\rm em}(z) - \alpha_{\rm em}]/\alpha_{\rm em} = -0.57 \pm 0.10 \times 10^{-5}$, where $\alpha_{\rm em} \equiv \alpha_{\rm em}(0)$ is the present value of the fine structure constant [2, 3]. Extensive analysis has yet to find a selection effect that can explain the sense and magnitude of the relativistic line-shifts underpinning these deductions. Further observational studies have been published in refs. [57, 58] using a different but smaller data set of 23 absorption systems in front of 23 VLT-UVES quasars at $0.4 \leq z \leq 2.3$ and have been analysed using an approximate form of the many-multiplet analysis techniques introduced in refs. [2, 3]. They obtained $\Delta\alpha_{\rm em}/\alpha_{\rm em} \equiv -0.6 \pm 0.6 \times 10^{-6}$; a figure that disagrees with the results of refs. [2, 3] . However, reanalysis is needed in order to understand the accuracy being claimed and ensure that all spectral lines are being identified. Other observational studies of lower sensitivity have also been made using OIII emission lines of galaxies and quasars. The analysis of data sets of 42 and 165 quasars from the SDSS gave the constraints $\Delta\alpha_{\rm em}/\alpha_{\rm em} \equiv 0.51 \pm 1.26 \times 10^{-4}$ and $\Delta\alpha_{\rm em}/\alpha_{\rm em} \equiv 1.2 \pm 0.7 \times 10^{-4}$ respectively for objects in the redshift range $0.16 \leq z \leq 0.8$ [59]. Observations of a single quasar absorption system at $z = 1.15$ by Quast et al [60] gave $\Delta\alpha_{\rm em}/\alpha_{\rm em} \equiv -0.1 \pm 1.7 \times 10^{-6}$, and observations of an absorption system at $z = 1.839$ by Levshakov et al [61] gave $\Delta\alpha_{\rm em}/\alpha_{\rm em} \equiv 2.4 \pm 3.8 \times 10^{-6}$. A preliminary analysis of constraints derived from the study of the OH microwave transition from a quasar at $z = 0.2467$, a method proposed by Darling [62], has given $\Delta\alpha_{\rm em}/\alpha_{\rm em} \equiv 0.51 \pm 1.26 \times 10^{-4}$, [63].A comparison of redshifts measured using molecules and atomic hydrogen in two cloud systems by Drinkwater et al [64] at $z = 0.25$ and $z = 0.68$ gave a bound of $\Delta\alpha_{\rm em}/\alpha_{\rm em} < 5 \times 10^{-6}$ and an upper bound on spatial variations of $\delta\alpha_{em}/\alpha_{\rm em} < 3 \times 10^{-6}$ over 3 Gpc at these redshifts. A new study comparing UV absorption redshifted into the optical with redshifted 21 cm absorption lines from the same cloud in a sample of 8 quasars by Tzanavaris et al [65]. This comparison probes the constancy of $\alpha^2 g_{\rm p} m_{\rm e}/m_{\rm p}$ and gives $\Delta\alpha_{\rm em}/\alpha_{\rm em} \equiv 0.18 \pm 0.55 \times 10^{-5}$ if we assume that the electron-proton mass ratio and proton g-factor, g_p, are both constant.

Observational bounds derived from the microwave background radiation structure [66] and Big Bang nucleosynthesis [39, 67] are not competitive at present (giving $\Delta\alpha_{\rm em}/\alpha_{\rm em} \lesssim 10^{-2}$ at best at $z \sim 10^3$ and $z \sim 10^9 - 10^{10}$) with those derived from quasar studies, although they probe much higher redshifts.

Other bounds on the possible variation of the fine structure constant have been derived from geochemical studies, although they are subject to awkward environmental uncertainties. The resonant capture cross-section for thermal neutrons by samarium-149 about two billion years ago ($z \simeq 0.15$) in the Oklo natural nuclear reactor has created a samarium-149:samarium-147 ratio at the reactor site that is depleted by the capture process ^{149}Sm+n $\rightarrow$ ^{150}Sm+γ to an observed value of only about 0.02 compared to the value of about 0.9 found in normal samples of samarium. The need for this capture resonance to be in place two billion years ago at an energy level within about 90 meV of its current value leads to very strong bounds on all interaction coupling constants that contribute to the energy level, as first noticed by Shlyakhter [68, 1]. The latest analyses by Fujii et al [69] allow two solutions (one consistent with no variation the other with a variation) because of the double-valued form of the capture cross-section's response to small changes in the resonance energy over the range of possible reactor temperatures: $\Delta\alpha_{\rm em}/\alpha_{\rm em} \equiv -0.8 \pm 1.0 \times 10^{-8}$ or $\Delta\alpha_{\rm em}/\alpha_{\rm em} \equiv 8.8 \pm 0.7 \times 10^{-8}$. The latter possibility does not include zero but might be excluded by further studies of other reactor abundances. Subsequently, Lamoureax [70] has argued that a better (non-Maxwellian) assumption about the thermal neutron spectrum in the reactor leads to 6σ lower bound on the variation of $\Delta\alpha_{\rm em}/\alpha_{\rm em} > 4.5 \times 10^{-8}$ at $z \simeq 0.15$.

Studies of the effects of varying a fine structure constant on the β-decay lifetime was first considered by Peebles and Dicke [71] as a means of constraining allowed variations in $\alpha_{\rm em}$ by studying the ratio of rhenium to osmium in meteorites. The β-decay $^{187}_{75}$Re$\rightarrow$ $^{187}_{76}$Os+$\bar{\nu}_{\rm e}$+e^{-} is very sensitive to $\alpha_{\rm em}$ and the analysis of new meteoritic data together with new laboratory measurements of the decay rates of long-lived beta isotopes has led to a time-averaged limit of $\Delta\alpha_{\rm em}/\alpha_{\rm em} = 8 \pm 16 \times 10^{-7}$ [72, 73] for a sample that spans the age of the solar system ($z \leq 0.45$). Both the Oklo and meteoritic bounds are complicated by the possibility of simultaneous variations of other constants which contribute to the energy levels and decay rates; for reviews see refs. [74, 75]. They also apply to environments within virialised structures that do not take part in the Hubble expansion of the universe and so it is not advisable to use them in conjunction with astronomical information from quasars without a theory that links the values of $\alpha_{\rm em}$ in the two different environments that differ in density by a factor of $O(10^{30})$. Detailed discussions of this problem when G and α vary have been made in refs. [76, 77, 78].

4 The Role of Inhomogeneities

All early studies of the cosmological consequences of varying constants have assumed that they vary homogeneously. Such an assumption is also implicit when laboratory data or solar system observations are used to constrain cosmological theories of varying G and α. In reality such a simple approach is very dangerous. Our local observations are made inside a gross cosmological overdensity – 10^{30} times denser than the mean density of the background universe – that is not taking part in the universal expansion. We should no more expect laboratory observations of the constancy of α to reflect what is happening on extragalactic scales than we should expect a measurement of the density of the Earth to give a good estimate of the density of the universe. In order to use our local observations effectively we need a theoretical description of how variations in, say, α will vary with the local density of matter as a result of the process of galaxy, star, and planetary formation. For example, when a cosmological overdensity separates out from the expansion of the universe, and collapses under its own gravity, it will eventually come into a stationary virial equilibrium. If α is a space-time variable it will continue changing in the background universe after it has ceased to change in the virialised protogalaxy with a density contrast of about 10^6 with respect to the background universe. In this way we see that the process of galaxy formation leads us to expect that any time variation in fundamental constants will be inevitably accompanied by a space variation that is potentially much more marked. In particular both α and $\dot{\alpha}$ will exhibit different values inside and outside galaxies and galaxy clusters. Moreover, we expect the residual time variations inside galaxies (and hence in terrestrial laboratories) to be significantly smaller than those to be found in extragalactic systems that take part in the expansion of the universe [77, 78]. In contrast, if we go to very large scales where we are observing very small fluctuations long before they collapse into clusters and galaxies, we can calculate the effects of spatially varying 'constants' on the isotropy of the microwave background radiation. The author has recently shown [79] that the evolution eqn. (7) means that spatial variations in α are driven by spatial variations in the matter density which in turn produce spatial variations in the gravitational potential. These potential variations create temperature anisotropies in the microwave background on large angular scales. The observational bound on these variations from the COBE and WMAP satellites allow us to conclude that in these theories spatial variations in α are bounded above by $\delta\alpha/\alpha < 2 \times 10^{-9}$. Very strong bounds can be also derived in this way on allowed spatial variations in G and $m_{\rm e}/m_{\rm p}$and in Bran-Dicke theory and the new theory for varying $m_{\rm e}/m_{\rm p}$ recently devised by Barrow and Magueijo [80].

These theoretical developments, together with the appearance of new observational probes of the constants of physics at high redshift, coupled with recent rapid progress in direct laboratory probes of the stability of atomic systems that depend sensitively on the value of the fine structure constant here and now, promise to create an exciting new focal point in our quest to understand the nature (as well as the number) of the fundamental constants of Nature.

Acknowledgements I would like to thank my collaborators João Magueijo, Håvard Sandvik, John Webb, Michael Murphy, Dagny Kimberly and David Mota for discussions and for their essential contributions to the work described here.

References

[1] J.D. Barrow, *The Constants of Nature: from alpha to omega,* Jonathan Cape, London (2002).

[2] M. Murphy, J. Webb, V. Flambaum, V. Dzuba, C. Churchill, J. Prochaska, A. Wolfe, Mon. Not. R. astr. Soc., 327, 1208 (2001)

[3] J.K. Webb, V.V. Flambaum, C.W. Churchill, M.J. Drinkwater J.D. Barrow, Phys. Rev. Lett. **82**, 884 (1999); J.K. Webb, M.T. Murphy, V.V. Flambaum, V.A. Dzuba, J.D. Barrow, C.W. Churchill, J.X. Prochaska, A.M. Wolfe, Phys. Rev. Lett. **87**, 091301 (2001).

[4] E. A. Milne, *Relativity, Gravitation and World Structure,* Clarendon, Oxford, (1935).

[5] J.B.S. Haldane, Nature 139, 1002 (1937) and 158, 555 (1944) and article in *New Biology*, No. 16, eds. M.L. Johnson et al, Penguin, London (1955).

[6] H. Weyl, Ann. Physik 59, 129 (1919).

[7] J.D. Barrow and F.J. Tipler, *The Anthropic Cosmological Principle*, Oxford UP, Oxford (1986).

[8] P.A.M. Dirac, Nature 139, 323 (1937).

[9] E. Teller, Phys. Rev. 73, 801 (1948).

[10] R.H. Dicke, Rev. Mod. Phys. 29, 355 (1957) and Nature 192, 440 (1961).

[11] C. Brans and R.H. Dicke, Phys. Rev. 124, 924 (1961).

[12] J.D. Barrow, Mon. Not. R. astr. Soc., 282, 1397 (1996).

[13] J.D. Bekenstein, Phys. Rev. D 25, 1527 (1982).

[14] H. B. Sandvik, J. D. Barrow and J. Magueijo, Phys. Rev. Lett. 88, 031302 (2002)

[15] J. D. Barrow, H. B. Sandvik and J. Magueijo, Phys. Rev. D 65, 063504 (2002).

[16] J. D. Barrow, H. B. Sandvik and J. Magueijo, Phys. Rev. D 65, 123501 (2002).

[17] J. D. Barrow, J. Magueijo and H. B. Sandvik, Phys. Rev. D 66, 043515 (2002).

[18] J. Magueijo, J. D. Barrow and H. B. Sandvik, Phys. Lett. B 541, 201 (2002)

[19] D. Kimberly and J. Magueijo, Phys. Lett. B 584, 8 (2004)

[20] D. Shaw and J.D. Barrow, gr-qc/0412135.

[21] T. Banks, M. Dine, M.R. Douglas, hep-ph/0112059.

[22] P. Langacker, G. Segre and M. Strassler, hep-ph/0112233; X. Calmet and H. Fritzsch, hep-ph/0112110; H. Fairbairn and T. Dent, hep-ph/0112279.

[23] W. Marciano, Phys. Rev. Lett. 52, 489 (1984); T. Banks, M. Dine and M.R. Douglas, Phys. Rev. Lett. 88 (2002) 131301; P. Langacker, G. Segre and M. Strassler, Phys.Lett. B528 (2002) 121-128; X. Calmet and H. Fritzsch, hep-ph/0112110.

[24] G. Gamow, Phys. Rev. Lett. 19, 759 (1967).

[25] K.P. Stanyukovich, Sov. Phys. Dokl. 7, 1150 (1963).

[26] F. Dyson, Phys. Rev. Lett. 19, 1291 (1967).

[27] A. Albrecht and J. Magueijo, Phys. Rev. D 59, 043516 (1999).

[28] J. Magueijo, Phys. Rev. D 62, 103521, (2000).

[29] S. Landau, P. Sisterna & H. Vucetich, Phys. Rev. D63, 081303(R) (2001).

[30] G. Dvali and M. Zaldarriaga, hep-ph/0108217.

[31] N. Chamoun, S.J. Landau and H. Vucetich, Phys. Lett. B 504, 1 (2001).

[32] D. Parkinson, B. Bassett and J.D. Barrow, 2003. Phys. Lett. B, 578, 235(2003).

[33] P.P. Avelino, C.J.A.P. Martins and J.C.R.E. Oliveira, Phys. Rev. D 70, 083506 (2004).

[34] J.D. Bekenstein, Phys.Rev. D 66, 024005 (2002).

[35] K.A. Olive and M. Pospelov, Phys. Rev. D 65, 085044 (2002); E. J. Copeland, N. J. Nunes and M. Pospelov, Phys.Rev. D 69, 023501 (2004); S. Lee, K.A. Olive and M. Pospelov, Phys.Rev. D 70, 083503 (2004).

[36] A. Shylakhter, Nature 264, 340 (1976).Y. Fujii et al, Nucl. Phys. B, 573, 377 (2000); T. Damour and F.J. Dyson, Nucl. Phys. B, 480, 37 (1996).

[37] R. Battye, R. Crittenden, J. Weller, Phys. Rev. D 63, 043505.

[38] P.P. Avelino et. al., Phys. Rev. D 62, 123508, (2000) and astro-ph/0102144 (2001)

[39] J.D. Barrow, Phys. Rev. D 35, 1805 (1987).

[40] J. Magueijo, J.D. Barrow and H. Sandvik, astro-ph/0202374.

[41] T. Damour and A. Polyakov, Nucl. Phys. B 423, 532 (1994).

[42] G.R. Dvali and M. Zaldarriaga, Phys.Rev.Lett. 88 (2002) 091303.

[43] T. Damour, F. Piazza, and G. Veneziano, Phys. Rev. D 66, 046007 (2002) and Phys. Rev. Lett. 89, 081601 (2002).

[44] J.D. Barrow, D. Kimberly and J. Magueijo, Class. Quantum Grav. 21, 4289 (2004).

[45] J.D. Barrow and D. Mota, Class.Quant.Grav. 19, 6197 (2002).

[46] B. Carter, in *Confrontation of Cosmological Theories with Observation*, ed., M.S. Longair, Reidel, Dordrecht, 1974), p291 and *Large Numbers in Astrophysics and Cosmology*, unpublished preprint, Inst. Theoretical Astronomy, Cambridge Univ., (1968).

[47] E. Tryon, Nature 246, 387 (1973).

[48] C.B. Collins and S.W. Hawking, Astrophys. J. 180, 317 (1973).

[49] J. Moffat, Int. J. Mod. Phys. D 2, 351 (1993) and astro-ph/0109350.

[50] J.D. Barrow, Phys. Rev. D 59, 043515 (1999).

[51] J.D. Barrow and J. Magueijo, Phys. Lett. B 447, 246 (1999).

[52] J. Magueijo, Phys. Rev. D 63, 043502, (2001).

[53] H. Marion, *et al.*, Phys. Rev. Lett. 90, 150801 (2003).

[54] S. Bize, et al., Phys. Rev. Lett. 90, 150802 (2003).

[55] M. Fischer, et al., Phys. Rev. Lett. 92, 230802 (2004).

[56] E. Peik et al., physics/0402132.

[57] H. Chand et al., Astron. Astrophys. 417, 853 (2004).

[58] R. Srianand et al., Phys. Rev. Lett. 92, 121302 (2004).

[59] J. Bahcall, C.L. Steinhardt, and D. Schlegel, Astrophys. J. 600, 520 (2004).

[60] R. Quast, D. Reimers and S.A Levashov, Astron. & Astrophys. 415, L72004 (2004).

[61] S.A. Levshakov, et al., astro-ph/0408188.

[62] J. Darling, Phys. Rev. Lett. 91, 011301 (2003).

[63] J. Darling, astro-ph/0405240.

[64] M.J. Drinkwater, J.K. Webb, J.D. Barrow and V.V. Flambaum, Mon. Not. R. Astron. Soc. 295, 457 (1998).

[65] P. Tzanavaris, J.K. Webb, M.T. Murphy, V.V. Flambaum, and S.J. Curran, astro-ph/0412649.

[66] P.P. Avelino et al., Phys. Rev. D 62, 123508 (2000) and Phys. Rev. D 64, 103505 (2001); R.A. Battye, R. Crittenden and J. Weller, Phys. Rev. D 63, 0453505 (2001); S. Landau, D.D. Harari, and M. Zaldarriaga, Phys. Rev. D 63, 083505 (2001); C. Martins et al., Phys. Lett. B 585, 29 (2004); G. Rocha et al., N. Astron. Rev. 47, 863 (2003); K. Sigurdson, et al., Phys. Rev. D 68, 103509 (2003).

[67] B.A. Campbell, and K.A Olive, Phys. Lett. B 345, 429 (1995); R.H. Cyburt, et al., astro-ph/0408033.

[68] A.I. Shlyakhter, Nature 264, 340 (1976).

[69] Y. Fujii, et al., Nucl. Phys. B 573, 377 (2000).

[70] S.K. Lamoureaux, Phys. Rev. D 69, 121701 (2004).

[71] P.J. Peebles and Dicke, R.H., Phys. Rev. 128, 2006 (1962).

[72] K.A. Olive et al., Phys. Rev. D 66, 045022 (2000).

[73] K. Olive and M. Pospelov, hep-ph/0110377.

[74] J.P. Uzan, Rev. Mod. Phys. 75, 403 (2003): J.-P. Uzan, astro-ph/0409424.

[75] K.A. Olive and Y-Z. Qian, Physics Today, pp. 40-5 (Oct. 2004).

[76] J.D. Barrow and C. O'Toole, Mon. Not. R. astron. Soc. 322, 585 (2001).

[77] D. Mota, and J.D. Barrow, Phys. Lett. B, 581, 141 (2004); J.D. Barrow and D.F. Mota, Class. Quantum Gravity 20, 2045 (2003).

[78] D. Mota and J.D. Barrow, Mon. Not. Roy. Astron. Soc. 349, 281 (2004).

[79] J.D. Barrow, Phys Rev. D 00, 000000, astro-ph/0503434.

[80] J.D. Barrow and J. Magueijo, astro-ph/0503222