On the relation between the isotropy of the CMB and the geometry of the universe

RASANEN, Syksy

Abstract

The near-isotropy of the cosmic microwave background (CMB) is considered to be the strongest indication for the homogeneity and isotropy of the universe, a cornerstone of most cosmological analysis. We derive new theorems which extend the Ehlers-Geren-Sachs result that an isotropic CMB implies that the universe is either stationary or homogeneous and isotropic, and its generalisation to the almost isotropic case. We discuss why the theorems do not apply to the real universe, and why the CMB observations do not imply that the universe would be nearly homogeneous and isotropic.

Reference


DOI : 10.1103/PhysRevD.79.123522
On the relation between the isotropy of the CMB and the geometry of the universe

Syksy Räisänen
Université de Genève, Département de Physique Théorique,
24 quai Ernest-Ansermet, CH-1211 Genève 4, Switzerland

The near-isotropy of the cosmic microwave background (CMB) is considered to be the strongest indication for the homogeneity and isotropy of the universe, a cornerstone of most cosmological analysis. We derive new theorems which extend the Ehlers-Geren-Sachs result that an isotropic CMB implies that the universe is either stationary or homogeneous and isotropic, and its generalisation to the almost isotropic case. We discuss why the theorems do not apply to the real universe, and why the CMB observations do not imply that the universe would be nearly homogeneous and isotropic.

PACS numbers: 04.20.-q, 04.40.-b, 98.80.-k, 98.80.Jk

I. INTRODUCTION

A central assumption in most cosmological analysis is that the universe is homogeneous and isotropic up to small perturbations, and thus described by a perturbed Friedmann-Robertson-Walker (FRW) model. The strongest observational indication of isotropy comes from the temperature of the cosmic microwave background (CMB), which is uniform at the 10^{-5} level in different directions (excepting the dipole of 10^{-3}) [1, 2].

The important question is how the CMB temperature is related to the spacetime geometry. The first step towards constraining the geometry without assuming that it is FRW was taken by Ehlers, Geren and Sachs, who proved that if the only matter is a radiation fluid with a perfectly isotropic distribution function, the spacetime is either stationary, FRW, or a special case with non-zero rotation and acceleration [3]. (This has been extended to arbitrary matter with an isotropic distribution [4].) The result was generalised by Stoeger, Maartens and Ellis, who showed that if the CMB temperature and its derivatives are almost isotropic everywhere in an expanding dust-dominated universe, and the observers are geodesic, the spacetime is almost FRW [5].

Counterexamples which violate some of the assumptions of [5] have been presented [6]. For cosmology, the main issue is the applicability of the theorems to realistic models of the present-day universe, briefly discussed in [5]. We first derive new theorems which relate the CMB temperature to the geometry of the universe in the case of perfect isotropy and in the case of small anisotropy, generalising the results of [3, 5]. We then discuss why neither the new nor the old theorems apply to the real universe, and how the CMB observations only indicate that the universe is statistically homogeneous and isotropic, not that it would be locally almost exactly homogeneous and isotropic (i.e. almost FRW).

II. THE COVARIANT FORMALISM

The four-velocity. We are interested in the CMB temperature measured by observers moving on timelike curves. For reviews of the covariant approach we use, see [4, 5]. The observer velocity is denoted \( u^\alpha \), and normalised as \( u_\alpha u^\alpha = -1 \). The derivative with regard to the proper time measured by the observers is given by \( \partial_\tau \equiv u^\alpha \nabla_\alpha \) and also denoted by an overdot. The tensor which projects on the space orthogonal to \( u^\alpha \) is \( h_{\alpha\beta} \equiv g_{\alpha\beta} + u_\alpha u_\beta \), where \( g_{\alpha\beta} \) is the metric. The derivative projected with \( h_{\alpha\beta} \) is denoted by \( \tilde{\nabla}_\alpha \equiv h^{\beta}_{\alpha} \nabla_\beta \).

The covariant derivative of \( u^\alpha \) can be decomposed as

\[
\nabla_\beta u_\alpha = \frac{1}{3} h_{\alpha\beta} \theta + \sigma_{\alpha\beta} + \omega_{\alpha\beta} - \dot{u}_\alpha u_\beta ,
\]

where \( \theta \equiv \nabla_\alpha u^\alpha \) is the volume expansion rate, \( \sigma_{\alpha\beta} \) is the traceless symmetric shear tensor, \( \omega_{\alpha\beta} \equiv \nabla_\beta u^\alpha + u_\beta \dot{u}^\alpha - \dot{u}_\alpha u_\beta \) is the vorticity tensor and \( \dot{u}^\alpha \) is the acceleration vector. The tensors \( \sigma_{\alpha\beta} \) and \( \omega_{\alpha\beta} \) and the vector \( \dot{u}^\alpha \) are orthogonal to \( u^\alpha \). Instead of \( \omega_{\alpha\beta} \), we can equivalently use the vorticity vector \( \omega^\alpha \equiv \frac{1}{2} \epsilon^{\alpha\beta\gamma} \omega_{\beta\gamma} \), where \( \epsilon^{\alpha\beta\gamma} \equiv \eta_{\alpha\beta\gamma} u^\delta \) is the volume element in the space orthogonal to \( u^\alpha \), \( \eta_{\alpha\beta\gamma} \) being the spacetime volume element.

The temperature. In the geometrical optics approximation, light travels on null geodesics. The null geodesic tangent vector, identified with the photon momentum, is denoted by \( k^\alpha \). It satisfies \( k_\alpha k^\alpha = 0 \) and \( k^\alpha \nabla_\alpha k^\beta = 0 \). The energy is the projection of the momentum onto the observer’s velocity,

\[
E = -u_\alpha k^\alpha .
\]

The photon momentum can be decomposed into an amplitude and the direction, and the direction can be split into components orthogonal and parallel to \( u^\alpha \),

\[
k^\alpha = E (u^\alpha + e^\alpha) ,
\]

where \( u_\alpha e^\alpha = 0 \), \( e_\alpha e^\alpha = 1 \).

In [7], the temperature was defined with the integrated brightness \( I \equiv \int_0^\infty dE E^3 f(E, x, e) \propto T^4 \), where \( f(E, x, e) \) is the phase space distribution function of the CMB photons. However, observations of the anisotropy of the brightness are only made in some small energy ranges around a few frequencies [1, 2]. To avoid assumptions about the spectrum at unobserved energies, rather than considering \( I \), we define the temperature using the
spectrum in the observed energy range. There the CMB distribution function has the blackbody shape to an accuracy of $10^{-4}$ \cite{10}, $f = 2(E/E(T(x,e)) - 1)^{-1}$. We can invert $f = F(E/T(x,e))$ to obtain $E(x,e,f) = T(x,e)F^{-1}(f)$. Neglecting interactions with matter after last scattering, the photons are essentially collisionless, so the distribution is conserved, $\frac{dE}{dv} = 0$, where $\frac{dE}{dv}$ is the derivative along the photon flow line in phase space. It follows that $\frac{dE}{dv} = \frac{dE}{dv} = k^a \nabla_a \ln E$, where the last derivative is on the spacetime manifold. It is therefore sufficient to consider the photon energy $E(x,e,f)$. We assume that $E$ is an analytic function. Decomposing $k^a \nabla_a E$ using (1), (2) and (3) on the one hand, and writing $\frac{k^a \nabla_a \theta}{\theta}$, where $\theta$, $\pi$, $h$ is on the spacetime manifold. It is therefore sufficient to consider the spacetime geometry. The behaviour of these quantities to the rest of the geometry.

III. THEOREMS

Isotropy theorem. We prove the following result about the implication of a perfectly isotropic CMB for the spacetime geometry:

Theorem 1. If geodesic observers at all $x$ measure photon energy $E(x,e,f)$ that is independent of the direction
e, the spacetime is either stationary or obeys the conditions (15), (14). If the anisotropic stress is zero (in particular, if the matter is an ideal fluid), (15), (14), reduce to the statement that the spacetime is FRW.

Assuming the observers to be geodesic implies \( \dot{u}_a = 0 \). Since \( E \) does not depend on \( e \) and harmonic functions of different degree are independent, it follows from (14) that
\[
e^\alpha \partial_\alpha E = 0, \quad \sigma_{\alpha\beta} = 0. \tag{18}
\]
Since \( e^\alpha \) is an arbitrary direction orthogonal to \( u^\alpha \), \( e^\alpha \partial_\alpha E = 0 \) is equivalent to \( \nabla_\alpha E = 0 \), which can be written as \( \nabla_\alpha E = -u_\alpha E \) by using the definition of \( h_{\alpha\beta} \). Now there are two possibilities. Either \( \dot{E} = 0 \) or \( \dot{E} \neq 0 \).

If \( \dot{E} = 0 \), it follows from (14) that \( \theta = 0 \). Together with \( \dot{u}_a = 0 \) and \( \sigma_{\alpha\beta} = 0 \), this is equivalent to \( \nabla_\alpha u_{\beta} = 0 \), so \( u^a \) is a Killing vector, and the spacetime is stationary. The spacetime is static if and only if \( u^a \) is also hypersurface orthogonal, which is equivalent to \( \omega_{\alpha\beta} = 0 \). These results hold independently of the matter content and the Einstein equation.

If \( \dot{E} \neq 0 \), we have \( \theta \neq 0 \). We can write \( u_\alpha = -\dot{E}^{-1} \nabla_\alpha E \). This (together with \( \dot{u}_a = 0 \)) means that the vorticity tensor has the form \( \omega_{\alpha\beta} = u_\alpha e_\beta \). It then follows from \( u^\alpha \omega_{\alpha\beta} = 0 \) that \( \omega_{\alpha\beta} = 0 \). (This is an application of Frobenius’ theorem (11).) Given \( u_\alpha = 0 \), the condition \( \nabla_\alpha E = 0 \) implies \( \nabla_\alpha E = 0 \), so from (14) we get \( \nabla_\alpha \theta = 0 \). From (3) we have \( H_{\alpha\beta} = 0 \). These results hold independently of the matter content and the Einstein equation. Assuming that the Einstein equation holds, the evolution equations (3–17) reduce to the following:
\[
\omega_{\alpha\beta} = 0, \quad \nabla_\alpha \theta = 0, \quad H_{\alpha\beta} = 0
\]
\[
\rho + \theta (\rho + p) = 0, \quad \dot{\theta} + \frac{1}{3} \theta^2 = -\frac{1}{2} (\rho + 3p)
\]
\[
\nabla_\alpha (\rho + 3p) = 0
\]
\[
\nabla_\beta \pi^\beta_\alpha = -\nabla_\alpha p, \quad \pi_{\alpha\beta} + \frac{2}{3} \theta \pi_{\alpha\beta} = 0
\]
\[
E_{\alpha\beta} = \frac{1}{2} \pi_{\alpha\beta}, \quad q_{\alpha} = 0. \tag{19}
\]
It follows that if the anisotropic stress is zero, the spacetime is FRW. In particular, this is true for an ideal fluid. The result that for an ideal fluid the conditions \( \dot{u}_a = 0, \sigma_{\alpha\beta} = \omega_{\alpha\beta} = 0 \) imply that the spacetime is FRW is well-known [8]. Further, for an ideal fluid the conditions \( \dot{u}_a = 0, \sigma_{\alpha\beta} = 0 \) imply that \( \theta \omega_{\alpha\beta} = 0 \) [12], so the spacetime is either stationary or FRW (or both).

Note that we have made no assumptions about the photon distribution function outside the observed frequency range or about the distribution of matter components other than the CMB photons, in contrast to [3,4,5].

*Almost isotropy’ theorem. We generalise the previous theorem to the case of small anisotropy:

**Theorem 2** If observers at all \( x \) measure photon energy \( E(x, e, f) \) such that the first three harmonic moments of \( (\partial_t + e^\alpha \partial_\alpha) \ln E \) and their derivatives are independent of the direction \( e \) to \( \mathcal{O}(\varepsilon) \) (where \( \varepsilon \ll 1 \) is a constant), and the matter is an ideal fluid to \( \mathcal{O}(\varepsilon) \), the spacetime is, to \( \mathcal{O}(\varepsilon) \), either stationary or FRW.

Note that the observers are not assumed to be geodesic. The detailed conditions on \( (\partial_t + e^\alpha \partial_\alpha) \ln E \) and the matter content are given in the proof below.

When \( E \) is not perfectly isotropic, knowing the anisotropy of \( E \) is not sufficient to place constraints on the local geometry, because the relation (14) involves the derivatives of \( E \). In fact, we only need assumptions about the derivatives of \( \ln E \). We expand their direction dependence in covariant harmonics:
\[
(\partial_t + e^\alpha \partial_\alpha) \ln E(x, e, f) \equiv A = \bar{A}(x) + A_\alpha(x) e^\alpha + A_{\alpha\beta}(x) e^\alpha e^\beta + \sum_{n=3}^{\infty} A_{\alpha_1...\alpha_n}(x) e^{\alpha_1}...e^{\alpha_n}. \tag{20}
\]
Since (14) involves at most two \( e^\alpha \) on the right-hand side, we do not need moments of \( A \) higher than two, corresponding to the dipole and the quadrupole (of the derivatives, not the energy itself). Decomposed into harmonic moments, the relation (14) reads
\[
\bar{A} = -\frac{1}{3} \theta, \quad A_\alpha = -\dot{u}_\alpha, \quad A_{\alpha\beta} = -\sigma_{\alpha\beta}. \tag{21}
\]
If we assume that the anisotropy of \( A \) is small, we have to say with respect to which scale this holds, since \( A \) is a dimensional quantity. We introduce a timescale \( L \), and assume that \( A_\alpha, A_{\alpha\beta} \lesssim \mathcal{O}(\varepsilon)L^{-1} \) where \( \varepsilon \ll 1 \) is a fixed constant. (The scale \( L \) need not be constant; in fact, when \( A \) is not also small, a natural choice of \( L \) is the local Hubble time \( 3H^{-1} = \dot{A}^{-1} \).) From (21) we have
\[
\dot{u}_\alpha \lesssim \mathcal{O}(\varepsilon)L^{-1}, \quad \sigma_{\alpha\beta} \lesssim \mathcal{O}(\varepsilon)L^{-1}. \tag{22}
\]
We consider two possibilities, either \( \bar{A} \lesssim \mathcal{O}(\varepsilon)L^{-1} \) or \( \bar{A} \) is of the same order as \( L^{-1} \).

If \( \bar{A} \lesssim \mathcal{O}(\varepsilon)L^{-1} \), we have from (21) the result \( \theta \lesssim \mathcal{O}(\varepsilon)L^{-1} \), which combined with (22) means \( \nabla_\alpha \omega_{\beta\gamma} \lesssim \mathcal{O}(\varepsilon)L^{-1} \). So \( u^\alpha \) is an ‘almost Killing vector’, and the spacetime is ‘almost stationary’ when viewed on timescales of \( L \) or shorter. If the vorticity is also \( \mathcal{O}(\varepsilon)L^{-1} \), the spacetime is ‘almost static’.

If \( \bar{A} = 0(1)L^{-1} \), we have \( \theta = 0(1)L^{-1} \). Analogously to the exactly isotropic case, we write \( u_\alpha = -\dot{A}^{-1} \nabla_\alpha \bar{A} + \dot{A}^{-1} \nabla_\alpha \bar{A} \). Assuming that \( \nabla_\alpha \nabla_\beta \bar{A} \lesssim \mathcal{O}(\varepsilon)L^{-3} \) and \( \bar{A} = 0(1)L^{-2} \), it follows that \( \omega_{\alpha\beta} \lesssim \mathcal{O}(\varepsilon)L^{-1} \). The conditions \( \dot{u}_\alpha, \sigma_{\alpha\beta}, \omega_{\alpha\beta} \lesssim \mathcal{O}(\varepsilon)L^{-1} \) do not necessarily imply that the spacetime is FRW to \( \mathcal{O}(\varepsilon) \) even if the matter is an ideal fluid, because the evolution equations (10–17) involve their derivatives, which may be large. We assume that \( \nabla_\alpha \bar{A}, \nabla_\alpha \nabla_\beta \bar{A}, \nabla_\alpha A_{\beta\gamma}, \nabla_\alpha \dot{A}_{\beta\gamma}, \nabla_\alpha A_{\beta\gamma} \) are all \( \lesssim \mathcal{O}(\varepsilon) \). Then (3–17) reduce at leading order to the FRW equations, and spatial derivatives and Weyl tensor components are \( \mathcal{O}(\varepsilon) \). In this sense, the spacetime is FRW to \( \mathcal{O}(\varepsilon) \). The assumption about \( \nabla_\alpha \bar{A} \) could be replaced with further assumptions about the derivatives of \( \bar{A}, A_\alpha \) and \( A_{\alpha\beta} \).
IV. DISCUSSION

The real universe. Theorem 1 follows from the directional distribution of photon energies, which is observable. However, since the observed CMB is not perfectly isotropic, the theorem is not relevant for the real universe. In contrast, theorem 2 relies on assumptions about the derivatives of the energy, which are not directly observed. It would be more accurate to say that the ‘almost isotropy’ theorem characterises the relation between the geometry and the CMB in nearly stationary or nearly FRW universes, rather than that it places constraints on the geometry from CMB observations. The same is true for the theorem of [4], where it was assumed that derivatives of the temperature anisotropy of up to third order are small. The smallness of the derivatives may seem to be just a technical assumption, and in [5] it was justified with the Copernican principle. However, the derivatives are related to the local geometry, and large local variations are not in contradiction with the assumption that our position is typical. Extending the Copernican principle to mean that local variations are small is equivalent to assuming that the universe is almost FRW, and this can be done for the geometry without going via the CMB.

In the real universe, there are large spatial variations in the geometry. For example, the volume expansion rate $\theta$ varies by a factor of order unity between underdense and overdense regions. Reading the relation (4) between the energy and the geometry the other way than we have done so far shows that the derivatives of $E$ are therefore large. This does not imply that the energy would have large spatial variations and thus be anisotropic (otherwise there would be no need to make assumptions about the derivatives of $E$ for the ‘almost isotropy’ theorem). This can be made transparent by integrating (4) to obtain $E(x, e, f) = E(x) \exp (-\int_s^{x_0} d\eta [\frac{1}{2} \theta + \dot{u}_\alpha e^\alpha + \sigma_{\alpha\beta} e^\alpha e^\beta])$, where the integral is from a source to the observer along the null geodesic which has the tangent vector $e^\alpha$. Here $\eta$ is a parameter along the geodesic defined by $d\eta \equiv \partial_t + e^\alpha \partial_\alpha$. The CMB photons start from the last scattering surface, and $E_s$ is given by the distribution $f$ and the constant decoupling temperature. This relation shows how the energy itself, instead of its derivatives, is related to the geometry. As long as the spatial variation in the geometry is uncorrelated over long distances, and the coherence scale is small compared to the magnitude of the derivatives, the anisotropy in $E$ is small [7].

Conclusion. A perfectly isotropic CMB seen by geodesic observers implies that the spacetime is stationary or FRW (or has the restricted form (18), (19)), but an almost isotropic CMB implies that the universe is almost stationary or almost FRW only with additional assumptions about the derivatives of the CMB temperature. These assumptions do not hold in the real universe. The observed isotropy of the CMB, coupled with the Copernican assumption and analyticity, indicates statistical homogeneity and isotropy, but not local homogeneity and isotropy. This is an important distinction, because a universe which is only statistically homogeneous and isotropic is not described by the FRW metric, and its expansion is not determined by the FRW equations. In particular, the assumption that the universe is FRW is crucial in deducing the existence of dark energy or modified gravity. Dropping this assumption, it may be possible to explain the observed deviation in the expansion of the universe at late times from the matter-dominated FRW prediction without new physics [7, 13, 14].

Acknowledgments. I thank Ruth Durrer for helpful discussions and comments on the manuscript, and Roy Maartens for comments and for pointing out an error in the first version of the paper.