Is the Universe homogeneous?

Roy Maartens

Department of Physics, University of Western Cape, Cape Town 7535, South Africa
Institute of Cosmology & Gravitation, University of Portsmouth, Portsmouth PO1 3FX, UK

The standard model of cosmology is based on the existence of homogeneous surfaces as the background arena for structure formation. Homogeneity underpins both general relativistic and modified gravity models and is central to the way in which we interpret observations of the CMB and the galaxy distribution. However, homogeneity cannot be directly observed in the galaxy distribution or CMB, even with perfect observations, since we observe on the past lightcone and not on spatial surfaces. We can directly observe and test for isotropy, but to link this to homogeneity, we need to assume the Copernican Principle. First, we discuss the link between isotropic observations on the past lightcone and isotropic spacetime geometry: what observations do we need to be isotropic in order to deduce spacetime isotropy? Second, we discuss what we can say with the Copernican assumption. The most powerful result is based on the CMB: the vanishing of the dipole, quadrupole and octupole of the CMB is sufficient to impose homogeneity. Real observations lead to near-isotropy on large scales – does this lead to near-homogeneity? There are important partial results, and we discuss why this remains a difficult open question. Thus we are currently unable to prove homogeneity of the Universe on large-scales, even with the Copernican Principle. However we can use observations of galaxies and clusters to test the Copernican Principle itself.

I. INTRODUCTION

The standard model of the Universe – the LCDM "concordance" model – is homogeneous, with structure formation described via perturbations. Given the assumption of homogeneity, and if GR correctly describes gravity, then the acceleration of the Universe is driven by dark energy. The homogeneous LCDM model is highly successful – a simple, predictive model that is compatible with all observations up to now. However, there is still no satisfactory description of the dark energy that is central to this model. This motivates the need to probe the foundations of the model. We can probe the assumption that GR holds on cosmological scales, by investigating modified gravity theories and by devising consistency tests of GR. This probe is only effective if we assume homogeneity. Alternatively, we can assume that GR holds and probe the assumption of homogeneity.[1]

A common misconception is that "homogeneity is obvious from the CMB and the galaxy distribution". In fact, *we cannot directly observe or test homogeneity* – since we observe down the past lightcone, and not on spatial surfaces that intersect that lightcone (see Fig. 1). We only see the CMB on a 2-sphere at one redshift, and galaxy surveys give us the galaxy distribution on 2-spheres of constant redshift. There are interesting and important analyses of the observed galaxy distribution to probe statistical homogeneity (e.g. [3, 4]), but these effectively assume an FLRW background geometry.

What we can directly test for is *isotropy of observations*. This then raises an important, but often overlooked question: *what observational quantities need to be isotropic in order to enforce isotropy of the geometry?* This question is addressed in Sec. II. To answer this question we need a fully nonlinear analysis, since we cannot assume a priori any symmetric background spacetime. For observations of the galaxy distribution, the answer is – we need isotropic angular distances, number counts, bulk velocities and lensing. Isotropy of the CMB by contrast does *not* in itself enforce spacetime isotropy.

In order to link isotropy to homogeneity, we have to assume the Copernican Principle (CP), i.e. that we are not at a special position in the Universe. The CP is not observationally based; it is an expression of the intrinsic limitation of observations from one spacetime location. We consider in Sec. III what can be done with the CP. If we have isotropy along one worldline, based on the observed galaxy distribution, then the CP leads to homogeneity. A more powerful result is that *homogeneity follows if all observers see isotropic angular distances up to third order in redshift.* The strongest basis for homogeneity comes from the CMB. This is often considered to be obvious – but it is far from straightforward to show that homogeneity of the metric follows if all observers see isotropic CMB. The proof requires the general nonlinear Einstein-Liouville equations. Remarkably, it is not necessary to assume full CMB isotropy for

---

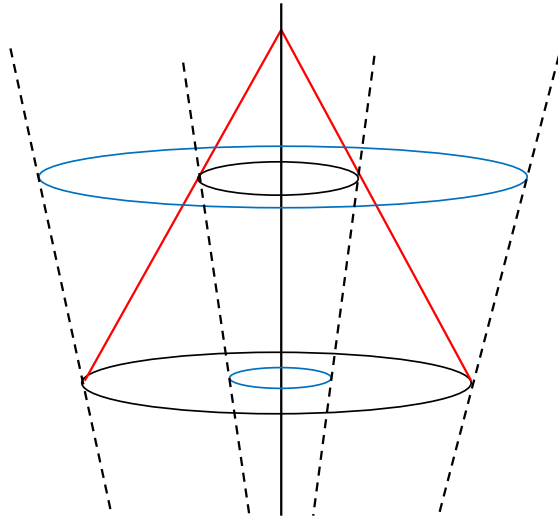[1] This paper builds on [1]. See also [2] for some related discussions.

FIG. 1: We observe down the past lightcone and therefore cannot directly confirm homogeneity.

each observer – it is enough that each observer *sees isotropy in the CMB only up to the octupole.* This is the most powerful observational basis for homogeneity currently known.

Of course the real CMB is not exactly isotropic, but nearly isotropic. Does it follow from near-isotropy that the Universe is nearly homogeneous, i.e. perturbed FLRW? This has so far only been shown with further assumptions on the gradients and time derivatives of CMB multipoles.

It is important to stress from the outset that there are two fundamental limitations:

- Isotropy and homogeneity of observables can only be meaningfully defined on large enough scales – and the nature of the transition scale is only poorly understood.
- Isotropy and homogeneity of observables can only be meaningfully defined in an average sense – and the problem of how to average in general relativity (and other metric theories that are intrinsically nonlinear) remains unsolved.

These unresolved issues are of crucial importance in cosmology, but they are not discussed here; instead, observations are treated as idealized.

II.   WHAT IS THE BASIS FOR ISOTROPY?

Here we consider the situation when the CP is *not* adopted. What is the observational basis for showing that spacetime is isotropic about the worldline of a single observer? Isotropy is directly observable and the best example is the CMB, which is isotropic about us to $\sim 10^{-5}$ (after the dipole is interpreted as due to our motion relative to the cosmic frame, and removed by a boost). Observations of the galaxy distribution do not have the same precision, but there is no evidence for anisotropy. First we look at observations of matter and then of the CMB.

A.   Isotropic matter distribution on the past lightcone of one observer

The dominant cosmological components – cold dark matter and dark energy – have not been independently observed. Unlike baryonic matter, the dark components are up to now only manifest via their gravitational effect. The distribution of dark matter is mapped by weak lensing surveys. But to relate the measured projected potential on the sky at each redshift to the dark matter, we require a specific model, such as a perturbed FLRW model. The dark matter 4-velocity is usually assumed to be aligned with that of baryonic matter – but this is also based on a perturbed FLRW model.

This unavoidably means that we must impose a model for these dark components – not merely their physical properties, but how they relate spatially to observed matter – in order to determine their distribution via cosmological

observations. A starting point [1] is to assume that the CDM 4-velocity is the same as the baryonic 4-velocity, and that we know the primordial ratio of CDM density to baryonic density, as well as the bias factor that relates the concentrations of CDM and baryons in clustered matter

$$\rho_{\rm c} \text{ known from } \rho_{\rm b} \text{ and } u_{\rm b}^a = u_{\rm c}^a := u^a. \tag{1}$$

If there is a modified gravity theory that avoids the need for CDM, then we do not need (1) – but we are likely to need other assumptions on the extra degrees of freedom that mimic CDM.

If dark energy is in the form of Λ, then we need to assume that its value is known from non-cosmological physics:

$$\Lambda \text{ known independently of cosmological observations.} \tag{2}$$

For quintessence and other more complicated forms of dark energy, we would need to assume how the dark energy field is distributed in spacetime – since we are not assuming a FLRW geometry a priori. If there is a modified gravity theory that avoids the need for dark energy, then (2) is not needed, but assumptions will likely be necessary on the extra degrees of freedom that mimic dark energy.

Given the assumptions on the dark components, what can we say about spacetime geometry if the matter distribution is isotropic on the past lightcone of the observer? Which observables need to be isotropic in order to deduce isotropic geometry? It turns out that 4 independent observables on the lightcone are exactly enough to impose isotropy of spacetime. The original result for a baryonic universe [5–7] may be updated to include CDM and Λ [1], incorporated via (1) and (2):

Isotropy of matter distribution on the lightcone → isotropy of spacetime geometry

If one fundamental observer comoving with the matter measures isotropy of (a) angular diameter distances, (b) number counts, (c) bulk velocities, and (d) lensing, in an expanding dust universe with Λ, then the spacetime is isotropic about the observer's worldline.

Note that isotropy of bulk velocities is equivalent to vanishing transverse velocities (proper motions) on the observer's sky. Isotropy of lensing means that there is no distortion of images, only magnification.

The proof of this result requires a non-perturbative approach – there is no background to perturb around. Since the data is given on the past lightcone of the observer, we need the full general metric, adapted to the past lightcones of the observer worldline $\mathcal{C}$. We define observational coordinates $x^\mu = (w, y, \theta, \phi)$, where $x^P = (\theta, \phi)$ are the celestial coordinates, $w = $ const are the past light cones on $\mathcal{C}$ ($y = 0$), normalized so that w measures proper time along $\mathcal{C}$, and y measures distance down the light rays $(w, \theta, \phi) = $ const (see Fig. 2). A convenient choice for y is $y = z$ (redshift) on the lightcone of here-and-now, $w = w_0$, and then keep y comoving with matter off the initial lightcone, so that $u^y = 0$. Then the matter 4-velocity and the photon wave-vector are

$$u^\mu = (1+z)(1, 0, V^P), \quad k_\mu = w_{,\mu}, \quad 1+z = u_\mu k^\mu, \tag{3}$$

where $V^P = \mathrm{d}x^P/\mathrm{d}w$ are the transverse velocity components on the observer's sky.

The metric of an arbitrary spacetime in observational coordinates is

$$\mathrm{d}s^2 = -A^2\mathrm{d}w^2 + 2B\mathrm{d}w\mathrm{d}y + 2C_P\mathrm{d}x^P\mathrm{d}w + D^2(\mathrm{d}\Omega^2 + L_{PQ}\mathrm{d}x^P\mathrm{d}x^Q) \tag{4}$$

$$A^2 = (1+z)^{-2} + 2C_P V^P + g_{PQ}V^P V^Q, \quad B = \frac{\mathrm{d}v}{\mathrm{d}y}, \tag{5}$$

where v is a null affine parameter, D is the angular diameter distance, and L_{PQ} determines the lensing distortion of images via the shear of lightrays,

$$\hat{\sigma}_{PQ} = \frac{D^2}{2B}\frac{\partial L_{PQ}}{\partial y}. \tag{6}$$

The number of galaxies in a solid angle $\mathrm{d}\Omega$ and a null distance increment $\mathrm{d}v = B\mathrm{d}y$ is

$$\mathrm{d}N = Sn(1+z)D^2 B\mathrm{d}\Omega\mathrm{d}y, \tag{7}$$

where S is the selection function and $n = \rho_{\rm m}/m$ is the number density.

Before specializing to isotropic observations, we identify how the observations in general and in principle determine the geometry of the past light cone $w = w_0$ of here-and-now, where $y = z$:
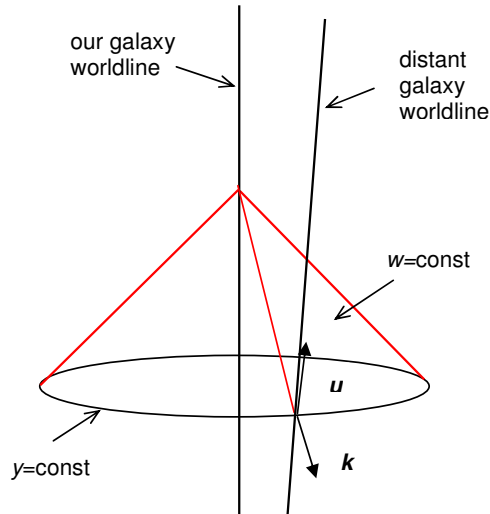
FIG. 2: Observational coordinates based on the past lightcones of the observer's worldline.

- Given the intrinsic properties and evolution of sources, observations in principle determine: (a) the angular diameter distance $D(w_0, z, x^P)$, (b) the lensing distortion of images, $L_{PQ}(w_0, z, x^R)$. Thus the metric components g_{PQ} on $w = w_0$ are determined.
- Given (1) and the selection function, the number counts $N(w_0, z, x^P)$ in principle determine $B(w_0, z, x^P)\rho_{\mathrm{m}}(w_0, z, x^P)$.
- In principle, observations over extended timescales determine the instantaneous transverse velocities $V^P(w_0, z, x^Q)$ of discrete sources. Transverse velocities of clusters are in principle determined by the polarization of scattered CMB photons [8].

It follows that, in principle and for idealized observations:

$$\text{Idealized data} \ \Rightarrow \ \{u^\mu, B\rho_{\mathrm{m}}, g_{PQ}\} \text{ on } w = w_0. \tag{8}$$

But this is insufficient to determine the geometry of the past lightcone, because we need $g_{0P} = C_P$ and we cannot separate out B and ρ_{m}. Without gravitational field equations, we are unable to fully determine the spacetime and matter on the past lightcone, even assuming perfect information from discrete-source observations. As a consequence, it is also impossible to test gravity theories directly:

Cosmological testing of gravity theories

Even with perfect observations, we cannot determine the spacetime geometry and matter on our past lightcone without gravitational field equations. Thus observations cannot directly test GR on cosmological scales, or test any alternative theories of large-scale gravity, without making assumptions about the spacetime geometry.

If GR holds, then (8) is exactly what is needed [5, 6]:

Matter observations → metric and matter on lightcone

Observational data (8) is exactly the information needed for Einstein's equations to determine B and C_P on $w = w_0$, so that the metric and matter are fully determined on the lightcone. Then the past Cauchy development of this data determines $g_{\mu\nu}, u^\mu, \rho_{\mathrm{m}}$ in the interior of the past lightcone.

If the matter observations are isotropic, then we can prove isotropy of spacetime and matter [5, 6]:

Isotropy of lightcone matter distribution for one observer → LTB

In an expanding dust region with Λ, if one fundamental observer measures isotropic angular diameter distances, number counts, bulk velocities, and lensing,

$$\frac{\partial D}{\partial x^P} = \frac{\partial N}{\partial x^P} = V^P = L_{PQ} = 0\,, \tag{9}$$

then the spacetime and matter are isotropic, i.e. the region is Lemaitre-Tolman-Bondi (LTB).

It is an open question how this result translates to the case of almost-isotropy of observations, i.e. *does this lead to almost-isotropy of the spacetime?*

B. Isotropic CMB for one observer

If the CMB is isotropic for one observer comoving with the matter, then along the observer's worldline $\mathcal{C}$ we have [see (17)]

$$f(x_\mathcal{C}, p) = F(x_\mathcal{C}, E), \quad F_{\mu_1 \cdots \mu_\ell}\big|_\mathcal{C} = 0 = \dot{F}_{\mu_1 \cdots \mu_\ell}\big|_\mathcal{C} \text{ for } \ell \geq 1. \tag{10}$$

In other words, all covariant multipoles of the distribution function beyond the monopole [see (A32)], and their time derivatives, must vanish along $\mathcal{C}$. By (A33), the radiation momentum density (from the dipole) and anisotropic stress (from the quadrupole) vanish: $q_r^\mu|_\mathcal{C} = 0 = \pi_r^{\mu\nu}|_\mathcal{C}$. However, without the Copernican assumption, we are not able to deduce directly the vanishing of spatial derivatives $\tilde{\nabla}_\nu F_{\mu_1 \cdots \mu_\ell}\big|_\mathcal{C}$, and then we cannot show isotropy of the spacetime geometry about $\mathcal{C}$.

Isotropy of the CMB alone is not sufficient, since the matter could in principle be anisotropic, even if this is physically unnatural. In order to rule out artificial counter-examples, it would be sufficient to *characterize the minimal matter isotropy that combines with CMB isotropy to give geometric isotropy.*

If we adopt the Copernican Principle, then the CMB alone leads to a powerful result, as discussed in the next section.

III. WHAT IS THE BASIS FOR SPATIAL HOMOGENEITY?

Homogeneity cannot be directly observed – we are effectively unable to move away in cosmic time or distance from here-and-now and hence cannot probe spatial variations on constant time slices; effectively, our observations only access the past lightcone of here-and-now. Direct observation thus cannot distinguish between an evolving homogeneous distribution of matter and inhomogeneity with a different time evolution – since the past light cone only accesses a 2-sphere in each constant-time slice (see Fig. 1). Thus we are forced to adopt the Copernican assumption. We first consider matter observations, then an exactly isotropic CMB, and finally the case of an almost-isotropic CMB.

A. Isotropic matter observations for all observers

If all observers see isotropy of the 4 matter observables, then we have geometric isotropy along all worldlines and thus homogeneity follows.

> **Isotropy of lightcone matter distribution for all observers → FLRW**
>
> In an expanding dust region with Λ, if all fundamental observers measure isotropic angular diameter distances, number counts, bulk velocities, and lensing, then the spacetime is FLRW.

This is an improved form of the Cosmological Principle – based on isotropy of specific observables and not on assumed geometric isotropy.

In fact there is a much stronger statement than this, based only on one observable, the angular diameter distance, and only for small redshifts [9]:

> **Isotropic distances up to $O(z^3)$ for all observers → FLRW**
>
> If all fundamental observers in an expanding spacetime region measure isotropic angular diameter distances up to third-order in a redshift series expansion, then the spacetime is FLRW in that region.

A covariant proof of this [1, 10] is based on a series expansion in a general spacetime, using the method of Kristian and Sachs [11]. The redshift may be expanded in terms of the angular diameter distance:

$$\begin{aligned} z = & \left[k^\mu k^\nu \nabla_\mu u_\nu\right]_O D + \frac{1}{2}\left[k^\mu k^\nu k^\alpha \nabla_\mu \nabla_\nu u_\alpha\right]_O D^2 \\ & + \frac{1}{6}\left[k^\mu k^\nu k^\alpha k^\beta \nabla_\mu \nabla_\nu \nabla_\alpha u_\beta + \frac{1}{2} k^\mu k^\nu k^\alpha k^\beta R_{\alpha\beta} \nabla_\mu u_\nu\right]_O D^3 + O(D^4). \end{aligned} \tag{11}$$

Here all terms are evaluated at the observer O in the unit direction e^μ of the lightray, where [see (3)]:

$$k^\mu = -(1+z)(u^\mu + e^\mu), \quad u_\mu e^\mu = 0, \; e_\mu e^\mu = 1. \tag{12}$$

The covariant derivative of the 4-velocity is decomposed as in (A5). The $O(D)$ coefficient is the observed Hubble rate (generalizing the FLRW quantity),

$$H_0^{\mathrm{obs}} = \left[k^\mu k^\nu \nabla_\mu u_\nu\right]_0 = \left[\frac{1}{3}\Theta + A_\mu e^\mu + \sigma_{\mu\nu} e^\mu e^\nu\right]_O. \tag{13}$$

Isotropy at lowest order, for all observers O, thus enforces vanishing acceleration and shear

$$A_\mu = 0 = \sigma_{\mu\nu}. \tag{14}$$

Using (14), the next coefficient becomes

$$\left[k^\mu k^\nu k^\alpha \nabla_\mu \nabla_\nu u_\alpha\right]_O = \frac{1}{6}\left[\left(2\Theta^2 + \rho_{\mathrm{m}} - 2\Lambda\right) - 2\tilde{\nabla}_\mu \Theta e^\mu + \left(E_{\mu\nu} + \omega_{\langle\mu}\omega_{\nu\rangle}\right)e^\mu e^\nu\right]_O, \tag{15}$$

where $\tilde{\nabla}_\mu$ is the covariant spatial derivative (A3), $E_{\mu\nu}$ is the electric part of the Weyl tensor (A8), and the angled brackets denote the spatial tracefree part (A2). Isotropy at $O(D^2)$ therefore imposes homogeneity of the expansion rate, $\tilde{\nabla}_\mu \Theta = 0$, and the condition $E_{\mu\nu} = -\omega_{\langle\mu}\omega_{\nu\rangle}$. However, this condition is identically satisfied by virtue of (14) and the shear propagation equation (A14). The constraint equations (A18), (A19) then show that $\mathrm{curl}\,\omega_\mu = 0$ and $H_{\mu\nu} = \tilde{\nabla}_{\langle\mu}\omega_{\nu\rangle}$, where $H_{\mu\nu}$ is the magnetic Weyl tensor. The $O(D^3)$ coefficient imposes further constraints, and we find that $0 = \tilde{\nabla}_\mu \rho_{\mathrm{m}} = \omega_\mu = E_{\mu\nu} = H_{\mu\nu}$. Putting everything together, we have a covariant characterization of FLRW.

B. Exactly isotropic CMB for all observers

It is commonly assumed that isotropy of the CMB for all observers leads obviously to FLRW, without the need for any proof. In fact, the proof is far from obvious, and requires a nonlinear analysis of the general Einstein-Liouville equations. The starting point is a pioneering mathematical result by Ehlers, Geren and Sachs [12]. They assumed that the only source of the gravitational field was the radiation, i.e. they neglected matter and Λ. This can be generalized to include self-gravitating matter and dark energy [1] (extending [13–18]):

CMB isotropy for all observers → FLRW
In a region, if
- collisionless radiation is exactly isotropic,
- the radiation four-velocity is geodesic and expanding,
- there is dust matter and dark energy in the form of Λ, quintessence or a perfect fluid,

then the metric is FLRW in that region.

The fundamental 4-velocity u^μ is the radiation 4-velocity, which has zero 4-acceleration and positive expansion:

$$A_\mu = 0, \quad \Theta > 0. \tag{16}$$

Isotropy of the radiation distribution about u^μ means that photon peculiar velocities are isotropic for comoving observers; thus in momentum space, the photon distribution depends only on components of the 4-momentum p^μ along u^μ, i.e., on the photon energy $E = -u_\mu p^\mu$:

$$f(x,p) = F(x,E), \quad F_{\mu_1\cdots\mu_\ell} = 0 \text{ for } \ell \geq 1. \tag{17}$$

In other words, all covariant multipoles of the distribution function beyond the monopole, defined in (A32), must vanish. In particular, (A33) shows that:

$$q_{\mathrm{r}}^\mu = 0 = \pi_{\mathrm{r}}^{\mu\nu}. \tag{18}$$

Equation (17) also implies that the radiation brightness octupole $\Pi_{\mu\nu\alpha}$ and hexadecapole $\Pi_{\mu\nu\alpha\beta}$ are zero. These are source terms in the anisotropic stress evolution equation, which is the $\ell = 2$ case of (A37). The general nonlinear form of the $\pi_{\mathrm{r}}^{\mu\nu}$ evolution equation is [13, 20]

$$\dot{\pi}_{\mathrm{r}}^{\langle\mu\nu\rangle} + \frac{4}{3}\Theta\pi_{\mathrm{r}}^{\mu\nu} + \frac{8}{15}\rho_{\mathrm{r}}\sigma^{\mu\nu} + \frac{2}{5}\tilde{\nabla}^{\langle\mu}q_{\mathrm{r}}^{\nu\rangle} + 2A^{\langle\mu}q_{\mathrm{r}}^{\nu\rangle} - 2\omega^\alpha \varepsilon_{\alpha\beta}{}^{(\mu}\pi_{\mathrm{r}}^{\nu)\beta}$$
$$+ \frac{2}{7}\sigma_\alpha{}^{\langle\mu}\pi_{\mathrm{r}}^{\nu\rangle\alpha} + \frac{8\pi}{35}\tilde{\nabla}_\alpha\Pi^{\mu\nu\alpha} - \frac{32\pi}{315}\sigma_{\alpha\beta}\Pi^{\mu\nu\alpha\beta} = 0. \tag{19}$$

Isotropy removes all terms on the left except the third, and thus enforces a shear-free expansion of the fundamental congruence:

$$\sigma_{\mu\nu} = 0\,. \tag{20}$$

We can also show that u^{μ} is irrotational as follows. Together with (16), momentum conservation for radiation, i.e., (A30) with $I = r$, reduces to

$$\tilde{\nabla}_{\mu}\rho_{\rm r} = 0\,. \tag{21}$$

Thus the radiation density is homogeneous relative to fundamental observers. Now we invoke the exact nonlinear identity for the covariant curl of the gradient, (A6):

$$\operatorname{curl}\tilde{\nabla}_{\mu}\rho_{\rm r} = -2\dot{\rho}_{\rm r}\omega_{\mu} \;\Rightarrow\; \Theta\rho_{\rm r}\omega_{\mu} = 0\,, \tag{22}$$

where we have used the energy conservation equation (A29) for radiation. By assumption $\Theta > 0$, and hence we deduce that the vorticity must vanish:

$$\omega_{\mu} = 0\,. \tag{23}$$

Then we see from the curl shear constraint equation (A19) that the magnetic Weyl tensor must vanish:

$$H_{\mu\nu} = 0\,. \tag{24}$$

Furthermore, (21) actually tells us that the expansion must also be homogeneous. From the radiation energy conservation equation (A29), and using (18), we have $\Theta = -3\dot{\rho}_{\rm r}/4\rho_{\rm r}$. On taking a covariant spatial gradient and using the commutation relation (A7), we find

$$\tilde{\nabla}_{\mu}\Theta = 0\,. \tag{25}$$

Then the shear divergence constraint, (A18), enforces the vanishing of the total momentum density in the fundamental frame,

$$q^{\mu} := \sum_{I} q_{I}^{\mu} = 0 \;\Rightarrow\; \sum_{I} \gamma_{I}^{2}(\rho_{I}^{*} + p_{I}^{*})v_{I}^{\mu} = 0\,. \tag{26}$$

The second equality follows from (A26), using the fact that the baryons, CDM and dark energy (in the form of quintessence or a perfect fluid) have vanishing momentum density and anisotropic stress in their own frames, i.e.,

$$q_{I}^{*\mu} = 0 = \pi_{I}^{*\mu\nu}\,, \tag{27}$$

where the asterisk denotes the intrinsic quantity (see Appendix A). If we include other species, such as neutrinos, then the same assumption applies to them. Except in artificial situations, it follows from (26) that

$$v_{I}^{\mu} = 0\,, \tag{28}$$

i.e., the bulk peculiar velocities of matter and dark energy [and any other self-gravitating species satisfying (27)] are forced to vanish – all species must be comoving with the radiation.

The comoving condition (28) then imposes the vanishing of the total anisotropic stress in the fundamental frame:

$$\pi^{\mu\nu} := \sum_{I} \pi_{I}^{\mu\nu} = \sum_{I} \gamma_{I}^{2}(\rho_{I}^{*} + p_{I}^{*})v_{I}^{\langle\mu}v_{I}^{\nu\rangle} = 0\,, \tag{29}$$

where we used (A27), (27) and (28). Then the shear evolution equation (A14) leads to a vanishing electric Weyl tensor

$$E_{\mu\nu} = 0\,. \tag{30}$$

Equations (26) and (29), now lead via the total momentum conservation equation (A12) and the E-divergence constraint (A20), to homogeneous total density and pressure:

$$\tilde{\nabla}_{\mu}\rho = 0 = \tilde{\nabla}_{\mu}p\,. \tag{31}$$

Equations (16), (20), (24), (25), (26), (29) and (31) constitute a covariant characterization of an FLRW spacetime. This establishes the EGS result, generalized from the original to include self-gravitating matter and dark energy, and presented in a fully covariant form. It is straightforward to include other species such as neutrinos. The critical assumption needed for all species is the vanishing of the intrinsic momentum density and anisotropic stress, i.e., (27). Equivalently, the energy-momentum tensor for the I-component should have perfect fluid form in the I-frame. The isotropy of the radiation and the geodesic nature of its 4-velocity then enforce the vanishing of (bulk) peculiar velocities v_I^μ. We emphasize that one does *not* need to assume that the matter or other species are comoving with the radiation – it follows from the assumptions on the radiation.

In fact this result can be dramatically strengthened on the basis of a theorem by Ellis, Treciokas and Matravers [19]: we do not need vanishing of all multipoles, but only the dipole, quadrupole and octupole! The key step is to show that the shear vanishes, without having zero hexadecapole – the quadrupole evolution equation (19) no longer automatically gives $\sigma_{\mu\nu} = 0$, and we need to find another way to show this. The ETM trick is to return to the Liouville multipole equation (A36). The $\ell = 2$ multipole of this equation, with $F_\mu = F_{\mu\nu} = F_{\mu\nu\alpha} = 0$, gives

$$\frac{12}{63}\frac{\partial}{\partial E}\left(E^5\sigma^{\mu\nu}F_{\mu\nu\alpha\beta}\right) + E^5\frac{\partial F}{\partial E}\,\sigma_{\alpha\beta} = 0\,. \tag{32}$$

We integrate over E from 0 to ∞, and use the convergence property $E^5F_{\mu\nu\alpha\beta} \to 0$ as $E \to \infty$. This gives

$$\sigma_{\alpha\beta}\int_0^\infty E^5\frac{\partial F}{\partial E}\mathrm{d}E = 0\,. \tag{33}$$

Integrating by parts, the integral reduces to $-5\int_0^\infty E^4F\mathrm{d}E$. Since $F > 0$, the integral is strictly negative, and thus we arrive at vanishing shear, $\sigma_{\mu\nu} = 0$. Then the proof above proceeds as before, incorporating matter and dark energy to extend the ETM result [1]:

> ### CMB partial isotropy for all observers → FLRW
>
> In a region, if
>
> - collisionless radiation has vanishing dipole, quadrupole and octupole,
>
> $$F_\mu = F_{\mu\nu} = F_{\mu\nu\alpha} = 0\,, \tag{34}$$
>
> - the radiation four-velocity is geodesic and expanding,
> - there is dust matter and dark energy in the form of Λ, quintessence or a perfect fluid,
>
> then the metric is FLRW in that region.

This is the most powerful observational basis that we have for background homogeneity and thus an FLRW background model.

C. The real universe: almost-isotropic CMB

In practice we can only observe approximate isotropy (see Fig. 3). Is the previous result stable – i.e., does almost-isotropy of the CMB lead to an almost-FLRW Universe? This would be the *realistic* basis for a spatially homogeneous Universe (assuming the Copernican Principle). It was shown to be the case, but subject to further assumptions, by [13, 21]:

> ### A more realistic basis for homogeneity
>
> In a region of an expanding Universe with dust and cosmological constant, if all observers comoving with the matter measure an almost isotropic distribution of collisionless radiation, and if some of the time and spatial derivatives of the covariant multipoles are also small, then the region is almost FLRW.

We emphasize that the perturbative assumptions are purely on the photon distribution, not the matter or the metric – and one has to prove that the matter and metric are then perturbatively close to FLRW. Once again, a nonperturbative analysis is essential, since we are trying to prove an almost-FLRW spacetime, and we cannot assume it a priori.
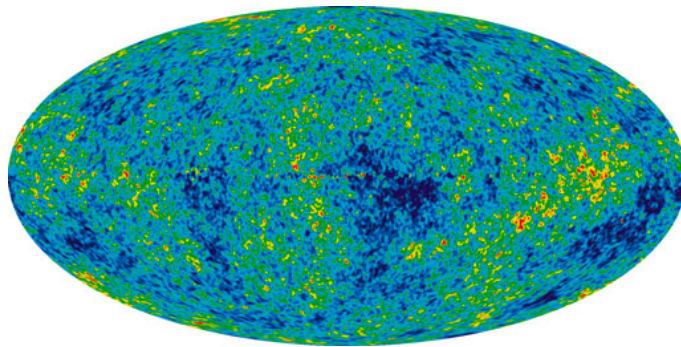
FIG. 3: The CMB temperature from WMAP 7-year data, showing anisotropies at the level $\sim 10^{-5}$. (Credit: NASA/ WMAP Science Team.)

Almost-isotropy of the photon distribution means that

$$F_{\mu_1 \cdots \mu_\ell}(x, E) = \mathcal{O}(\epsilon), \quad \ell \geq 1, \tag{35}$$

where ϵ is a (dimensionless) smallness parameter. The brightness multipoles Π_{M_ℓ} have dimensions of energy density and we therefore normalize them to the monopole $\Pi = \rho_{\rm r}/4\pi$, so that $\Pi_{M_\ell}/\Pi = \mathcal{O}(\epsilon)$.

The task is to show that the relevant kinematical, dynamical and curvature quantities, suitably non-dimensionalized, are $\mathcal{O}(\epsilon)$. For example, the dimensionful kinematical quantities may be normalized by the expansion, $\sigma_{\mu\nu}/\Theta, \omega_\mu/\Theta$. The proof then follows the same pattern as the proof above of the exact result – except that at each stage, we need to show that quantities are $\mathcal{O}(\epsilon)$ rather than equal zero.

However, in order to show this, we need smallness not just of the multipoles, but also of some of their derivatives. Smallness of the multipoles does not directly imply smallness of their derivatives, and we have to assume this [18, 22]. It remains a difficult open problem whether these additional assumptions may be removed. If all observers measure small multipoles, then it may be possible to use almost-isotropy of the matter distribution to show that the time and space derivatives on cosmologically significant scales must also be small.

A number of experiments has been proposed to test the Copernican Principle by looking for violations of isotropy at events down our past lightcone, as discussed in the next section. The almost-isotropic CMB result then gives a framework for probing inhomogeneities via such observations. Indeed, these tests may provide a way of constraining spatial gradients of the low-ℓ multipoles.

In addition, it may be possible to strengthen the almost-isotropic result above by proving that it is sufficient for only the first 3 multipoles and their derivatives to be small. This would represent a more realistic foundation for almost-homogeneity.

IV. TESTING THE COPERNICAN PRINCIPLE

Although we cannot directly probe homogeneity by observations, we can test for violations of the CP, which is at the foundation of homogeneity. If we find no violation, then the indirect evidence for homogeneity is strengthened. However, if even one significant violation of the CP is discovered, then homogeneity will have been disproved. There are two broad classes of observational tests of the CP, one based on consistency relations that hold in FLRW and the other based on using galaxy clusters as probes of the anisotropy seen in the CMB from positions at cosmological distances down our past lightcone.

A. Consistency of distances and expansion rate

The effective standard candles provided by supernovae observations lay the basis for a consistency test of homogeneity. There are two geometric effects on distance measurements: the curvature bends null geodesics and the expansion changes radial distances. These are coupled in FLRW models via

$$D_L(z) = \frac{(1+z)}{H_0\sqrt{-\Omega_{K0}}} \sin\left(\sqrt{-\Omega_{K0}} \int_0^z \frac{\mathrm{d}z'}{H(z')/H_0}\right), \tag{36}$$

and then one can combine the Hubble rate and distance data to find the curvature today:

$$\Omega_{K0} = \frac{H^2(z)[(1+z)D'_L(z) - D_L(z)]^2 - H_0^2(1+z)^4}{H_0^2(1+z)^2 D_L^2(z)}. \tag{37}$$

This relation is independent of all other cosmological parameters, including dark energy – and it is also independent of the theory of gravity. It can be used at a single redshift to determine Ω_{K0}. Furthermore, it is the basis for a CP test, proposed by [23] – since Ω_{K0} is independent of z, we can differentiate (37) to get the consistency relation:

$$\begin{aligned}
\mathcal{K}(z) \; := \; & H_0^2(1+z)^4 + H^2(z)\left[(1+z)^2\left\{D_L(z)D''_L(z) - D'^2_L(z)\right\} + D_L^2(z)\right] \\
& + (1+z)H(z)H'(z)D_L(z)\left[(1+z)D'_L(z) - D_L(z)\right] \\
= \; & 0 \;\; \text{for FLRW geometry.}
\end{aligned} \tag{38}$$

Note that $\mathcal{K} = 0$ for *any* FLRW geometry, independent of curvature, dark energy, matter content, and theory of gravity. In realistic models we should expect $|\mathcal{K}(z)| \sim 10^{-5}$, reflecting perturbations from large-scale structure formation. Significantly larger values indicate a breakdown of homogeneity:

$$\mathcal{K}(z) \;\; \text{significantly different from } 0 \;\; \Rightarrow \;\; \text{non-FLRW universe.} \tag{39}$$

Carrying out this test should not be more difficult than carrying out dark energy measurements of $w(z)$ from SNIa data, which require $H'(z)$ from distance measurements or the second derivative $D''_L(z)$.

This is the simplest test of homogeneity, and its implementation should be regarded as a high priority. Another CP test involves the time drift of the cosmological redshift [24], but this will only be feasible on a much longer timescale.
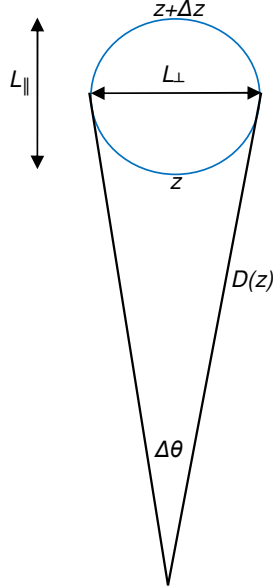

FIG. 4: Radial and transverse BAO scales.

Finally the BAO feature itself provides in principle a CP test. Future large-volume surveys will allow the detection of the BAO scale in both radial and transverse directions. The physical lengths in radial and transverse directions of a feature with redshift extent Δz and subtending an angle $\Delta\theta$ are

$$L_\parallel = \frac{\Delta z}{(1+z)H(z)}, \quad L_\perp = D(z)\Delta\theta, \tag{40}$$

where D is angular diameter distance and H is the expansion rate in the radial direction (see Fig. 4). The two scales are equal for the BAO feature in an FLRW background, where expansion is isotropic at all points. Any significant disagreement between the radial and transverse BAO scales would signal a breakdown of remote isotropy and thus of homogeneity:

$$\frac{L_\parallel}{L_\perp} - 1 \;\; \text{significantly different from } 0 \;\; \Rightarrow \;\; \text{non-FLRW universe.} \tag{41}$$

B. Sunyaev-Zeldovich effect: temperature of scattered CMB photons

Galaxy clusters with their hot ionized intra-cluster gas, act via scattering of CMB photons like giant mirrors that carry information about the last scattering surface seen by the cluster (see Fig. 5). In other words, clusters allow us in principle to indirectly probe inside our past lightcone.

CMB photons are scattered into our line of sight, thus inducing spectral distortions in the CMB temperature that we observe [25, 26]. The thermal Sunyaev-Zeldovich (SZ) effect from the thermal motion of electrons reflects the monopole seen by the cluster. If the blackbody temperatures of CMB photons arriving at the cluster from points inside our lightcone are significantly different from the blackbody temperature that we directly observe, then there will be a significant distortion. Such a signal would indicate that the cluster sees a significantly anisotropic CMB, hence violating the CP and homogeneity [27]. (This test has been applied to a class of LTB models by [28].)

The bulk radial motion of the cluster gas induces a kinetic SZ signal that reflects the CMB dipole seen by the cluster. If this is large, then there would be a violation of the CP and homogeneity. This has been applied to classes of LTB models by [29, 30], but it is in fact more generally applicable as a test of the CP.

In summary,

$$\text{Large (non-perturbative) thermal or kinetic SZ temperature effect} \quad \Rightarrow \quad \text{non-FLRW universe.} \tag{42}$$
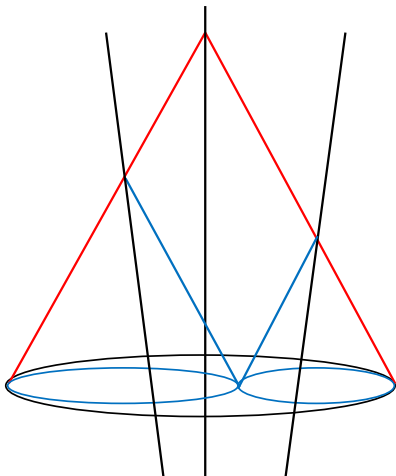

FIG. 5: CMB photons from inside our past lightcone are scattered into our line of sight by the ionized gas in galaxy clusters, thus carrying information about the level of isotropy seen by the clusters.

C. Sunyaev-Zeldovich effect: polarization of scattered CMB photons

Analogous to the SZ effect on CMB temperature, there is an SZ effect on CMB polarization [8]. The cluster bulk transverse velocity, and the CMB monopole, quadrupole and octupole, as seen by the cluster, induce modifications in CMB polarization via scattering off the cluster gas. These effects have been computed in FLRW models [31–33]. The SZ polarization signals potentially contain more information than the SZ temperature signals, and this has been proposed for the standard FLRW model as a way to map the CMB quadrupole seen at remote locations, thus lessening the cosmic variance [33–36].

From a more general viewpoint, the SZ effect on polarization is in principle a powerful probe of the CP and homogeneity. In principle, if CMB polarization data indicate large transverse cluster velocities, or large modifications to the primordial polarization signal, this could signal a violation of the CP and homogeneity:

$$\text{Large (non-perturbative) SZ polarization effects} \quad \Rightarrow \quad \text{non-FLRW universe.} \tag{43}$$

V. CONCLUSIONS

It is important to re-examine the basic assumptions of the standard concordance model, especially in view of the problems raised by the fine-tuned and unnatural nature of dark energy. Here we have re-examined the central assumption of homogeneity. We argued that homogeneity cannot be established or directly tested by observations, given the inherent limitations of lightcone-based data. By contrast, isotropy can be directly probed by observations. We constructed the observational coordinates necessary to answer the question of which observables on the past lightcone are needed to prove isotropy. It turns out that 4 independent matter observables – angular diameter distances, number counts, lensing distortion and transverse velocities – are exactly sufficient to determine the spacetime geometry. Isotropy of these 4 observables along one worldline imposes isotropy of the spacetime geometry. Surprisingly, isotropy of the CMB along one worldline does not in itself lead to isotropic geometry.

To establish homogeneity, we are forced to adopt the Copernican Principle (CP). Isotropy of the 4 matter observables for all observers leads to homogeneity – giving an observational version of the Cosmological Principle. A surprising and more powerful result is that isotropy of only the angular diameter distances for all observers, and only to $O(z^3)$, enforces homogeneity.

The strongest observational basis for homogeneity comes from the CP combined with the high isotropy of the CMB. The main additional assumption needed is that the CMB rest frame is geodesic. We outlined a covariant nonlinear proof and generalization of the original EGS result. And we highlighted the remarkable ETM result, i.e. that the vanishing of the dipole, quadrupole and octupole is sufficient to enforce homogeneity. The realistic case, with almost-isotropy of the CMB, does not lead to almost-homogeneity without additional assumptions on some of the derivatives of the multipoles. It remains an open problem whether these assumptions can be avoided, possibly using further information from almost-isotropy of matter observations.

Although we cannot directly observe homogeneity, we can test the CP at the foundation of homogeneity, using observations that carry information from inside our past lightcone. We described how this can be achieved via consistency relations between distances and expansion rates, using supernova and BAO data, and via Sunyaev-Zeldovich effects from galaxy clusters on the CMB temperature and polarization. Up to now, none of these tests has yet indicated a breakdown of the CP and thus a violation of homogeneity. But the further advance of high-precision data will provide new opportunities to apply and extend these critical tests.

In summary: the standard homogeneous model of cosmology is successful, predictive and up to now robust against all the observational data, and against current tests of the CP. The problem of a satisfactory explanation for dark energy may be resolved by advances in particle physics and quantum gravity. It is certainly reasonable to assume that the LCDM model is a good description of the Universe. Nevertheless it is also necessary to continuing probing the foundations of the model – not only the assumption of homogeneity, but also other critical questions, such as the problem of averaging and the need for a detailed understanding of light propagation in a lumpy Universe.

Appendix A: Nonlinear field and Boltzmann equations

For convenience we repeat the summary of the 1+3 covariant Lagrangian formulation of the field equations and Boltzmann equation in full nonlinear generality, as given in [1]. For a chosen 4-velocity field u^μ, the fundamental tensors are

$$h_{\mu\nu} = g_{\mu\nu} + u_\mu u_\nu, \quad \varepsilon_{\mu\nu\alpha} = \eta_{\mu\nu\alpha\beta} u^\beta, \tag{A1}$$

where $h_{\mu\nu}$ projects into the instantaneous rest space of comoving observers, and $\varepsilon_{\mu\nu\alpha}$ is the projection of the spacetime alternating tensor $\eta_{\mu\nu\alpha\beta} = -\sqrt{-g}\delta^0{}_{[\mu}\delta^1{}_\nu \delta^2{}_\alpha \delta^3{}_{\beta]}$. The projected symmetric tracefree (PSTF) parts of vectors and rank-2 tensors are

$$V_{\langle\mu\rangle} = h_\mu{}^\nu V_\nu, \quad S_{\langle\mu\nu\rangle} = \left\{ h_{(a}{}^\alpha h_{\nu)}{}^\beta - \frac{1}{3} h^{\alpha\beta} h_{\mu\nu} \right\} S_{\alpha\beta}. \tag{A2}$$

The skew part of a projected rank-2 tensor is spatially dual to the projected vector, $S_\mu = \frac{1}{2}\varepsilon_{\mu\nu\alpha} S^{[\nu\alpha]}$, and then any projected rank-2 tensor has the decomposition $S_{\mu\nu} = \frac{1}{3} S h_{\mu\nu} + \varepsilon_{\mu\nu\alpha} S^\alpha + S_{\langle\mu\nu\rangle}$, where $S = S_{\alpha\beta} h^{\alpha\beta}$. The covariant derivative ∇_μ defines 1+3 covariant time and spatial derivatives:

$$\dot{J}^{\mu\cdots}{}_{\cdots\nu} = u^\alpha \nabla_\alpha J^{\mu\cdots}{}_{\cdots\nu}, \quad \tilde{\nabla}_\alpha J^{\mu\cdots}{}_{\cdots\nu} = h_\alpha{}^\beta h^\mu{}_\kappa \cdots h_\nu{}^\tau \nabla_\beta J^{\kappa\cdots}{}_{\cdots\tau}. \tag{A3}$$

The projected derivative $\tilde{\nabla}_\mu$ defines a covariant PSTF divergence, $\tilde{\nabla}^\mu V_\mu$, $\tilde{\nabla}^\nu S_{\mu\nu}$, and a covariant PSTF curl,

$$\mathrm{curl}\, V_\mu = \varepsilon_{\mu\nu c} \tilde{\nabla}^\nu V^\alpha, \quad \mathrm{curl}\, S_{\mu\nu} = \varepsilon_{\alpha\beta(\mu} \tilde{\nabla}^\alpha S_{\nu)}{}^\beta. \tag{A4}$$

The relative motion of comoving observers is encoded in the PSTF kinematical quantities: the volume expansion rate, 4-acceleration, vorticity and shear, given respectively by

$$\Theta = \tilde{\nabla}^\mu u_\mu, \ A_\mu = \dot{u}_\mu, \ \omega_\mu = \mathrm{curl}\, u_\mu, \ \sigma_{\mu\nu} = \tilde{\nabla}_{\langle\mu} u_{\nu\rangle} \ \Leftrightarrow \ \nabla_\nu u_\mu = \frac{1}{3}\Theta h_{\mu\nu} + \varepsilon_{\mu\nu\alpha}\omega^\alpha + \sigma_{\mu\nu} - A_\mu u_\nu. \tag{A5}$$

Key nonlinear identities are

$$\mathrm{curl}\,\tilde{\nabla}_\mu \psi \ := \ \varepsilon_{\mu\nu\alpha}\tilde{\nabla}^\nu \tilde{\nabla}^\alpha \psi = -2\dot{\psi}\omega_\mu, \tag{A6}$$

$$h_\mu{}^\nu (\tilde{\nabla}_\nu \dot{\psi}) - \tilde{\nabla}_\mu \dot{\psi} \ = \ \dot{\psi} A_\mu - \left(\frac{1}{3}\Theta h_{\mu\nu} + \sigma_{\mu\nu} + \varepsilon_{\mu\nu\kappa}\omega^\kappa \right) \tilde{\nabla}^\nu \psi. \tag{A7}$$

The PSTF dynamical quantities describe the sources of the gravitational field: the (total) energy density $\rho = T_{\mu\nu} u^\mu u^\nu$ and isotropic pressure $p = \frac{1}{3} h_{\mu\nu} T^{\mu\nu}$ (including Λ), momentum density $q_\mu = -T_{\langle\mu\rangle\nu} u^\nu$, and anisotropic stress $\pi_{\mu\nu} = T_{\langle\mu\nu\rangle}$, where $T_{\mu\nu}$ is the total energy-momentum tensor. The Weyl tensor splits into PSTF gravito-electric and gravito-magnetic fields

$$E_{\mu\nu} = C_{\mu\alpha\nu\beta} u^\alpha u^\beta, \quad H_{\mu\nu} = \frac{1}{2}\varepsilon_{\mu\alpha\beta} C^{\alpha\beta}{}_{\nu\kappa} u^\kappa. \tag{A8}$$

The Ricci and Bianchi identities,

$$\nabla_{[\mu}\nabla_{\nu]} u_\alpha = R_{\mu\nu\alpha\beta} u^\beta, \quad \nabla^\beta C_{\mu\nu\alpha\beta} = -\nabla_{[\mu}\left\{ R_{\nu]\alpha} - \frac{1}{6} R g_{\nu]\alpha} \right\}, \tag{A9}$$

produce the fundamental evolution and constraint equations governing the covariant quantities. Einstein's equations are incorporated via the algebraic replacement $R^{\mu\nu} = T^{\mu\nu} - \frac{1}{2} T_\alpha{}^\alpha g^{\mu\nu}$. The resulting equations, in fully nonlinear form and for a general source of the gravitational field, are:

Evolution:

$$\dot{\rho} + (\rho + p)\Theta + \tilde{\nabla}^\mu q_\mu = -2A^\mu q_\mu - \sigma^{\mu\nu}\pi_{\mu\nu} \,, \tag{A10}$$

$$\dot{\Theta} + \frac{1}{3}\Theta^2 + \frac{1}{2}(\rho + 3p) - \tilde{\nabla}^\mu A_\mu = -\sigma_{\mu\nu}\sigma^{\mu\nu} + 2\omega_\mu\omega^\mu + A_\mu A^\mu \,, \tag{A11}$$

$$\dot{q}_{\langle\mu\rangle} + \frac{4}{3}\Theta q_\mu + (\rho + p)A_\mu + \tilde{\nabla}_\mu p + \tilde{\nabla}^\nu\pi_{\mu\nu} = -\sigma_{\mu\nu}q^\nu + \varepsilon_{\mu\nu\alpha}\omega^\nu q^\alpha - A^\nu\pi_{\mu\nu} \,, \tag{A12}$$

$$\dot{\omega}_{\langle\mu\rangle} + \frac{2}{3}\Theta\omega_\mu + \frac{1}{2}\mathrm{curl}\,A_\mu = \sigma_{\mu\nu}\omega^\nu \,, \tag{A13}$$

$$\dot{\sigma}_{\langle\mu\nu\rangle} + \frac{2}{3}\Theta\sigma_{\mu\nu} + E_{\mu\nu} - \frac{1}{2}\pi_{\mu\nu} - \tilde{\nabla}_{\langle\mu}A_{\nu\rangle} = -\sigma_{\alpha\langle\mu}\sigma_{\nu\rangle}{}^\alpha - \omega_{\langle\mu}\omega_{\nu\rangle} + A_{\langle\mu}A_{\nu\rangle} \,, \tag{A14}$$

$$\dot{E}_{\langle\mu\nu\rangle} + \Theta E_{\mu\nu} - \mathrm{curl}\,H_{\mu\nu} + \frac{1}{2}(\rho + p)\sigma_{\mu\nu} + \frac{1}{2}\dot{\pi}_{\langle\mu\nu\rangle} + \frac{1}{6}\Theta\pi_{\mu\nu} + \frac{1}{2}\tilde{\nabla}_{\langle\mu}q_{\nu\rangle}$$
$$= -A_{\langle\mu}q_{\nu\rangle} + 2A^\alpha\varepsilon_{\alpha\beta(\mu}H_{\nu)}{}^\beta + 3\sigma_{\alpha\langle\mu}E_{\nu\rangle}{}^\alpha - \omega^\alpha\varepsilon_{\alpha\beta(\mu}E_{\nu)}{}^\beta - \frac{1}{2}\sigma^\alpha{}_{\langle\mu}\pi_{\nu\rangle\alpha} - \frac{1}{2}\omega^\alpha\varepsilon_{\alpha\beta(\mu}\pi_{\nu)}{}^\beta \,, \tag{A15}$$

$$\dot{H}_{\langle\mu\nu\rangle} + \Theta H_{\mu\nu} + \mathrm{curl}\,E_{\mu\nu} - \frac{1}{2}\mathrm{curl}\,\pi_{\mu\nu} = 3\sigma_{\alpha\langle\mu}H_{\nu\rangle}{}^\alpha - \omega^\alpha\varepsilon_{\alpha\beta(\mu}H_{\nu)}{}^\beta$$
$$- 2A^\alpha\varepsilon_{\alpha\beta(\mu}E_{\nu)}{}^\alpha - \frac{3}{2}\omega_{\langle\mu}q_{\nu\rangle} + \frac{1}{2}\sigma^\alpha{}_{(\mu}\varepsilon_{\nu)\alpha\beta}q^\beta \,. \tag{A16}$$

Constraint:

$$\tilde{\nabla}^\mu\omega_\mu = A^\mu\omega_\mu \,, \tag{A17}$$

$$\tilde{\nabla}^\nu\sigma_{\mu\nu} - \mathrm{curl}\,\omega_\mu - \frac{2}{3}\tilde{\nabla}_\mu\Theta + q_\mu = -2\varepsilon_{\mu\nu\alpha}\omega^\nu A^\alpha \,, \tag{A18}$$

$$\mathrm{curl}\,\sigma_{\mu\nu} + \tilde{\nabla}_{\langle\mu}\omega_{\nu\rangle} - H_{\mu\nu} = -2A_{\langle\mu}\omega_{\nu\rangle} \,, \tag{A19}$$

$$\tilde{\nabla}^\nu E_{\mu\nu} + \frac{1}{2}\tilde{\nabla}^\nu\pi_{\mu\nu} - \frac{1}{3}\tilde{\nabla}_\mu\rho + \frac{1}{3}\Theta q_\mu = \varepsilon_{\mu\nu\alpha}\sigma^\nu{}_\beta H^{\alpha\beta} - 3H_{\mu\nu}\omega^\nu + \frac{1}{2}\sigma_{\mu\nu}q^\nu - \frac{3}{2}\varepsilon_{\mu\nu\alpha}\omega^\nu q^\alpha \,, \tag{A20}$$

$$\tilde{\nabla}^\nu H_{\mu\nu} + \frac{1}{2}\mathrm{curl}\,q_\mu - (\rho + p)\omega_\mu = -\varepsilon_{\mu\nu\alpha}\sigma^\nu{}_\beta E^{\alpha\beta} - \frac{1}{2}\varepsilon_{\mu\nu\alpha}\sigma^\nu{}_\beta\pi^{\alpha\beta} + 3E_{\mu\nu}\omega^\nu - \frac{1}{2}\pi_{\mu\nu}\omega^\nu \,. \tag{A21}$$

The energy and momentum conservation equations are (A10) and (A12). The total dynamical quantities have contributions from all dynamically significant particle species:

$$T^{\mu\nu} = \sum_I T_I^{\mu\nu} = \rho u^\mu u^\nu + ph^{\mu\nu} + 2q^{(\mu}u^{\nu)} + \pi^{\mu\nu} \,, \quad T_I^{\mu\nu} = \rho_I^* u_I^\mu u_I^\mu + p_I^* h_I^{\mu\nu} + 2q_I^{*(\mu}u_I^{\nu)} + \pi_I^{*\mu\nu} \,, \tag{A22}$$

where $I = \mathrm{r,n,b,c,\Lambda}$ labels the species. The asterisk on the dynamical quantities $\rho_I^*, \cdots$ denotes that these quantities are measured, not in the u^μ-frame, but in the I-frame, whose 4-velocity is given by

$$u_I^\mu = \gamma_I\left(u^\mu + v_I^\mu\right) \,, \quad v_I^\mu u_\mu = 0 \,, \quad \gamma_I = \left(1 - v_I^2\right)^{-1/2} \,. \tag{A23}$$

The fully nonlinear equations for the I dynamical quantities as measured in the fundamental u^μ-frame are:

$$\rho_I = \rho_I^* + \left\{\gamma_I^2 v_I^2\left(\rho_I^* + p_I^*\right) + 2\gamma_I q_I^{*\mu}v_{I\mu} + \pi_I^{*\mu\nu}v_{I\mu}v_{I\nu}\right\} \,, \tag{A24}$$

$$p_I = p_I^* + \frac{1}{3}\left\{\gamma_I^2 v_I^2\left(\rho_I^* + p_I^*\right) + 2\gamma_I q_I^{*\mu}v_{I\mu} + \pi_I^{*\mu\nu}v_{I\mu}v_{I\nu}\right\} \,, \tag{A25}$$

$$q_I^\mu = q_I^{*\mu} + (\rho_I^* + p_I^*)v_I^\mu + \left\{(\gamma_I - 1)q_I^{*\mu} - \gamma_I q_I^{*\nu}v_{I\nu}u^\mu + \gamma_I q_I^{*\nu}v_{I\nu}v_I^\mu\right.$$
$$\left. + \gamma_I^2 v_I^2\left(\rho_I^* + p_I^*\right)v_I^\mu + \pi_I^{*\mu\nu}v_{I\nu} - \pi_I^{*\nu\alpha}v_{I\nu}v_{I\alpha}u^\mu\right\} \,, \tag{A26}$$

$$\pi_I^{\mu\nu} = \pi_I^{*\mu\nu} + \left\{-2u^{(\mu}\pi_I^{*\nu)\alpha}v_{I\alpha} + \pi_I^{*\nu\alpha}v_{I\nu}v_{I\alpha}u^\mu u^\nu + 2\gamma_I v_I^{\langle\mu}q_I^{*\nu\rangle}\right.$$
$$\left. - 2\gamma_I q_I^{*\alpha}v_{I\alpha}u^{(\mu}v_I^{\nu)} - \frac{1}{3}\pi_I^{*\alpha\beta}v_{I\alpha}v_{I\beta}h^{\mu\nu} + \gamma_I^2\left(\rho_I^* + p_I^*\right)v_I^{\langle\mu}v_I^{\nu\rangle}\right\} \,. \tag{A27}$$

The terms in braces are the nonlinear corrections that vanish in the standard perturbed FLRW case. The total dynamical quantities in (A10)–(A21), are given by

$$\rho = \sum_I \rho_I \,, \quad p = \sum_I p_I \,, \quad q^\mu = \sum_I q_I^\mu \,, \quad \pi^{\mu\nu} = \sum_I \pi_I^{\mu\nu} \,. \tag{A28}$$

Assuming that the species are non-interacting, they each separately obey the energy and momentum conservation equations (A10) and (A12):

$$\dot{\rho}_I + (\rho_I + p_I)\Theta + \tilde{\nabla}_\mu q_I^\mu = -2A_\mu q_I^\mu - \sigma_{\mu\nu}\pi_I^{\mu\nu}\,, \tag{A29}$$

$$\dot{q}_I^{\langle\mu\rangle} + \frac{4}{3}\Theta q_I^\mu + (\rho_I + p_I)A^\mu + \tilde{\nabla}^\mu p_I + \tilde{\nabla}_\nu\pi_I^{\mu\nu} = -\sigma^\mu{}_\nu q_I^\nu + \varepsilon^\mu{}_{\nu\alpha}\omega^\nu q_I^\alpha - A_\nu\pi_I^{\mu\nu}\,, \tag{A30}$$

where the I-quantities are given by (A24)–(A27).

The covariant kinetic theory description starts by splitting the photon 4-momentum as

$$p^\mu = E(u^\mu + e^\mu)\,,\quad e^\mu e_\mu = 1\,,\ e^\mu u_\mu = 0\,. \tag{A31}$$

Here $E = -u_\mu p^\mu$ is the energy and $e^\mu = p^{\langle\mu\rangle}/E$ is the direction, as measured by a comoving fundamental observer. Then the photon distribution function is decomposed into covariant harmonics via the expansion

$$f(x,p) = f(x,E,e) \ = \ F + F_\mu e^\mu + F_{\mu\nu}e^\mu e^\nu + \cdots = \sum_{\ell\geq 0} F_{M_\ell}(x,E)e^{\langle M_\ell\rangle}\,, \tag{A32}$$

where $M_\ell := \mu_1\mu_2\cdots\mu_\ell$ and $e^{M_\ell} := e^{\mu_1}\cdots e^{\mu_\ell}$. The PSTF multipoles F_{M_ℓ} are a covariant alternative to the usual expansion in spherical harmonics. The energy-momentum tensor is $T_{\rm r}^{\mu\nu}(x) = \int p^\mu p^\nu f(x,p)\mathrm{d}^3 p$, so the first 3 multipoles define the radiation dynamical quantities (in the u^μ-frame):

$$\rho_{\rm r} = 4\pi\int_0^\infty E^3 F\,\mathrm{d}E\,,\ \ q_{\rm r}^\mu = \frac{4\pi}{3}\int_0^\infty E^3 F^\mu\,\mathrm{d}E\,,\ \ \pi_{\rm r}^{\mu\nu} = \frac{8\pi}{15}\int_0^\infty E^3 F^{\mu\nu}\,\mathrm{d}E\,. \tag{A33}$$

Thus $\Pi = (1/4\pi)\rho_{\rm r}, \Pi^\mu = (3/4\pi)q_{\rm r}^\mu, \Pi^{\mu\nu} = (15/8\pi)\pi_{\rm r}^{\mu\nu}$, where the brightness multipoles are

$$\Pi_{\mu_1\cdots\mu_\ell} = \int E^3 F_{\mu_1\cdots\mu_\ell}\mathrm{d}E\,. \tag{A34}$$

The collisionless Boltzmann (or Liouville) equation is

$$\frac{\mathrm{d}f}{\mathrm{d}v} := p^\mu\frac{\partial f}{\partial x^\mu} - \Gamma^\mu{}_{\alpha\beta}p^\alpha p^\beta\frac{\partial f}{\partial p^\mu} = 0\,, \tag{A35}$$

where $p^\mu = \mathrm{d}x^\mu/\mathrm{d}v$, and we neglect polarization. The covariant multipoles of $\mathrm{d}f/\mathrm{d}v$ are given by

$$\begin{aligned}
\frac{1}{E}\left(\frac{\mathrm{d}f}{\mathrm{d}v}\right)_{M_\ell} =\ & \dot{F}_{\langle M_\ell\rangle} - \frac{1}{3}\Theta E F'_{M_\ell} + \tilde{\nabla}_{\langle\mu_\ell}F_{M_{\ell-1}\rangle} + \frac{(\ell+1)}{(2\ell+3)}\tilde{\nabla}^\mu F_{\mu M_\ell} \\
& - \frac{(\ell+1)}{(2\ell+3)}E^{-(\ell+1)}\left[E^{\ell+2}F_{\mu M_\ell}\right]'A^\mu - E^\ell\left[E^{1-\ell}F_{\langle M_{\ell-1}}\right]'A_{\mu_\ell\rangle} \\
& - \ell\omega^\nu\varepsilon_{\nu\alpha(\mu_\ell}F_{M_{\ell-1})}{}^\alpha - \frac{(\ell+1)(\ell+2)}{(2\ell+3)(2\ell+5)}E^{-(\ell+2)}\left[E^{\ell+3}F_{\mu\nu M_\ell}\right]'\sigma^{\mu\nu} \\
& - \frac{2\ell}{(2\ell+3)}E^{-1/2}\left[E^{3/2}F_{\nu\langle M_{\ell-1}}\right]'\sigma_{\mu_\ell\rangle}{}^\nu - E^{\ell-1}\left[E^{2-\ell}F_{\langle M_{\ell-2}}\right]'\sigma_{\mu_{\ell-1}\mu_\ell\rangle}\,,
\end{aligned} \tag{A36}$$

where a prime denotes $\partial/\partial E$. This is a fully nonlinear expression. Multiplying by E^3 and integrating over all energies leads to the brightness multipole evolution equations:

$$\begin{aligned}
0 =\ & \dot{\Pi}_{\langle M_\ell\rangle} + \frac{4}{3}\Theta\Pi_{M_\ell} + \tilde{\nabla}_{\langle\mu_\ell}\Pi_{M_{\ell-1}\rangle} + \frac{(\ell+1)}{(2\ell+3)}\tilde{\nabla}^\nu\Pi_{\nu M_\ell} \\
& - \frac{(\ell+1)(\ell-2)}{(2\ell+3)}A^\nu\Pi_{\nu M_\ell} + (\ell+3)A_{\langle\mu_\ell}\Pi_{M_{\ell-1}\rangle} \\
& - \ell\omega^\nu\varepsilon_{\nu\alpha(\mu_\ell}\Pi_{M_{\ell-1})}{}^\alpha - \frac{(\ell-1)(\ell+1)(\ell+2)}{(2\ell+3)(2\ell+5)}\sigma^{\nu\alpha}\Pi_{\nu\alpha M_\ell} \\
& + \frac{5\ell}{(2\ell+3)}\sigma^\nu{}_{\langle\mu_\ell}\Pi_{M_{\ell-1}\rangle\nu} - (\ell+2)\sigma_{\langle\mu_\ell\mu_{\ell-1}}\Pi_{M_{\ell-2}\rangle}\,.
\end{aligned} \tag{A37}$$

The monopole equation is just the energy conservation equation, i.e., (A29) with $I = r$, the dipole is the momentum conservation equation (A30), with $I = r$, and the quadrupole is (19).

[1] C. Clarkson and R. Maartens, Class. Quant. Grav. **27** (2010) 124008 [arXiv:1005.2165].
[2] G. F. R. Ellis, arXiv:1103.2335.
[3] P. Sarkar, J. Yadav, B. Pandey and S. Bharadwaj, arXiv:0906.3431.
[4] F. Sylos Labini, arXiv:1103.5974.
[5] R. Maartens, PhD thesis, University of Cape Town (1980).
[6] G. F. R. Ellis, S. D. Nel, R. Maartens, W. R. Stoeger and A. P. Whitman, Phys. Rept. **124** (1985) 315.
[7] R. Maartens and D. R. Matravers, Class. Quantum Grav. **11** (1994) 2693.
[8] R. A. Sunyaev and Y. B. Zeldovich, Mon. Not. Roy. Astron. Soc. **190**, 413 (1980).
[9] W. Hasse and V. Perlick, Class. Quantum Grav. **16** (1999) 2559.
[10] C. A. Clarkson, Ph.D. thesis, University of Glasgow (1999) [arXiv:astro-ph/0008089].
[11] J. Kristian and R. K. Sachs, Astrophys. J. **143** (1966) 379.
[12] J. Ehlers, P. Geren and R. K. Sachs, J. Math. Phys. **9** (1968) 1344.
[13] W. R. Stoeger, R. Maartens and G. F. R. Ellis, Astrophys. J. **443** (1995) 1.
[14] R. Treciokas and G. F. R. Ellis, Commun. Math. Phys. **23** (1971) 1.
[15] J. J. Ferrando, J. A. Morales and M. Portilla, Phys. Rev. D **46** (1999) 578.
[16] C. A. Clarkson and R. K. Barrett, Class. Quantum Grav. **16** (1999) 3781 [arXiv:gr-qc/9906097].
[17] C. A. Clarkson and A. A. Coley, Class. Quant. Grav. **18** (2001) 1305 [arXiv:astro-ph/0102172].
[18] S. Rasanen, Phys. Rev. D **79** (2009) 123522 [arXiv:0903.3013].
[19] G. F. R. Ellis, R. Treciokas and D. R Matravers, Ann. Phys. **150** (1983) 487.
[20] R. Maartens, T. Gebbie and G. F. R. Ellis, Phys. Rev. D **59** (1999) 083506 [arXiv:astro-ph/9808163].
[21] R. Maartens, G. F. R. Ellis and W. R. Stoeger, Phys. Rev. D **51** (1995) 1525 [arXiv:astro-ph/9501016].
[22] S. Räsänen, JCAP **0902** (2009) 011 [arXiv:0812.2872].
[23] C. Clarkson, B. Bassett and T. H.-C. Lu, Phys. Rev. Lett. **101** (2008) 011301 [arXiv:0712.3457].
[24] J.-P. Uzan, C. Clarkson and G.F.R. Ellis, Phys. Rev. Lett. **100** (2008) 191303 [arXiv:0801.0068].
[25] R. A. Sunyaev and Y. B. Zeldovich, Astrophys. Space Sci. **7** (1970) 3.
[26] R. A. Sunyaev and Y. B. Zeldovich, Comments Astrophys. Space Phys. **4** (1972) 173.
[27] J. Goodman, Phys. Rev. D **52** (1995) 1821 [arXiv:astro-ph/9506068].
[28] R. R. Caldwell and A. Stebbins, Phys. Rev. Lett. **100** (2008) 191302 [arXiv:0711.3459].
[29] J. Garcia-Bellido and T. Haugbølle, JCAP **0804** (2008) 003 [arXiv:0807.1326].
[30] P. Zhang and A. Stebbins, arXiv:1009.3967.
[31] S. Y. Sazonov and R. A. Sunyaev, Mon. Not. Roy. Astron. Soc. **310**, 765 (1999) [arXiv:astro-ph/9903287].
[32] A. Challinor, M. Ford and A. Lasenby, Mon. Not. Roy. Astron. Soc. **312** (2000) 159 [arXiv:astro-ph/9905227].
[33] M. Kamionkowski and A. Loeb, Phys. Rev. D **56**, 4511 (1997) [arXiv:astro-ph/9703118].
[34] N. Seto and M. Sasaki, Phys. Rev. D **62**, 123004 (2000) [arXiv:astro-ph/0009222].
[35] D. Baumann and A. Cooray, New Astron. Rev. **47**, 839 (2003) [arXiv:astro-ph/0304416].
[36] E. F. Bunn, Phys. Rev. D **73**, 123517 (2006) [arXiv:astro-ph/0603271].