Cosmic microwave background anisotropy

Nick Kaiser* & Joseph Silk†

* Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, UK
† Astronomy Department, University of California, Berkeley, California 94720, USA

Current hypotheses for the origin of structure in the Universe lead to predictions of the amplitudes of anisotropies in the cosmic microwave background radiation. The dipole anisotropy is related to density fluctuations on large scales and to other determinations of our motion relative to distant galaxies. Observation and theory are coming tantalizingly close to measuring the elusive anisotropy, or to revealing that our ideas about the origin of galaxies and large-scale structures are in need of substantial revision.

STUDIES of the angular structure of the microwave background radiation play two distinct roles in cosmology. First, there are those anisotropies that are intrinsic to the microwave background. These provide an unsurpassed glimpse of the early Universe, propagating freely to us from an epoch when the Universe was much simpler than it is today. Galaxy clusters had not formed, nor, most probably, had galaxies themselves. Detection of anisotropies in the background radiation would provide a unique measure of the primordial density fluctuations from which the present large-scale structure of the Universe developed. Yet despite continuing and ever-improving measurements, no such anisotropies have been discovered. Here we describe some of the current hypotheses for the origin of structure in the Universe, and the calculations which have been made to predict the amplitude of the microwave anisotropy. A different test of cosmological theories is provided by the dipole anisotropy of the microwave background, which arises from our motion relative to a uniformly expanding 'cosmic frame' defined by the mean state of motion of matter at great distances. We describe this anisotropy as extrinsic, and infer a solar motion which can be used to constrain both the large-scale density fluctuations in which we are embedded and independent determinations of our motion relative to distant galaxies. In the latter part of this review we describe the present status of these studies.

Origin of structure

Large-scale structure originated, according to the primaeval hypothesis, from infinitesimal density fluctuations in the earliest instants of the Big Bang. Inflation[1] has provided a tentative answer to the question of the origin of these fluctuations[2-5]. Quantum fluctuations are amplified to large scales during the exponential expansion phase which characterizes the inflationary epoch. The inflation is driven by vacuum energy during a first-order phase transition which occurs as the energy of matter decays. The universe evolves from the symmetric state of false vacuum, when the fundamental forces are unified, to the asymmetric state of the true vacuum. During the inflationary phase, described by de Sitter space-time, there is no particle horizon. The fluctuations are boosted to a scale-invariant amplitude, up to scales comparable to the radius of the inflated space-time bubble, which greatly exceeds our present horizon, and models can be constructed that yield density fluctuations at horizon crossing of

$$(\delta\rho/\rho)_{\rm H} \approx 10^{-4} \qquad (1)$$

Values much larger ($(\delta\rho/\rho)_{\rm H} \approx 1$) would be disastrous for cosmology, producing black holes and large inhomogeneities, and much smaller values would not lead to any galaxy formation. It is remarkable that constraints on the precise value of $(\delta\rho/\rho)$ can be inferred from constraints on temperature fluctuations $\delta T/T$ in the cosmic microwave background. Indeed, all currently acceptable cosmological theories require large-scale structure to have originated from infinitesimal fluctuations at early phases of the Big Bang, and these inevitably leave a signature in the form of temperature fluctuations $\delta T/T$. Whether the amplitude (1) will emerge naturally from a quantum gravity model is not known. Because quantum gravity is poorly understood, most approaches to this problem begin when a classical description first becomes valid, well below the Planck energy of $\sim 10^{19}$ GeV. According to one school of thought, the inflationary potential can be designed to yield the correct $\delta\rho/\rho$, which can then be interpreted as an otherwise unknown particle physics parameter. This process results in a scale-invariant spectrum of gaussian density fluctuations. Another view is that the initial fields that describe the energy content of the Universe are chaotic and highly variable in space and time[6]. Beginning with this initial space-time foam, an anthropic argument is applied to infer that a very special region, with sufficiently regular fields, will inevitably exist and be able to inflate sufficiently to dominate space-time, producing a highly isotropic (and nearly homogeneous) universe. More than one epoch of inflation could have occurred[7], and a wide range of possible primordial fluctuation spectra can arise.

One intriguing consequence is that the primordial fluctuations might be formed exclusively on small scales, far smaller than those of galaxies or galaxy clusters. Galaxy formation can still occur, however, if nonlinear structures develop on stellar mass scales; an amplification process can then produce large-scale structure[8,9]. This happens as follows: massive stars form, explode as supernovae, and sweep up large shells of ambient matter which subsequently fragment on much larger scales, into pieces with the mass of star clusters. If sufficient energy is to be released by dying stars at this stage for the process to repeat, amplifying to larger and larger scales, then formation of structure on galactic mass scales can be achieved. Galactic binding energy, in such schemes, may be said to be of mediaeval, rather than primaeval, origin: it is secondary, having been induced by the explosions altering the distribution of matter.

In the primaeval picture, the excess binding energy of a cluster of galaxies can be traced back to a primordial excess of density in the same co-moving region at the inflationary epoch, whereas in the mediaeval picture this binding energy is imposed at a very late epoch, when the region is well inside the horizon. There is yet another hypothesis, the 'cosmic string' model[10,11], in which the binding energy fluctuations on a given scale are induced by the dynamics of the string network as that scale enters the horizon.

In each of these models there is a link between the very earliest instants of the Big Bang and the large-scale structure of the Universe today. The microwave background fluctuations result from the linear growth regime of the primordial fluctuations, and thus bypass much of the uncertainty associated with nonlinear evolution. Observations of the microwave background anisotropy provide a unique probe which may help to distinguish between the various models for the origin of galaxies.

Intrinsic anisotropy

Figure 1 shows what one may learn by studying the anisotropy of the cosmic microwave radiation. The microwave photons stream freely to the observer ('here and now') from the surface of last scattering, which in the standard model of the Big Bang occurs in the redshift range $z \approx 1{,}100$–$1{,}500$. The last scattering surface has a finite thickness Δz, which is due to the noninstantaneous recombination of the hydrogen and to the residual ionization level which freezes out. A transmission factor per unit redshift interval, $(\mathrm{d}\tau/\mathrm{d}z)\exp(-\tau)$, defines the effective profile of the last scattering surface, where $\tau = \int_0^z n_e \sigma_T c\,\mathrm{d}t$ is the optical depth to redshift z, σ_T is the Thomson cross-section and n_e is the electron density. This function can be approximated by a gausian which peaks at $z = 1{,}065$, with a gaussian width $\Delta z = 80$ (ref. 12). The corresponding co-moving distance that describes the thickness of the last scattering surface is $\Delta L \approx 7h^{-1}\,\Omega^{-1/2}$ Mpc (where h is the Hubble constant divided by $100\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$).

The finite width of this shell smooths out fluctuations on angular scales $\Delta\theta \leq 8\,\Omega^{-1/2}$ arc min. Even larger-scale primordial fluctuations are not necessarily preserved: a source of energetic photons or heat could re-ionize the Universe or increase its ionization over the low post-recombination value ($n_e/n_H \leq 10^{-3}$) and thereby shift the last scattering surface towards lower redshift. As the thermal history of the Universe at the relevant epoch ($1{,}000 \geq z \geq 30$) is uncertain, there is considerable ambiguity in the interpretation of fine-scale anisotropy. This ambiguity is avoided if we search over angular scales exceeding that subtended by the particle horizon at the latest possible epoch of last scattering. No causal process could erase such structure. This angular scale is $\sim(\Omega/z)^{1/2}$ radians, where Ω is the cosmological density parameter. As the minimum redshift at which significant scattering could have occurred ($\tau \geq 1$) is ~ 30 if the baryon density parameter $\Omega_b \leq 0.1$, we infer that an observational strategy aimed at fluctuations over angular scales of $\geq 10°$ would involve the least number of assumptions about the early evolution of the Universe.

The corresponding co-moving linear scale is $\sim 1{,}000$ Mpc. At recombination, a linear scale of 100 Mpc subtends an angle of $\sim\Omega h$ degrees. Observations of the background radiation fluctuations therefore nicely complement studies of galaxy clustering, which provide firm evidence for structure only on scales of up to a few Mpc, in probing the large-scale structure of the Universe. Thus, the interpretation of large-angle anisotropy is independent of the uncertain ionization history, but is crucially dependent on the assumed extrapolation of density fluctuations to large scales by means of equation (1). For small-scale anisotropy the situation is reversed. Regardless of the angular scale considered, the radiation has streamed freely since the Universe was very homogeneous, at which time it was well described by linear fluctuation theory. Thus, in contrast with studies of galaxy formation and clustering[13–16], the complexities of nonlinear evolution are completely bypassed. The detection of anisotropy on any angular scale could provide a unique and direct measure of the primordial fluctuations from which large-scale structure evolved.

Predicting $\Delta T/T$. A variety of calculations predicting $\Delta T/T$ can be found in the recent literature; they differ mainly in the assumed material content of the Universe, the assumed initial fluctuations and the normalization to present-day structure. In order to make sense of the issues here, we first describe one set of calculations, which we refer to somewhat arbitrarily as the standard model, and then describe variations on this theme.

The 'standard' model: cold dark matter with isentropic gaussian fluctuations. The fluctuations in this model have been studied in detail[17,18]. The Universe is assumed to have three components: a cold, collisionless fluid (this might represent axions, for example); the baryonic gas, initially hot and fully ionized; and the photons and neutrinos, described by a distribution function, with the former being coupled to the gas by electron scattering. The initial state is assumed to be isentropic, which means that, for any two types of particle A and B, the ratio of the number density of As to Bs, n_A/n_B, is a constant independent of position. This terminology arose in the context of a universe containing only radiation and plasma, for which the ratio of the number of photons to the number of baryons n_γ/n_B is a measure of the entropy (per baryon). Such perturbations are sometimes referred to as 'adiabatic' because they can be generated by compressing all of the particles in some region, producing a net density perturbation $\delta\rho(r) \equiv \rho(r) - \rho$. Strictly speaking, one must also specify the initial peculiar velocity field, but this is usually dispensed with by stipulating that the perturbation should contain only growing modes.

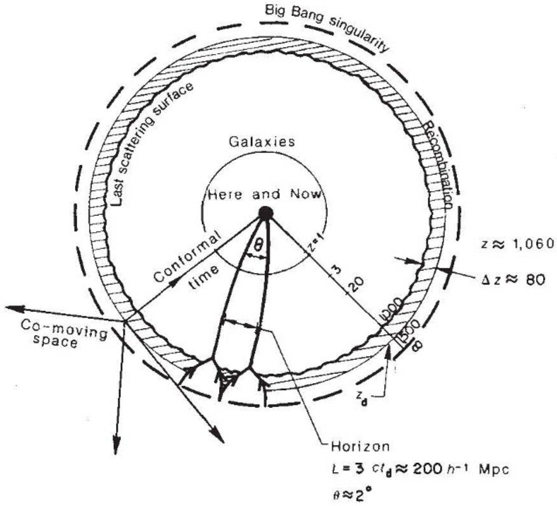

Fig. 1 A co-moving space/conformal time diagram of the Big Bang. The observer ('here and now') is at the centre; the Big Bang singularity is the outermost dashed circle, and the horizon scale is schematically indicated at last scattering.

A 'gaussian' field is one which can be generated as a linear sum of spatial Fourier components with random phases. Such a field is specified by its power spectrum, which in this case is taken to be the Zel'dovich spectrum, $|\delta_k|^2 = |\delta\rho/\rho|_k^2 = Ak$ (where k is the wavenumber of a Fourier component and A is a constant to be specified by a normalization described below), which results in binding energy fluctuations that are scale-invariant.

With the initial conditions thus specified, the spatial Fourier modes can evolve independently until some time after recombination. This results in final linear fluctuations in the matter and radiation which are now decoupled. The angular dependence of the sky fluctuations is now completely specified. The amplitude, however, remains to be fixed. This 'normalization' is performed by requiring that the density fluctuations be of sufficient amplitude to generate the observed large-scale structure. This is conventionally measured by the integral of the galaxy–galaxy correlation function: $J_3(x) = 3/x^3 \int_0^x \xi_{\text{gal-gal}}(r) r^2\,\mathrm{d}r$, where x is a measure of the separation between an arbitrarily selected pair of galaxies. Note that despite the fact that $\xi(r) \gg 1$ on small scales, this integral satisfies linear theory[19]. One evaluates this at a scale where the fluctuations are linear but large enough so that $\xi(r)$ is well known. For example, at $x = 10h^{-1}$ Mpc, the r.m.s. fluctuations are $\sim 50\%$. This normalization assumes that the galaxies are 'fair tracers' of the density field. Although this is a fairly natural assumption, this need not necessarily be the case. If the Universe has closure density ($\Omega = 1$), then it seems that galaxies must be more strongly clustered than mass, at least in the dense clusters where virial analysis is applied to estimate mass to light ratios. In considering these models, some allowance must be made for this, somewhat reducing the amplitude of $\delta T/T$.

There are two final complications. Uncertainty in the Hubble constant H_0 leads to uncertainty in $\delta T/T$. Specifically, for cold dark matter, a decrease in H_0 leads to an increase in L_{eq} (the radius of the co-moving sphere encompassing the causal horizon at the epoch of equality of matter and radiation densities), and thus to an increase in $\delta T/T$. Similarly, if $\Omega<1$, $\delta T/T$ is increased, approximately in proportion to Ω^{-1}. This is because the Universe becomes curvature-dominated at redshift $\sim 1/\Omega$, and fluctuation growth effectively ceases. One is now in a position where $\delta T/T$ can be predicted. The detailed coupling of matter and radiation must, of course, be followed. This involves solving the perturbed Boltzmann equation for the photon intensity today[20,21]. Rather than give any details here, we will simply indicate in a qualitative manner the magnitude of the various terms that contribute to $\delta T/T$.

On small scales, corresponding to a co-moving scale that subtends the thickness of the last scattering surface at ~10 Mpc, there are adiabatic temperature fluctuations. These were first predicted in 1967[22,23], and would be of order $(1/3)\delta\rho/\rho$ were recombination instantaneous and complete. Their angular scale is typically 5-10 arc min, and fluctuations of smaller angular scale do not survive recombination. The amplitude of these small-scale fluctuations is reduced, because of the finite thickness of the last scattering (LS) surface, to $\delta T/T \approx (\delta\rho/\rho)_{LS}/30$. On larger scales, the random motions of fluctuations induce Doppler shifts. These dominate $\delta T/T$ at co-moving scales $d \approx L_{rec}$ (the radius of the co-moving sphere encompassing the causal horizon at recombination) at ~50-100 Mpc, corresponding to angular scales of 1-2°, and are of order $\delta T/T \approx (\delta\rho/\rho)_{LS}(d/ct)_{LS}$, where c is the speed of light and t is the time elapsed since the Big Bang. On the largest scales, the gravitational potential difference between last scattering and now generates gravitational redshifts, of order $\delta T/T = (\frac{1}{3})(\delta\rho/\rho)_0(d/ct)_0^2$. These fluctuations, also first predicted in 1967[24], are contributed by co-moving scales d of 10^3-10^4 Mpc, and dominate over angular scales greater than a few degrees[18].

The result of all this is a gaussian sky pattern (the gaussian nature of the fluctuations being preserved in linear theory), with roughly equal power in each logarithmic interval of angular wavenumber. The r.m.s. amplitude is a function of global cosmological parameters such as Ω, h and Ω_b, and constraints can be placed on these by comparison with the stringent small-scale anisotropy limits. High-density models are generally consistent with these limits, but low-density models can be excluded. There are two reasons for this: first, there is reduced growth of $\Delta\rho/\rho$ at late times if $\Omega \ll 1$, and second, the fluctuations appear at a more favourable angular scale. The results we have described were obtained numerically. The anisotropy on large scales, much greater than the horizon size at decoupling, for these isentropic fluctuations can be calculated analytically[24] (see ref. 25 for an application to cold dark matter). The effect is analogous to the usual gravitational redshift effect, with photons being redshifted (blueshifted) if they were last scattered in overdense (underdense) regions. The amplitude of $\Delta T/T$ is one third of the newtonian peculiar potential ϕ evaluated at the point of emission.

In these calculations it is assumed that the Universe is not re-ionized after recombination. There is some justification for this in that, for most reasonable choices of parameters, the spectrum of density fluctuations is such that very little matter forms nonlinear structure at a sufficiently early epoch to plausibly re-ionize the Universe[17,18]. We now describe some variations on this 'standard' model.

Hot dark matter. An interesting alternative is to replace the cold dark matter (axions, photinos, or the like) with hot dark matter such as massive neutrinos. The main effect of this is to erase post-recombination density fluctuations on scales smaller than that of superclusters (the angular dependence of $\Delta T/T$ being largely unaffected because the $\Delta T/T$ fluctuations on small scales are in any case smeared out)[17,26]. The normalization to J_3 (for galaxies) is now questionable, because galaxy formation is a secondary phenomenon in a neutrino-dominated universe, arising from pancake fragmentation[27]. An alternative normalization is to require that 'pancake' collapse should occur sufficiently early to form quasars and the most distant galaxies. For a neutrino mass of ~30 eV, the predicted anisotropy is compatible with the current upper limits.

Isocurvature fluctuations. An alternative to the isentropic fluctuations discussed so far is the 'isocurvature' mode. These fluctuations are, in a sense, orthogonal to the isentropic modes because, while the latter have fluctuations in the total energy (and therefore in space curvature) but no variation in the relative abundances of the various particles, fluctuations can be realized by spatially varying the equation of state on some initial $t=$ constant hypersurface in a universe which was hitherto absolutely homogeneous. This process conserves energy locally, so any excess in one species is balanced by a deficit in the others. It is the constancy of the spatial curvature on this initial hypersurface which gives rise to the terminology employed here. These fluctuations are not, strictly speaking, isothermal, but are often referred to as 'entropy' or 'isothermal' fluctuations.

Isocurvature fluctuations in a universe dominated by cold dark matter have recently been investigated[28,29]. The small-scale anisotropy is very similar to that arising from isentropic fluctuations. The large-scale anisotropy, however, is about an order of magnitude larger. This increase is a consequence of two factors: the first is that the spectrum of density fluctuations is flatter than in the isentropic case, because no sub-horizon growth occurs in the radiation era. Thus, with the same normalization, the value of $\Delta\rho/\rho$ at large scales is about twice as large.

The origin of the other factor can be understood as follows. Imagine that, at some early time $z \gg z_{eq}$, we impose a positive perturbation to the axion density $\Delta\rho_A/\rho_A = \Delta n_A/n_A = \varepsilon$ in some large spherical region destined to enter the horizon at some time after z_{EQ}. We must also impose a smaller compensating negative perturbation in the radiation density, $\Delta\rho_\gamma/\rho_\gamma = -(\rho_A/\rho_\gamma)\varepsilon$, so that ρ_{TOTAL} is unperturbed. When this region enters the horizon, the density of the radiation will be negligible, having been redshifted, so we will have a positive perturbation to the proper mass enclosed: $(\Delta M/M)_{proper} = \Delta n_A/n_A = \varepsilon$. However, from considerations of causality, we know that the gravitational mass as registered by an external observer is unperturbed. There is no paradox here; the perturbed region must now be sitting in a potential well of depth $\phi = \varepsilon$ so that $M_{grav} = M_{proper}(1-\phi)$ is unchanged. If the perturbation were isentropic then we would see the Sachs-Wolfe term, $(\Delta T/T)_{SW} = -\phi/3 = -\varepsilon/3$. However, this perturbation is not isentropic: we initially increased the number of axions but the number of photons was almost unaltered. Thus, relative to an isentropic state, we have a negative entropy perturbation, $\Delta n_\gamma/n_\gamma = -\varepsilon$, which gives rise to a temperature fluctuation $(\Delta T/T)_{entropy} = 4(\Delta n_\gamma/n_\gamma)/3 = 4\varepsilon/3$. We see then that the usual Sachs-Wolfe term is augmented by an entropy perturbation of the same sign but four times larger, and the net result is to increase $(\Delta T/T)_{TOTAL}$ by a factor of ~5. With the conventional normalization, the predicted anisotropy is incompatible with the limits of Fixen *et al.*[30].

The re-ionized Universe. If re-heating were to occur, the Universe would become transparent over approximately one expansion time at the mean redshift of last scattering. The photons we see would then arrive from a broad range of redshifts, $\Delta z \approx z_{ls}$. The temperature anistropy due to the large-scale structure in such a re-ionized universe is smaller than that predicted to arise in a un-ionized universe. This is because, along any line of sight through the cosmic photosphere, we will see fluctuations due to many regions adding incoherently and this results in statistical cancellation.

The dominant source of anisotropy in a re-ionized universe is the peculiar motion of the last scatterers. The calculation of the anisotropy is greatly simplified by two approximations which are expected to hold to high accuracy. First, at the epoch of last scattering, the density fluctuations that give rise to the currently observed large-scale structure were still in the linear regime, and second, the dynamical effect of the radiation on the matter was almost certainly negligible. Hence we need only calculate the effect of the matter on the radiation.

The statistical cancellation due to the projection of density fluctuations along the line of sight depends on the statistical nature of the density perturbation field $\delta(k)$. If this density field was gaussian (that is, if the Fourier components $\delta(k)$ had random phases), then the temperature fluctuations on the sky, $\Delta(\theta)$, will also be a gaussian field. The power spectrum of this field, Δ^2, is related to the power spectum of density fluctuations by $\Delta^2 \propto k^{-4}\delta^2(k)$, where the constant of proportionality depends on the ionization history of the Universe[31]. A consequence of the fact that the scale of the fluctuations is much smaller than the thickness of the photosphere is that the fluctuations $\Delta(k)$ at some angular wavenumber k are determined solely by the value of δ^2 at the corresponding spatial wavenumber.

A useful measure of the amplitude of the temperature fluctuations at angle θ is $(\delta T/T)_\theta \approx \sqrt{k^2\Delta^2(k)}$, where $k \approx 1/\theta$. Similarly, the amplitude of the density fluctuation on scale λ is $(\Delta\rho/\rho)_\lambda \approx [k^3\delta^2(k)]^{1/2}$. Thus, to order of magnitude, we have $(\Delta T/T) \approx (H\lambda)^{5/2}(\Delta\rho/\rho)$ (where H, λ and $\delta\rho/\rho$ are evaluated at z_{ls}). Equivalently, we can write $(\Delta T/T) \approx (1/\sqrt{N})\phi_\lambda$, where ϕ_λ is the gravitational potential fluctuation on scale λ and N is the number of regions of size $\sim\lambda$ seen along a line of sight. One can infer that the anisotropy due to fluctuations of a given scale λ is smaller for smaller z_{ls} (as N is then increased), and ϕ_λ is approximately equal to the square of the velocity dispersion of the final nonlinear structures produced. As this velocity dispersion appears to be a non-decreasing function of mass (and therefore of λ), the result tells us that the greatest anisotropy is generated by the largest structures we see today. If galaxies fairly trace the mass distribution, then the anticorrelation seen in the redshift survey suggests that the power spectrum falls rapidly for $k \leq 2\pi/60h^{-1}$ Mpc. These fluctuations would generate $\Delta T/T \approx (3\text{-}6)\times 10^{-6}$ on an angular scale of $\sim 1°$. (We have here assumed $\Omega \approx 0.2$ in accord with the assumption that galaxies trace the mass). In a high-density universe ($\Omega \approx 1$), the amplitude would be smaller by a factor of ~ 3 and the anisotropy would appear at an angular scale of ~ 10 arc min.

On larger scales, the amplitude of $\Delta T/T$ depends on the extrapolation of the power spectrum. If this is assumed to be a power law with $\delta_k^2 \propto k^n$, then $\Delta T/T \propto \theta^{(-n+2)/2}$ and would increase with angular scale if $n \leq 2$. The estimate we have given above is the minimal anisotropy implied by the structure observed today, and it does not require any extrapolation of the power spectrum to large scales. The predicted anisotropy is well below the current upper limits on the relevant scales.

So far we have not specified the source of energy which re-ionizes the Universe. An interesting possibility is that these pre-galactic sources of radiation may, if they are sufficiently inhomogeneous, generate the large-scale structure[32]. A constraint on such theories is that, in addition to the 'Doppler scattering' fluctuations we have discussed above, the radiation released would probably have been degraded to microwave frequencies and would appear anisotropic on a scale of a few degrees, this being the angular size subtended by the width of the photosphere.

This effect was investigated in ref. 33. The assumptions are that a population of sources turn on and burn at redshift z_*, that these sources are inhomogeneous on some scale λ_* and are uncorrelated on larger scales (note that there is no assumed initial density inhomogeneity), and that these sources produce a substantial fraction of the microwave background. The diffusion or streaming of this radiation away from the sources then generates density fluctuations by the pressure of the radiation of the gas. This process can transfer momentum, and generate density fluctuations, on scales up to the radiation Jeans length, which is comparable to the scale of the largest structures we see today. The redshift of burning was taken[33] to be $z_* \approx$ 200–500; if z_* is lower than this, then radiation drag becomes inefficient, and for higher z the luminosity of the sources must exceed the Eddington luminosity. This process could generate the observed large-scale structure if the initial scale of radiation inhomogeneity is of the order of a galactic mass scale. The transfer of radiation smooths out the initially inhomogeneous radiation on scales up to the horizon scale at last scattering. The final microwave background anistropy predicted (assuming thermalization of the radiation by dust) appears at an angular scale of a few degrees and is comparable to the present upper limits on this scale.

On larger scales the temperature anisotropy decreases as $\Delta T/T \propto \phi^{-1}$, and so any quadrupole anisotropy would be very small. This $\Delta T/T \propto \phi^{-1}$ behaviour at large angles is the hallmark of spontaneous theories of formation of structure. Most 'primordial' theories of large-scale structure invoke a Zel'dovich spectrum which would generate large-angle $\Delta T/T$ fluctuations which are scale-independent. Thus observation of large and comparable quadrupole, octopole and higher-order moments would rule out spontaneous generation theories. Unfortunately, our Galaxy gives rise to anisotropy, with significant low-order multipoles. It may then be that one would be able to detect the Zel'dovich anistropy only on intermediate angular scales, where one could test for a dependence of $\Delta T/T$ on galactic latitude.

Anisotropy from cosmic strings. All of the mechanisms we have discussed so far give rise to gaussian fluctuations, either by virtue of the assumed initial conditions or, as in the re-ionized universe, where we have many fluctuation regions adding incoherently, so that the central limit theorem should apply. An interesting alternative is supplied by the 'cosmic string' hypothesis[10,11] (for a review of this subject see ref. 34). The perturbations provided by cosmic strings are inherently non-gaussian and result in step-like discontinuities snaking across the microwave sky[35-37]. The pattern is like that arising from a Zel'dovich spectrum, scale-invariant on large scales. In this picture, one would naturally expect the Universe to be re-ionized, so the theory is relatively immune to the constraints from small-scale anisotropy. In order to detect the distinctive features predicted in this model, one would need to scan a region of sky at least $\sim 10°$ square with reasonably good resolution.

ΔT/T *in the explosive model.* Explosive amplification[8,9] begins with rare seeds, either massive objects ($\sim 10^6$ solar masses) or possibly galaxies themselves, and results in a foam-like distribution of galaxies: swept-out voids separate thin shells, sheets and ridges where shells intersect, and there most of the galaxies form. The typical void size can be up to ~ 5 Mpc in diameter if sufficient explosive energy is available. Although this falls short of the observed voids in the galaxy distribution, which are ~ 25 Mpc across[38], the qualitative resemblance of the bubble-like structure of the galaxy distribution to that predicted in explosive amplification is tantalizing, and this model is a serious contender for explaining galaxy formation.

In fact, the voids are not completely empty, but contain hot shocked gas at a temperature of $\geq 10^8$ K. In the pancake fragmentation model, the physics of galaxy formation is very similar to that in the Ostriker-Cowie-Ikeuchi model[8,9]. Pancakes collapse to form thin, dense sheets of gas which fragment into galaxies, leaving behind large regions of tenuous hot gas. This hot gas can result in observable fluctuations in the cosmic microwave background, by means of the Sunyaev-Zel'dovich effect[39]. Microwave photons Compton-scatter against the hot electrons and acquire a perturbation in temperature, in the Rayleigh-Jeans part of the spectrum, of $\Delta T/T \approx -(2kT/m_ec^2)\tau_{es}$, where k is the Boltzmann constant, m_e is the electron mass and τ_{es} is the scattering optical depth. $\Delta T/T$ for an individual pancake is proportional to the pressure, which is constant in the shocked region and depends only on the comoving wavelength of the pancake and the epoch of pancake collapse. Although the distortion from an individual pancake at $z \approx 1$ is small ($\Delta T/T \approx 10^{-7}$), the contribution from the many ($N > 100$) pancakes along a typical line of sight increases the variance of the temperature fluctuations by a factor of $N^{1/2}$ between different lines of sight. Hogan[40] performed a similar calculation for the explosive galaxy formation model. He found that if the observed large-scale structure as measured by the galaxy autocorrelation function is generated explosively, small-scale anisotropy comparable to the Uson-Wilkinson limit[41] is produced. Vishniac and Ostriker[42] generalized this argument further, by noting that if the binding energy of the gaseous components of newly bound structures is thermalized, regardless of its origin (whether explosive or gravitational), then temperature fluctuations of amplitude $\sim 10^{-5}$ result on sub-arc-minute scales. Primordial fluctuations are greatly diminished on these scales due to the smearing associated with the finite thickness of the last scattering surface, and such secondary fluctuations are expected to dominate.

Searches for intrinsic anisotropy. Since the discovery of the microwave background radiation[43], many attempts have been made to search for angular anisotropy. These searches have used a variety of techniques to probe a range of angular scales spanning more than three orders of magnitude. It is often convenient, in the discussion of temperature anisotropies, to decompose the sky temperature pattern into spherical harmonics; these provide a complete and orthogonal set of modes. As we have mentioned, the $l = 1$ mode is probably dominated by our locally generated peculiar velocity, which may also generate a small $l = 2$ component. Any other anisotropies must be intrinsic. Observations at different angular scales suffer from different problems and require different observing strategies; it is useful to distinguish between 'large' angular scales for which $l \leq 10$, 'intermediate' angular scales with $l \approx 10$–100, and 'small' angular scales with $l \geq 100$.

For large and intermediate angular scales it is necessary to raise the observing apparatus above the bulk of the atmosphere. Anisotropometers have been deployed from balloons[30,44–46], from U-2 aircraft[47] and from a satellite[48]. These observations have typically employed beam-widths of a few to ten degrees and have obtained coverage of a significant fraction of the whole sky. Initially there was much emphasis on the possibility of a quadrupole anisotropy, and observers obtained estimates of the quadrupole tensor components from their data. These observations also contain useful information about higher-angular-frequency fluctuations on scales down to the beam-width. Although there have been claims of positive detections of quadrupole anisotropy at levels significantly above that of the noise, none of these have withstood the test of time, and the current upper limit from the Soviet RELIC experiment on any $l = 2$ mode is $\Delta T/T \leq 3 \times 10^{-5}$ with 95% confidence. The limit on the $l = 4$ mode is 5×10^{-5} (95% confidence). These limits are much smaller than the dipole amplitude, and this fact gives some support to the belief that most, if not all, of the dipole anisotropy is extrinsic or local, since one would expect any intrinsic dipole anisotropy to be of similar amplitude to the other low-order moments.

On intermediate angular scales there has also been one claimed positive detection of anisotropy[44], although few details of the method of analysis were given and no confidence level was quoted. Other workers have quoted upper limits of $\Delta T/T \leq 3 \times 10^{-5}$. More recently, an upper limit[49] of $\Delta T/T \leq 5 \times 10^{-4}$ (95% confidence) has been set at a beam-width of 2°. An upper limit of 5×10^{-5} (95% confidence) has been set over ~6–12° in a recent dedicated experiment at 3 cm wavelength (R. D. Davies *et al.*, in preparation). At small angular scales the emission from the atmosphere is less of a problem, and searches for anisotropy on scales of a few arc min have been made using conventional ground-based radio telescopes. The comparison of the results from these various observations, and their relation to quantities which can easily be predicted from theoretical models, are hampered by the variety of observing strategies employed. In order to achieve the high sensitivity required, it is necessary to perform a differential measurement; many different choices of beam geometry have been used. One possibility would be to allow the beam to drift across the sky and measure the r.m.s. fluctuations in the receiver output. This would then tell us the r.m.s. fluctuation in the sky temperature after this has been convolved with the beam pattern. In practice, however, a more common strategy is to repeatedly switch between two or more patches of sky and subtract the signal. If only two patches are observed, the output is simply the autocorrelation function of the (convolved) sky brightness. If, as is usual, the beam throw is comparable to the beam-width, then the output is a useful measure of the fluctuations on this angular scale. Such an apparatus would also have some sensitivity to fluctuations on much larger angular scales, because the apparatus measures the gradient of long-wavelength fluctuations. The most stringent current upper limits to small-scale anisotropy are those obtained by Uson and Wilkinson[41], using a double beam subtraction technique. They quote an upper limit $\Delta T/T \leq 2 \times 10^{-5}$ at $\theta \approx 4.5$ arc min, again with 95% confidence. This type of measurement is even less sensitive to large angular fluctuations because the beam geometry measures the second derivative of the sky pattern at large angles. Finally, we mention some searches for anisotropy on 'very small' angular scales (a few arc s), using the Very Large Array[50]. The sensitivity is somewhat poorer than that obtained on larger angular scales, with upper limits of 10^{-3}, 8×10^{-4} and 5×10^{-4} (95% confidence) at angular scales of 18, 30 and 60 arc s, respectively. These results have now been superseded by new, deeper survey at 6 cm, which improves on these limits by factors of 4 and 8 at 18 and 60 arc s, respectively[51]. Similar results have been obtained independently by H. M. Martin and R. B. Partridge (in preparation).

We also expect the microwave sky to display some linear polarization. To date, polarimetry searches have been carried out only at large angular scales[52,53], these searches having been stimulated by the suggestion[54] of a quadrupole polarization pattern in an anisotropic universe. Re-ionization can enhance this effect[55]. Polarization with a much smaller coherence angle will also arise due to inhomogeneity. The amplitude of the polarization pattern is of the order of the temperature anisotropy on a scale corresponding to the width of the last scattering shell[56]. If temperature anisotropy is ever detected, then polarization studies will provide an important test of the interpretation of such a result. It may be that temperature anisotropy searches will eventually be limited by confusion from faint sources. If this is the case, then polarimetry may provide the best way to view the primordial fluctuations.

Confidence intervals

The figures quoted above for upper limits on sky temperature fluctuations are a highly reduced form of the initial data, which typically consist of a set of N temperature difference estimates D_i together with estimates σ_i of the experimental uncertainty associated with each D_i. Before we consider the implications of the quoted upper limits it is important to understand the limitations of these statistics. To this end we will briefly discuss the method by which these confidence limits are arrived at and highlight some of the associated problems.

We will consider, for illustrative purposes, the simplified, though not entirely unrealistic example in which the D_i are measured in widely separated fields, and so can be considered to be statistically independent, and in which the instrumental variances are all equal: $\sigma_i^2 = \sigma_0^2$. In this case, the commonly adopted procedure[57] is to calculate the statistic $\chi^2 \equiv \Sigma D_i^2/\sigma_o^2$, and to ask, first, if there is firm evidence for fluctuations in excess of that produced by the receiver. Assuming that there is not, we then ask what is the value of σ_{ul}^2 such that, if the true sky r.m.s. temperature difference is $\sigma_{sky} = \sigma_{ul}$, a value of χ^2 as low as that observed will occur only a fraction $1-c$ of the time. The number of σ_{ul}^2 is our upper limit on the sky variance at confidence level c.

One problem with this method is that it requires a rather precise estimate of the instrumental variance σ_0 (ref. 58). For reasonably large N, the central limit theorem tells us that χ^2 will have a gaussian distribution with mean $N(\sigma_0^2+\sigma_{sky}^2)/\sigma_0^2$ and variance $2N(\sigma_0^2+\sigma_{sky}^2)^2/\sigma_0^2$. The upper limit is then $\sigma_{ul}^2 = \sigma_0^2\{(\chi_0^2/N)/(1-\alpha\sqrt{2}/N)-1\}$, where χ_0^2 is the value actually observed and α is the number of standard deviations associated with the confidence level c (for example, $\alpha = 1.65$ for $c = 95\%$). χ^2/N will be close to unity, $\chi^2/N = (1+O(N^{-1/2}))$, and so $\sigma_{ul}^2 \approx \sigma_0^2 N^{-1/2}$. But the statistic χ^2 is inversely proportional to the estimate of σ_0^2, so σ_{ul}^2 to be reliable, any systematic functional errors in σ_0^2 must be much smaller than $N^{-1/2}$. If we 'generously' overestimate σ_0^2, we will obtain much too low a value for σ_{ul}^2 and we will be in danger of excluding the true hypothesis.

Another problem is that, by our definition of σ_{ul}^2, we will reject the true hypothesis a fraction $1-c$ of the time. Thus, if 20 upper limits at 95% confidence are found in the literature than the 'best' upper limit is likely to lie below the true value. The actual situation will be worse than this if there is any selection pressure to favour the appearance of 'good' results. This might occur if there is any tendency, either of a voluntary or involuntary nature, not to publish results which are deemed to be uninteresting compared to other published results. Such selection effects cause the number of independent samples of event space to exceed the number of published results. The rate of sampling of event space can also be boosted by combining sets of data from different runs, or if the decision to extend or curtail the experiment is made after any data have been collected.

If for these or other reasons, the value of σ_{ul}^2 obtained has any influence on the publication of results, then the value of these constraints for rational decision making is debased. Ultimately, the only practical way to ensure that results are free of such bias is for the decision to publish or not to be made before the experiment is performed (with a blank space in the manuscript for σ_{ul}^2 to be inserted, of course). It is after all only rational that the criterion for acceptance of such a paper should be the sensitivity of the apparatus, rather than the random number σ_{ul}^2 which results at the end of the day.

Unreasonably restrictive upper limits may then occur by chance, or they may even be selected for. However, one should not simply accept this as a painful fact of life. Our procedure is guaranteed to reject the true hypothesis 5% of the time (for 95% confidence limits), but in any particular instance the data may give some indication as to whether this is such a case. For example, if $\sigma_{sky}^2 = 0$, then negative upper limits will occur. In such a case, we can be certain that we have been 'lucky'. There are other cases in which a positive limit results on the variance but in which one would be reluctant to reject the hypothesis that $\sigma_{sky}^2 \geq \sigma_{ul}^2$. These are those cases in which, while our statistic indicates that the result is unlikely under the hypothesis $\sigma_{sky}^2 > \sigma_{ul}^2$, the result is not significantly more likely under any alternative hypothesis. The problems associated with various statistical procedures for treating null results will be discussed elsewhere (A. N. Lazenby and N.K. in preparation), but for the present we note that the hallmark of such a spurious result is an unusually low value of χ^2. Visual inspection of the data presented in ref. 41 suggests this may have been the case in this instance. This is unfortunate because, taken at face value, this result rules out some interesting hypotheses.

Implications

Intermediate and large scales. The most immediate implication of these negative results is that we inhabit a universe that is currently uniform to a high degree of precision, at least on scales greater than the horizon size at last scattering. This inference is particularly certain because considerations of causality assure us that there is no mechanism which can render the Universe isotropic on scales exceeding the horizon scale at last scattering. Although we can observe only that region of space which is inside our present horizon, we can invoke the copernican principle to exclude such possibilities as that we inhabit an inhomogeneous universe which just happens to be isotropic about the Earth. We can also infer that the Universe is very close to homogeneous on scales much larger than the present horizon size[59]. If our universe is nearly homogeneous today, then it must have been prepared in a state with an extremely small amount of 'growing-mode' inhomogeneity present at early times; this preparation may have involved an inflationary phase[1]. The process which prepared the Universe may also have imprinted small departures from perfect homogeneity which could later grow to produce galaxies and other structure in the Universe. If this is the case, then it is most natural that the fluctuations should have curvature (specific binding energy) fluctuations which are independent of scale. Were this not the case, then it would be necessary to impose a special scale (that of curvature nonlinearity) in the initial conditions and this scale would not be many orders of magnitude different from the scale of galaxies. It is hard to see how the microphysics of the early Universe could produce such a scale.

On large scales the spectrum maintains its initial form. We can estimate the level of anisotropy which should be observed on intermediate and large angular scales, because the temperature fluctuations should be at least as large, to order of magnitude, as the specific binding energies of the largest structures we see today. Clusters have internal velocity dispersions of $v \approx 1000\ \mathrm{km\ s^{-1}}$ and so have specific binding energy $\phi \approx (v/c)^2 \approx 10^{-5}$. The binding energies of superclusters are thought to have similar magnitude and so we predict $\Delta T/T \approx 10^{-5}$. This is not much smaller than the current upper limits on these angular scales and so it seems that, with a modest improvement in the observational situation, we have the hope of either ruling out this picture or confirming the presence of these scale-free temperature fluctuations.

To obtain a quantitative constraint on $\Delta\rho/\rho$ on these scales, one can use the relation of Sachs and Wolfe[24] for the temperature fluctuation induced by a plane-wave density ripple. A convenient parameter to specify the amplitude of the perturbation is ϕ, the newtonian specific binding energy. The result of Sachs and Wolfe is simply that $\Delta T/T = \phi/3$. We have plotted the constraint on $(\Delta\rho/\rho)_\lambda$ using this relation, and the observed upper limits are shown in Figure 2. This constraint assumed that the initial fluctuation spectrum was initially isentropic.

Small scale. Observations on intermediate and large angular scales provide the best constraint on the amplitude of very-long-wavelength density fluctuations. On these scales, however, there is very little direct evidence for clustering of matter. On scales of $\leq 10h^{-1}$ Mpc we see strong clustering, at least of luminous material. The importance of small scale anisotropy studies is that they probe those length scales on which we actually see clustering, and so they provide a more direct test of theories of galaxy formation. Unfortunately, however, the interpretation of smallscale anisotropy is model-dependent: in order to apply this test one must first choose the nature of the initial fluctuations (for example, 'isothermal' or 'adiabatic'), assume some spectrum (such as a power law) and then propagate these fluctuations from some sufficiently early time until after the epoch at which the matter and radiation decouple. This linear evolution depends on the properties of the major constituents of the Universe at the time, so the abundances and masses of these particles must be specified. Up to this point the amplitude of the initial fluctuations is arbitrary. The next, and most uncertain, step is to normalize the calculation by requiring that the density fluctuations should have grown to the presently observed amplitude by today. One can then compare the predicted anisotropy with observations and thereby constrain the initial parameter space.

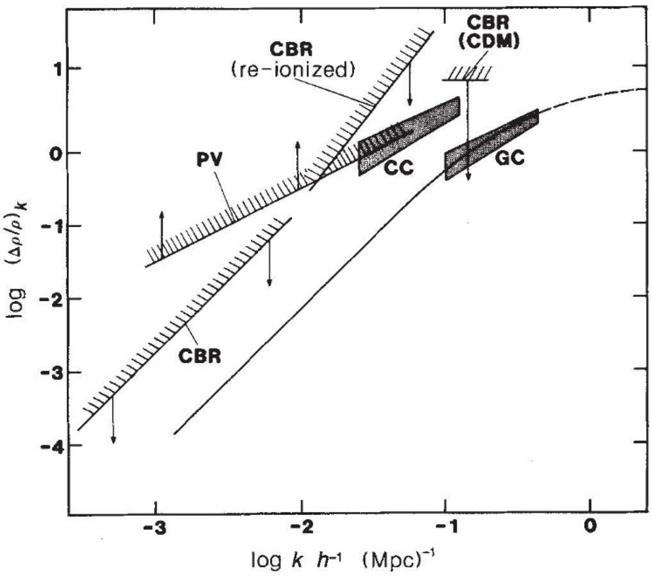

Fig. 2 Constraints on the power spectrum of density fluctuations in an Einstein-de Sitter ($\Omega=1$) universe. $(\Delta\rho/\rho)_k \equiv k^{3/2}\delta_k/\sqrt{2\pi}$. The contribution to the r.m.s. density fluctuations per logarithmic interval of k, is plotted versus k, where $k(\text{Mpc}^{-1})$ is comoving wave number, $\delta_k=\int d^3k\, e^{ik.x}\,\Delta\rho(x)/\rho$, and x labels the spatial coordinates of an arbitrary point. Specifically, the mass fluctuation within a sphere of radius x is $\delta M/M=\int k^2\, dk|\delta_k|^2 W(kx)/\sqrt{2\pi}$, where the window function $W(kx)\approx 1$ for $kx\leq 1$ and $W(kx)\approx (kx)^{-4}$ for $kx\gg 1$. CBR, large- and intermediate-scale microwave anisotropy; we have taken $\Delta T/T\leq 3\times10^{-5}$ for $\theta\leq 10$ arc min and $\Delta T/T\leq 10^{-4}$ for $\theta\approx$ a few degrees. CBR(CDM), small-scale anisotropy limits in a universe dominated by cold dark matter. CBR(re-ionized), anisotropy limits if the intergalactic medium is re-ionized at large redshift and secondary fluctuations are generated as discussed in the text (see discussion of 'the re-ionized universe'). PV, constraint from peculiar velocity; power spectrum of galaxy clustering; CC, power spectrum of cluster clustering; CDM, theoretical preduction for density fluctuations in a cold dark matter model with $h=0.5$. In a low-density universe ($\Omega=0.2$), large- and intermediate-scale anisotropy limits are ~3 times lower; peculiar velocity limits are ~3 times higher; small-scale anisotropy limits are ~10 times lower.

In recent years, there have been interesting theoretical and observational developments concerning the small-scale anisotropy. On the theoretical side there has been a rise in popularity of models in which the bulk of the gravitating matter is in a form which is otherwise very weakly coupled to other forms of matter. These models generate a much smaller level of anisotropy than models in which the density is dominated by baryons. Meanwhile, the observational limits have improved, and now not only rule out baryon-dominated models (or at least those with 'adiabatic' fluctuations) but also place useful constraints on models involving weakly interacting 'dark' matter.

The normalization that has been used in these calculations is that the density contrast in spheres of radius $8h^{-1}$ Mpc should have unit variance today. This normalization is chosen because the fluctuations in galaxy counts in spheres of this size have unit variance and because it is thought that fluctuations on these scales should be described fairly accurately by linear theory, and so should be accessible to theoretical calculation. Ideally, it would be desirable to normalize on even larger scales where linear theory would be even more accurate, but unfortunately the estimates of fluctuations in galaxy counts on larger scales become very uncertain.

With this choice of normalization, the small-scale anisotropy limit allows one to exclude 'cold dark matter' models with $\Omega=0.2$ (see the constraint labelled CBR (CDM) in Fig. 2). This conclusion is particularly interesting because this is the value for Ω implied if galaxies fairly trace the mass distribution. If the value of Ω is significantly greater than 0.2, then galaxies are, for some reason, more strongly clustered than the matter. If this is the case, then it is inconsistent to use the conventional normalization. Inclusion of this effect would lower the predicted $\Delta T/T$ still further in addition to the approximate Ω^{-1} scaling (by a factor of ~2)[60-62]; to detect such tiny fluctuations would require a great improvement over the current upper limits.

The possibility of ruling out a high-density Universe dominated by cold dark matter from small-angle anisotropy is therefore fairly remote. Large-angular-scale searches present a more hopeful prospect, although in this case the interpretation depends crucially on the assumed slope of the primordial spectrum of fluctuations. As we have mentioned, there are reasons to prefer the Zel'dovich, scale-free curvature spectrum. To rule out, or to confirm at the several-sigma level, the existence of the consequent scale-free temperature fluctuations would be a significant achievement.

Extrinsic anisotropy

The microwave sky has now been almost completely mapped at an angular resolution of a few degrees. The only convincing feature which has been observed to date is the 'dipole' anisotropy. Averaging together the two most precise of the recent measurements yields a dipole anisotropy with amplitude $\delta T/T=1.2\times10^{-3}$. The usual interpretation of this anisotropy is that the Sun is moving at a velocity $v_{\text{solar}}\approx 380\text{ km s}^{-1}$ towards an apex ~45° away from the Virgo cluster. The reason for this interpretation, rather than as a combination of extrinsic and intrinsic anisotropy, is that this dipole is at best an order of magnitude larger than the quadrupole or other low-order spherical harmonics on the microwave sky, and such an intrinsic pattern seems highly implausible. It may be possible to test this interpretation if the intrinsic quadrupole is really very small because, in addition to the dipole of amplitude v/c, our motion will also result[63] in an aligned quadrupole of amplitude v^2/c^2.

This measurement of the solar motion provides us with a constraint on large-scale density fluctuations because, if these induce an r.m.s. amplitude of $\gg 300\text{ km s}^{-1}$, we must live in a very special position[64,65]. Unless there is some anthropic argument to favour the existence of observers in regions which seem to be special only in so far as they have an unusually low dipole anisotropy, we can exclude such hypotheses as unlikely. These hypothetical large-scale density fluctuations should be well described by linear theory, and we then have $v_{\text{rms}}\approx HL(\Delta\rho/\rho)_{\text{rms}}$, for fluctuations on length scale L. If we assume for the moment that the density fluctuations are a gaussian process, then the large-scale velocities will also be gaussian-distributed. We then have an upper limit at 95% confidence of $v_{\text{rms}}\leq 3v_{\text{solar}}$ as only ~5% of all observers have a velocity less than one third of the r.m.s. value. We have assumed a gaussian distribution, but this result is not very sensitive to the form of the tails of the actual distribution because we are calculating the probability of very small values of v/v_{rms}. The result would be misleading, however, if the probability $P(v)\neq v_{\text{rms}}^{-3}$ for $v\ll v_{\text{rms}}$, as would be the case if a small fraction of observers have $v\gg v_{\text{rms}}$ while most are nearly at rest. Note that we have ignored contributions to v_{solar} from small scales which are nonlinear ($\Delta\rho/\rho\gtrsim 1$). This is legitimate if we have no detailed knowledge of the form of these velocities, as the inclusion of small-scale $\Delta\rho/\rho$ will only broaden the distribution $P(v)$. But as we do have some knowledge of the Sun's motion in the Galaxy and the Galaxy's motion in the Local Group, we should incorporate this information. In doing so, we will be able to 'filter out' the nonlinear velocities and concentrate on those linear velocities of interest.

The first step in this process is to allow for the rotation of the Galaxy. This results in a larger velocity, $v_{\text{galaxy}}\approx 550\text{ km s}^{-1}$, and therefore a weaker constraint than we would have if we had no knowledge of the galactic rotation. The reason for this is that the galactic rotation correction is almost parallel to the solar motion and is therefore unlikely (it means that we have one of the lowest dipole anisotropies of all observers in the Galaxy). However unlikely this state of affairs is *a priori*, it is hard to imagine that the galactic rotation estimate is wildly wrong, and so we must fully incorporate this information. The next step is to subtract the motion of the Galaxy with respect to other galaxies in the vicinity. Here, as we shall see, the situation is not so clear-cut. As before, the *a priori* expectation is that the velocity will decrease. However, as with the galactic rotation we should prepare ourselvles for possible surprises.

We determine the galactic motion with respect to a sample of galaxies around us by means of a 'Tully-Fisher' or 'Faber-Jackson' relation, or some analogous intrinsic correlation between a distance-dependent and a distance-independent quality. These 'L-σ' relations, or analogous relations, are established for a set of galaxies which we believe to be at the same distance from us (for example, in a cluster of galaxies). If we then assume that the correlation found is universal, this allows us to correct redshifts to true distances in a statistical manner by minimizing residuals. The validity of a velocity thereby obtained depends crucially on the assumed universality of the relation. If there are systematic differences in the zero-point of the relations from cluster to cluster, then these will be misinterpreted as peculiar velocities. An effect of this kind must be present at some level. In principle one can test for this by applying the method to clusters at a wide range of distances. Any spurious velocities will increase linearly with distance. These effects, if present at a significant level, do not however present any fundamental limit to the depth of sample for which one can determine the net velocity, as the number of clusters increases sufficiently quickly that the errors remain manageable.

The important question, for our purpose, is whether the source of our peculiar velocity has been localized within the sample. If so, we obviously get a stronger constraint on $\Delta\rho/\rho$ at large scales. It seemed, following the publication by Hart and Davies[66] that this was indeed the case. They studied a sample of galaxies with $v \leq 5{,}000$ km s^{-1} using a radio variant of the Tully-Fisher (TF) technique, and found that the net motion of these galaxies with respect to the MBR was very small: $v_{HD} = 130 \pm 70$ km s^{-1}. This result stands in contrast to the earlier result by Rubin *et al.*[67] for a sample of similar depth, although doubts about this result had already been expressed[68,69]. However, the Hart and Davies sample has now been reanalysed[70], and yields a much greater net motion of ≥ 700 km s^{-1}. An independent sample of similar depth was analysed by de Vaucouleurs and Peters[71], who found an intermediate value, $v_{dV} \approx 300$ km s^{-1}. More recently, Collins *et al.*[72] and Burstein *et al.*[73,74] have found net streaming velocities of ~700 km s^{-1} using slightly deeper samples, and are in approximate agreement with the original result of Rubin *et al.*[67]. However, these results had been preceded by those of Aaronson *et al.*[75], who used the infrared TF technique to analyse a sample of galaxies lying in the declination strip visible from Arecibo. They found a net motion of ≤ 200 km s^{-1}, although the geometry of their sky coverage results in increased uncertainty for a velocity vector in the direction favoured by Burstein *et al.*[73,74] and by Collins *et al.*[72]. Another interesting result to emerge from the Aaronson *et al.*[75] sample is that the scatter of individual cluster velocities about the Hubble flow (or of any offsets in the TF relation interpreted as velocities) is very small. They obtained a limit of ~200 km s^{-1} on the amplitude of such motions, roughly consistent with the individual cluster motions of ~400 km s^{-1} (for a different sample) inferred by Burstein *et al.*[75], and suggesting that the infrared TF technique has very low intrinsic scatter.

The differences between the solutions found, both for net motion and for individual cluster motions, give some measure of the uncertainty associated with these methods. Certainly, the question of whether the source of the galactic motion has been localized remains in doubt. It is probably premature to try to use these data to constrain $\Delta\rho/\rho$ on very large scales. The most reliable datum we have is for the galactic motion of ~550 km s^{-1} with respect to the microwave frame, and this results in the line labelled PV in Fig. 2. If a consistent picture for large streaming motions does emerge then this will imply, according to standard theory, sizeable density fluctuations ($\gtrsim 30\%$) on a scale of $\gtrsim 100h^{-1}$ Mpc. Such density fluctuations present a challenge to theorists: for instance, it appears that such fluctuations do not arise naturally in the cold dark matter model. A theory in which non-gaussian fluctuations emerge naturally, such as cosmic strings combined with hot dark matter, offers some prospect of simultaneously allowing large streaming motions and low $\delta T/T$. It should someday be possible to directly measure, or possibly exclude, such density fluctuations with large-scale redshift surveys.

Conclusions

The microwave background isotropy limits, as well as the measured dipole anisotropy and large-scale galaxy distribution, have already succeeded in severely constraining inflationary cosmological models. Indeed, the only survivors require an appeal to rather more contrived models than discussed here. One possibility requires a biasing scheme in which the observed galaxy distribution samples the gravitational effects of a universe with $\Omega = 0.1$, while the true Ω is unity[14,61]. Biasing schemes generally predict that ~90% of the Universe should contain 'failed galaxies', with the baryonic matter forming luminous galaxies only in the rare fluctuation peaks that sample ~10% of the Universe[76,77]. One alternative involves appeal to the vacuum energy density (or cosmological constant Λ): with $\Omega_{vac} = 0.2$, where $\Lambda \equiv \Omega_{vac}$, a combination of cosmological constant and cold dark matter can reduce both $\delta T/T$ and the large-scale peculiar velocity field to acceptable levels, while allowing light to be a good tracer of mass[78]. Another option is to let the cold dark matter (with $\Omega = 1$) decay into relativistic weakly interacting particles at a recent epoch (after galaxies have formed), thereby allowing astronomical measurements today to effectively sample a low-Ω universe[79,80]. This scheme apparently fails to produce flat galactic rotation curves[87] and generates an unacceptable feature in the galaxy correlation function[88], but it is consistent with $\delta T/T$ limits[81]. One could also revive hot dark matter in some non-baryonic guise such as the massive neutrino, with nonlinearity of the hot dark matter at a very recent ($z \approx 1$) epoch, as required by the various constraints such as $\xi(r)$ and the pairwise galaxy peculiar velocity distribution, but now seeding the Universe with large-amplitude small-scale fluctuations (cold dark matter, cosmic strings?) in order to allow some, if not all, galaxies to form at an earlier epoch[7,82,83]. Finally, the wakes produced by cosmic strings could themselves generate large sheet-like structures which might dominate the large-scale galaxy distribution[84-86]. String models naturally yield low (but ultimately observable) $\delta T/T$.

It is of course conceivable that we are completely on the wrong track. Perhaps the primordial fluctuation spectrum was not described by a scale-invariant power law, nor was it gaussian, nor even adiabatic. One might then resort to explosive amplification of primordial seeds to explain galaxy formation and leave little trace of the primordial initial conditions. This still poses the problem, however, of the origin of the large-scale clustering of the galaxies. We do observe large-scale correlations (superclusters and voids): surely these trace initial conditions. If our interpretation of the CBR in the standard Big Bang model is correct, and there is no real alternative, then we inevitably expect to see some residue of these initial conditions in the background radiation anisotropy. It would be too perverse of Nature to have thrown down an impenetrable screen which renders such fluctuations invisible. Eventual detection of $\delta T/T$ on some angular scale is inevitable, and it will surely elucidate one of the most challenging mysteries of the Big Bang theory, namely the origin and the formation of large-scale structure.

1. Guth, A. H. *Phil. Trans. R. Soc.* **A307**, 141-147 (1982).
2. Hawking, S. W. *Phys. Lett.* **115B**, 295-297 (1982).
3. Starobinskii, A. A. *Phys. Lett.* **117B**, 175-178 (1982).
4. Guth, A. & Pi, S.-Y. *Phys. Rev. Lett.* **49**, 1110-1113 (1982).
5. Bardeen, J. M., Steinhardt, P. J. & Turner, M. S. *Phys. Rev.* **D28**, 679-693 (1983).
6. Linde, A. *Phys. Lett.* **129B**, 177-181 (1983).
7. Silk, J. & Turner, M. *Phys. Rev.* **D.** (in the press).
8. Ostriker, J. P. & Cowie, L. *Astrophys. J.* **243**, L127-L132 (1981).
9. Ikeuchi, S. *Publs astr. Soc. Japan* **33**, 211-222 (1981).
10. Zel'dovich, Ya. B. *Mon. Not. R. astr. Soc.* **192**, 663-667 (1980).
11. Vilenkin, A. *Phys. Rep.* **121**, 263-315 (1985).
12. Wyse, R. G. & Jones, B. J. *Astr. Astrophys.* **149**, 144-150 (1985).
13. Blumenthal, G., Faber, S., Primack, J. & Rees, M. *Nature* **311**, 517-525 (1984).
14. Davis, M., Efstathiou, G., Frenck, C. S. & White, S. D. M. *Astrophys. J.* **292**, 371-394 (1985).
15. Frenk, C., White, S. D. M. & Davis, M. *Astrophys. J.* **271**, 417-430 (1983).
16. Efstathiou, G. & Silk, J. *Fundam. Cosmic Phys.* **9**, 1-138 (1983).
17. Bond, J. R. & Efstathiou, G. *Astrophys. J.* **285**, L45-L49 (1984).
18. Vittorio, N. & Silk, J. *Astrophys. J.* **285**, L39-L44 (1984).
19. Peebles, P. J. E. & Groth, E. *Astr. Astrophys.* **53**, 131-140 (1976).
20. Peebles, P. J. E. & Yu, J. T. *Astrophys. J.* **162**, 816-836 (1970).
21. Wilson, M. L. & Silk, J. *Astrophys. J.* **243**, 14-25 (1981).
22. Silk, J. *Nature* **215**, 1155-1156 (1967).
23. Silk, J. *Astrophys. J.* **151**, 459-471 (1968).
24. Sachs, R. K. & Wolfe, A. M. *Astrophys. J.* **147**, 73-90 (1967).
25. Peebles, P. J. E. *Astrophys. J.* L1-L5 (1982).
26. Bond, J. R. & Szalay, A. S. *Astrophys. J.* **274**, 443-468 (1983).
27. Zel'dovich, Ya. B. *Astr. Astrophys.* **5**, 84-89 (1970).
28. Efstathiou, G. & Bond, J. R. *Mon. Not. R. astr. Soc.* **218**, 103-121 (1986).
29. Suto, Y., Sato, K. & Kodama, H. *Astrophys. J.* **292**, L1-L4 (1985).
30. Fixen, D. J., Cheng, E. S. & Wilkinson, D. T. *Phys. Rev. Lett.* **50**, 620-622 (1983).
31. Kaiser, N. *Astrophys. J.* **282**, 374-381 (1984).
32. Hogan, C. *Mon. Not. R. astr. Soc.* **202**, 1101-1126 (1983).
33. Hogan, C. J. & Kaiser, N. *Astrophys. J.* **274**, 7-19 (1983).
34. Vilenkin, A. *Phys. Rep.* **131**, 263-315 (1985).
35. Kaiser, N. & Stebbins, A. *Nature* **310**, 391-393 (1984).
36. Gott, J. R. *Astrophys. J.* **288**, 422-427 (1985).
37. Brandenberger, R. H. & Turok, N. *Phys. Rev.* **D** (in the press).
38. de Lapparent, V., Geller, M. J. & Huchra, J. P. *Astrophys. J.* **302**, L1-L5 (1986).
39. Szalay, A., Bond, J. R. & Silk, J. in *Formation and Evolution of Galaxies and Large Structures in the Universe* (eds Audouze, J. & Tran Thanh Van, J.) 101-128 (Reidel, Dordrecht, 1984).
40. Hogan, C. J. *Astrophys. J.* **284**, L1-L4 (1984).
41. Uson, J. M. & Wilkinson, D. T. *Nature* **312**, 427-429 (1984).
42. Ostriker, J. & Vishniac, E. *Astrophys. J.* **306**, L51-L54 (1986).
43. Penzias, A. A. & Wilson, R. W. *Astrophys. J.* **142**, 419-421 (1965).
44. Melchiorri, F., Melchiorri, B., Ceccarelli, C. & Pietranera, L. *Astrophys. J.* **250**, L1-L4 (1981).
45. Fabbri, R., Guidi, I., Melchiorri, F. & Natale, V. *Phys. Rev. Lett.* **44**, 1563-1566 (1980).
46. Lubin, P., Villela, T., Epstein, G. & Smoot, G. *Astrophys. J.* **298**, L1-L5 (1985).
47. Smoot, G. F., Gorenstein, N. V. & Muller, R. A. *Phys. Rev. Lett.* **39**, 898-901 (1976).
48. Strukov, I. A. & Skulachev, D. P. *Soviet Astr. Lett.* **10**, 3-13 (1984).
49. Mandolesi, N. *et al. Nature* **319**, 751-753 (1986).
50. Fomalont, E. B., Kellermann, K. I. & Wall, J. V. *Astrophys. J.* **277**, L23-L26 (1984).
51. Kellerman, K. I., Fomalont, E. B., Weistrop, D. & Wall, J. *Highlights in Astronomy* (Reidel, Dordrecht, in the press).
52. Lubin, P. M. & Smoot, G. F. *Astrophys. J.* **245**, 1-17 (1981).
53. Lubin, P. M., Melese, P. & Smoot, G. F. *Astrophys. J.* **273**, L51-L54 (1983).
54. Rees, M. J. *Astrophys. J.* **153**, L1-L5 (1986).
55. Negroponte, J. & Silk, J. *Phys. Rev. Lett.* **44**, 1433-1437 (1980).
56. Hogan, C. J., Kaiser, N. & Rees, M. J. *Phil. Trans. R. Soc.* **A307**, 97-109 (1982).
57. Boynton, P. E. & Partridge, R. B. *Astrophys. J.* **181**, 243-253 (1973).
58. Lazenby, A. N. & Davies, R. D. *Mon. Not. R. astr. Soc.* **203**, 1137-1169 (1983).
59. Grishchuk, L. P. & Zel'dovich, Ya. B. *Sov. Astr. Lett.* **22**, 125-129 (1978).
60. Bardeen, J. M. in *Proc. Inner Space-Outer Space Symp.* (eds Kolb, E. W., Turner, M. S., Lindley, D., Olive, K. & Seckel, D.) 212-217 (University of Chicago Press, 1986).
61. Kaiser, N. in *Proc. Inner Space-Outer Space Symp.* (eds Kolb, E. W., Turner, M. S., Lindley, D., Olive, K. & Seckel, D.) 258-263 (University of Chicago Press, 1986).
62. Bardeen, J. M., Bond, J. R., Kaiser, N. & Szalay, A. *Astrophys. J.* **304**, 15-61 (1986).
63. Sunyaev, R. A. & Zel'dovich, Ya. B. *Mon. Not. R. astr. Soc.* **190**, 413-420 (1980).
64. Kaiser, N. *Astrophys. J.* **273**, L17-L20 (1983).
65. Vittorio, N. & Silk, J. *Astrophys. J.* **293**, L1-L4 (1985).
66. Hart, L. & Davies, R. D. *Nature* **297**, 191-196 (1982).
67. Rublin, V. C., Thonnard, N., Ford, W. K. & Robert, M. S. *Astr. J.* **81**, 719-737 (1976).
68. Fall, S. M. & Jones, B. J. T. *Nature* **262**, 547-460 (1976).
69. Schechter, P. L. *Astr. J.* **82**, 569-576 (1977).
70. Stavely-Smith, L. thesis, Univ. Manchester (1985).
71. de Vaucouleurs, G. & Peters, W. L. *Astrophys. J.* **287**, 1-16 (1984).
72. Collins, C. A., Joseph, R. D. & Robertson, N. A. *Nature* **320**, 506-508 (1986).
73. Burstein, D. *et al.* in *Galaxy Distances and Deviations From the Hubble Scale* (eds Madore, B. & Tully, R. B.) (Reidel, Dordrecht, in the press).
74. Lynden-Bell, D. *Q. Jl R. astr. Soc.* **27**, 319-327 (1986).
75. Aaronson, M. *et al. Astrophys. J.* **302**, 536-563 (1986).
76. Silk, J. *Astrophys. J.* **207**, 1-8 (1985).
77. Rees, M. J. *Mon. Not. R. astr. Soc.* **213**, 75P-81P (1985).
78. Peebles, P. J. E. *Astrophys. J.* **284**, 439-444 (1985).
79. Turner, M., Steigman, G. & Krauss, L. *Phys. Rev. Lett.* **52**, 2090-2093 (1984).
80. Gelmini, G., Schramm, D. & Valle, J. *Phys. Lett.* **146B**, 311-317 (1984).
81. Vittorio, N. & Silk, J. *Phys. Rev. Lett.* **54**, 2269-2272 (1985).
82. Achulli, S., Occhionero, F. & Scaramella, R. *Astrophys. J.* **299**, 577-582 (1985).
83. Umemusa, M. & Ikeuchi, S. *Astrophys. J.* **299**, 583-592 (1985).
84. Silk, J. & Vilenkin, A. *Phys. Rev. Lett.* **53**, 1700-1703 (1984).
85. Vachaspati, T. *Phys. Rev. Lett.* **57**, 1655-1657 (1986).
86. Brandenberger, R., Stebbins, A., Turok, N., Veeraraghavan, S. & Silk, J. *Astrophys. J.* (submitted).
87. Flores, R., Blumenthal, G. R., Dekel, A. & Primack, J. R., preprint SCIPP 86/64 (1986).
88. Veeraraghavan, S., Suto, Y. & Silk, J., *Astrophys. J.* (submitted).

ARTICLES

A molecular gradient in early *Drosophila* embryos and its role in specifying the body pattern

Paul M. Macdonald* & Gary Struhl*

Department of Biochemistry and Molecular Biology, Harvard University, 7 Divinity Ave, Cambridge, Massachusetts 02138, USA

After fertilization, the protein products of the Drosophila *homeobox gene* caudal (cad) *accumulate in a concentration gradient spanning the anteroposterior axis of the developing embryo. Mutations in the* cad *gene that reduce or eliminate the gradient cause abnormal zygotic expression of at least one segmentation gene* (fushi tarazu) *and alter the global body pattern.*

ONE recurring dilemma in developmental biology is the conflict between mosaic models of embryonic patterning which invoke rigid localizations of qualitatively distinct determinants, and regulative models, which postulate morphogen gradients or inductive cascades of cell-cell interactions. In the case of insect development, the classic experiments of Sander[1-4] and others[5-9] have ruled out mosaic models, except possibly for the segregation of the germ line[10], and argued strongly for anteroposterior gradients of morphogens that specify the body pattern. Further support for such a gradient model has come from the analysis of maternal effect mutations, such as *bicaudal*, *oskar* and *bicoid*, which alter the global segment pattern[11-18]. But concrete demonstrations of graded morphogens have been frustratingly absent.

* Present address: Howard Hughes Medical Institute, Center for Neurobiology and Behaviour, Columbia University College of Physicians and Surgeons, 722 West 168 Street, New York, New York 10032.

During the past few years, molecular studies have led to the identification of a conserved structural domain, the homeobox, common to many genes which control segment development[19,20]. In general, homeobox containing genes have tightly restricted spatial roles[21-25]; hence, they are first activated in specific regions of the embryo around the time segments are established[19,26-33]. Recently, we (unpublished observations) and others[34,35] have identified a new homeobox gene, *caudal* (*cad*)[34], which has the unusual property of being expressed during oogenesis and early embryogenesis. Hence, it seemed possible that this atypical gene might have a more global role in organizing the segment pattern. Indeed, circumstantial evidence for such a role has been obtained by Mlodzik *et al.*[34] and Levine *et al.*[35] who observed that *cad* transcripts form an anteroposterior concentration gradient during the syncytial blastoderm stage when the segment pattern is being set up.

We have used antibodies directed against the *cad* protein, together with mutations in the *cad* gene, to examine its develop-