Structure of structure formation theories

Wayne Hu and Daniel J. Eisenstein
Institute for Advanced Study, Princeton, New Jersey 08540
(Received 2 October 1998; published 23 March 1999)

We study the general structure of models for structure formation, with applications to the reverse engineering of the model from observations. Through a careful accounting of the degrees of freedom in covariant gravitational instability theory, we show that the evolution of structure is completely specified by the stress history of the dark sector. The study of smooth, entropic, sonic, scalar anisotropic, vector anisotropic, and tensor anisotropic stresses reveals the origin, robustness, and uniqueness of specific model phenomenology. We construct useful and illustrative analytic solutions that cover cases with multiple species of differing equations of state relevant to the current generation of models, especially those with effectively smooth components. We present a simple case study of models with phenomenologies similar to that of a ΛCDM model to highlight reverse-engineering issues. A critical-density universe dominated by a single type of dark matter with the appropriate stress history can mimic a ΛCDM model exactly. [S0556-2821(99)04208-3]

PACS number(s): 95.35.+d

I. INTRODUCTION

How does one reverse engineer a model for structure formation from observed phenomena? How unique is such an inversion? How robust are the phenomenological distinctions between broad classes of models? With the wealth of high-precision cosmological data expected in the near future from the cosmic microwave background (CMB), galaxy surveys, and the high redshift universe, the simple *ab initio* models for structure formation currently considered may be ruled out, forcing us to confront these difficult issues. In this paper, we take the first steps toward answering this question by examining from a general standpoint what makes a model for structure formation behave as it does in linear perturbation theory.

A model for structure formation is completely specified by its initial conditions and the full temporal and spatial behavior of the stresses in its dark sector. The dark sector contains the elements in the model that do not interact with the photons at any observable redshift. It can include, but is not limited to, cold dark matter (CDM), neutrinos, and cosmological defects.

Unfortunately, the stress history of the dark sector is by definition not directly observable. Its effects come filtered through gravity as mediated by metric fluctuations. The translation of metric fluctuations into observables in the CMB and evolution of structure is well understood. Therefore, the main hurdle in the task of reconstructing a model from observations is to understand how stress histories translate into metric fluctuations and vice versa.

Our general philosophy here is to start from elements of the cosmological model that will likely survive the onslaught of data: general relativity and a universe whose deviations from homogeneity and isotropy are initially small. We proceed down the theory pipeline to existing models of structure formation, making explicit the places where assumptions are made and hence could be altered. Where possible, we provide analytic solutions and approximations that highlight certain generic behavior and phenomena. These solutions are useful for describing the behavior of the existing models and are in most cases new or substantially more general than those found in the literature. In particular, we derive master solutions for models which contain multiple components with arbitrary equations of state and smooth, entropic, anisotropic, and sonic stresses.

We begin in Sec. II with an overview of the basic elements of a structure formation theory and their traditional classification in terms of their initial conditions, perturbation type, and clustering properties. We then present a concise but general treatment of linear perturbation theory in Sec. III and gauge issues in Sec. IV. Although these are well-studied subjects (see e.g. [1–3]), our treatment has several pedagogical and practical virtues. It keeps careful track of the degrees of freedom available to structure formation models and hence provides a unified treatment applicable to all models, including those containing exotic matter like scalar fields or cosmological defects. We also explicitly maintain general covariance such that the equations apply, and may be easily specialized, to any choice of coordinates or gauge. In Sec. V, we define general classes of stress perturbations and present an overview of their conversion into observables.

The remainder of the paper deals with stress histories on a case by case basis. The simplest case involves smooth stresses, and we present detailed analytic solutions in Sec. VI that apply to a wide range of models—from simple cosmological constant and cold dark matter models (ΛCDM) to massive neutrino and scalar field models. Pure anisotropic, entropic, and sonic stresses are treated in Sec. VII and mixed cases in Sec. VIII. To highlight reverse-engineering issues, we study single-component, critical-density dark matter models with phenomenologies that mimic the ΛCDM model in Sec. IX. We conclude in Sec. X by re-examining the traditional classification scheme of Sec. II in light of the phenomenological distinctions uncovered in this work.

II. CLASSIFICATION OF THEORIES

A. Initial conditions

Perhaps the most fundamental difference between models for structure formation lies with their initial conditions. Cur-

rently, inflation is the only known means of laying down large-scale density or curvature perturbations in the early universe. Indeed, inflation in the more general sense of a period of superluminal expansion is required for the causal generation of large-scale power. It provides a means of connecting parts of the universe that are currently space-like separated, i.e. outside the current particle horizon. Models with initial curvature perturbations are usually called ''adiabatic'' models.

All other causal models begin with no density or curvature fluctuations on large scales and are hence called ''isocurvature'' models. In these models, stress gradients causally move matter around inside the horizon to form large-scale structure.

The generation mechanism is also responsible for determining the spectrum and statistics of the fluctuations. The simplest inflationary models predict a nearly scale-invariant and Gaussian distribution of fluctuations [4] but higher order effects can break scale-invariance and generate non-Gaussianity [5,6]. Defect perturbations are intrinsically non-Gaussian but are typically also scale-invariant in the generalized sense of ''scaling'' [7]. We are primarily concerned here with the evolution of fluctuations from their initial state through the linear regime and do not consider these issues further. Note that changes in the spectrum of perturbations are simple to include in linear theory as evolutionary effects can be factored out into so-called ''transfer functions.''

B. Perturbation type

The perturbation type for the metric and matter fluctuations is the next most important distinction. A general linear fluctuation can be decomposed into scalar, vector and tensor components. These manifest themselves as density, vorticity, and gravitational wave perturbations respectively and do not interact in linear theory. The scalar modes are the only ones that grow through gravitational instability. Vector modes, on the other hand, always decay with the expansion. They can only be actively generated by shearing (or anisotropic) stress in the manner. Tensor modes are intermediate. Left to themselves, they propagate as gravity waves, but they generate and can be generated by transverse-traceless (quadrupolar) stresses in the matter.

The simplest inflationary models possess only scalar (''S'') fluctuations, and tensor fluctuations are generally cosmologically negligible in models with energy scales substantially below the Planck scale [8]. However, models whose initial conditions contain both scalar and tensor (''ST'') fluctuations are possible. Models with only S or ST generally have stresses that may be defined as functions of the metric, density and velocity perturbations and hence may be viewed as ''passive'' responses through equations of state.

''Active'' models have stresses that are a consequence of complex internal dynamics in the dark sector that cannot be simply specified as responses to gravitational perturbations. Although this definition is not precise for scalar and tensor modes, the very presence of vector modes indicates an active source because these must be continuously generated to have an observable effect. Nevertheless, such models generally have all three types of perturbations (''SVT'').

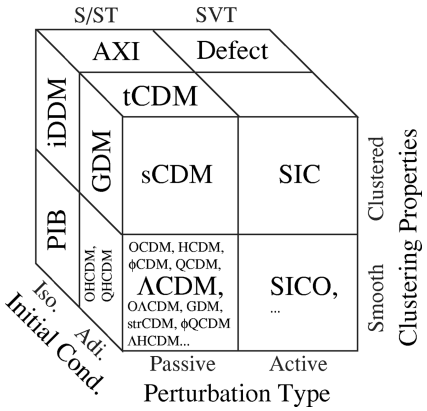

FIG. 1. Taxonomy of structure formation. Models can be classified by their initial conditions (adiabatic or isocurvature), perturbation type (passive or active), and clustering properties of the dark matter (clustered or smooth on large-scale structure scales). Passive fluctuations involve stress responses to other perturbations and can support scalar (''S'') or scalar and tensor (''ST'') components. Active stresses generate fluctuations and generally possess vector components as well (''SVT''). We will examine the extent to which this traditional categorization is useful in predicting model phenomenology.

C. Clustering properties of dark matter

Finally, the nature of the dark components affects the evolution of perturbations. We define as ''dark'' any component that interacts with the CMB photons only gravitationally. Thus, even massless neutrinos are classified as dark matter in this scheme.

Stresses in the dark components change the evolution of the mean density with time and the response of the matter to gravitational compression. We will loosely type models whose expansion rate is driven by a compressible type of matter (on scales relevant to cosmological structures) as ''clustered'' models and those which possess matter that is incompressible as ''smooth'' models. We shall see that this distinction is in fact rather inexact as it is not time invariant: essentially all models pass through phases when they would be considered smooth or clustered on the relevant scale.

D. Phenomenology

The key to understanding the phenomenology of a given model for structure formation is the evolution of metric fluctuations, in particular the Newtonian gravitational potential. Its qualitative behavior is determined by the initial conditions, perturbation type, and dark matter content of the model. We illustrate this taxonomy scheme in Fig. 1.

The behavior of the gravitational potential is directly related to the evolution of density perturbations through the Poisson equation. Once its evolution is determined as a function of scale, not only is the present large-scale structure of the universe determined but also the whole time history of structure formation. The latter is important for predicting the properties and abundances of high-redshift objects.

The gravitational potential also generates CMB anisotropies through gravitational redshifts [9] and is the ultimate source of all anisotropies from scalar perturbations. CMB phenomenology can be essentially read off of the time evolution of the gravitational potential [10], although this involves understanding the back reaction from density perturbations in the CMB itself [11].

Similar but simpler considerations apply for vector and tensor metric perturbations. They also generate anisotropies via gravitational redshifts but do not have unstable modes and hence do not affect large-scale structure formation in linear theory.

The difference between adiabatic and isocurvature models plays a direct role in metric evolution because initial curvature (or gravitational potential) perturbations are present in one and absent in the other. The perturbation type changes the ratio of CMB anisotropies to large-scale structure. Finally, the dark matter properties affect the evolution of the gravitational potentials. Smooth components by definition do not contribute to the gravitational potential but do contribute to the expansion rate. They slow down the growth of structure and cause the gravitational potential to decay. Hence they decrease the amount of structure and increase the large-angle anisotropies of the CMB.

In summary, the observable properties of structure formation models are encapsulated in the time evolution of the metric fluctuations. This in turn is governed by the stress properties of the matter through both its initial conditions and intrinsic properties.

E. Current model zoo

The archetypal model for structure formation is the standard cold dark matter model (sCDM), which is an adiabatic, passive, and clustered model. Here, an initial scale-invariant spectrum of adiabatic scalar (''S'') perturbations collapses via the gravitational instability of pressureless cold dark matter. Although this model is no longer viable from an observational standpoint, it predicts phenomena sufficiently similar to the observations to act as a good starting point for model building. One of its failings is that it predicts too much small-scale power for the level of CMB anisotropies demanded by the Cosmic Background Explorer (COBE) detection.

A simple variation of the sCDM model that attempts to address this problem involves tilting the initial spectrum of scalar perturbations (tCDM) to reduce small-scale power relative to large. Under certain inflationary scenarios, this brings about the addition of tensor perturbations that further reduce small-scale density perturbations relative to the COBE detection. Such a model would be a ''ST'' variant of SCDM.

The second class of variations involves changing the matter content so as to suppress the clustering of matter, yielding a ''smooth'' variant. The prototypical example is the ΛCDM model, where an additional component of matter that does not cluster replaces most of the CDM. Another example is the ''open'' OCDM model where spatial curvature plays the role of the smooth component. Those two represent examples where the additional component is smooth on all scales and for all time by definition. Variants where the matter is only smooth on small scales include the hot and cold dark matter (HCDM) model [e.g. [12] also called C+HDM and mixed dark matter (MDM)] with a component of hot dark matter, the ϕCDM model [13] with a scalar field component ϕ that tracks the background behavior of the matter, and QCDM [14,15] with a general scalar field (''quintessence''). In a string-dominated universe (strCDM) [16], the string network plays the role of a smooth component with the same equation of state as spatial curvature. Of course, one can have multiple smooth species as well, e.g. OΛCDM. The generalized DM (GDM) class of models [17] phenomenologically parametrizes all such models.

Replacing the CDM with GDM of a different equation of state but no stress perturbations is a phenomenological possibility (GDM) suggested by [17]. This is an adiabatic, passive, and clustered variant of CDM. We will use the designation ''CDMv'' to represent all such variants of the CDM model.

Isocurvature models have been proposed as alternatives to the ''CDMv'' class of models. The simplest examples are those in which the initial stress fluctuations are established by balancing the density perturbations of two different types of matter. Examples include the axion isocurvature (AXI) [18] model and the primordial isocurvature baryon (PIB) model [19], where radiation density fluctuations are balanced by axions and baryons, respectively. The simplest versions involve only scalar fluctuations and hence are passive (''S'') models. Models with and without smooth cosmological constant or spatial curvature components have been proposed. Versions with Gaussian power-law initial conditions are observationally challenged [18,21] but more complicated variations exist [20]. Based on our work, one of us has constructed an isocurvature decaying dark matter (iDDM) model that defies conventional wisdom on isocurvature models and solves these observational problems [22].

Finally, topological defect models such as strings and textures fall into the isocurvature class but have fluctuations that are active (''SVT''). The simplest versions obey scaling and

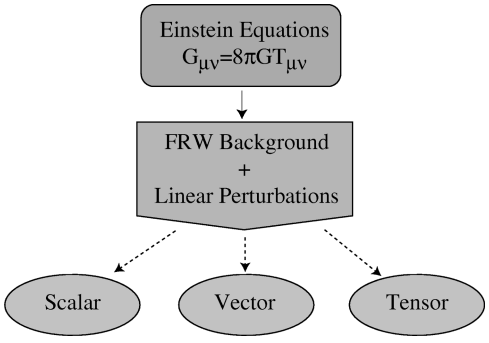

FIG. 2. Scalar, vector, tensor decomposition. At the top of the tree of possibilities for structure formation models is the assumption that general relativity holds in the cosmological context and the universe is homogeneous and isotropic in the mean with linear perturbations initially. Without further assumptions, the linear fluctuations may be expanded in scalar, vector, and tensor modes that do not interact while the fluctuations remain linear.

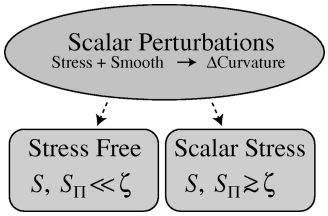

FIG. 3. Scalar perturbations. It is useful to subdivide the stress-free class of scalar perturbations from the general possibilities. A ''stress-free'' perturbation has dimensionless stresses (S,S_Π) that are much smaller than the comoving curvature perturbation ζ. Note that the stress-free perturbation condition does not preclude background or ''smooth'' stress.

have only clustering matter but fail to generate enough large-scale structure for the observed CMB anisotropies [23,24]. Models with a smooth Λ component have been proposed to alleviate these problems [25].

Hybrid models can also be constructed. If defects form after the inflationary epoch, one has a model with adiabatic initial conditions and active perturbations. A string model with inflation and cold dark matter (SIC) is a concrete example [26]. One can also add in smooth components, e.g. spatial curvature (SICO).

Clearly, the existing models do not even qualitatively exhaust the possibilities open to structure formation models. In the rest of the paper, we conduct an examination of these possibilities beginning with general principles and explicitly stating the assumptions that are made in obtaining the models described as well as their generalizations. We summarize this analysis in a series of flowcharts (Figs. 2, 3, 4, 5, 6, and 12.)

III. COVARIANT PERTURBATION THEORY

A. General definitions

We assume that the background is described by a Friedmann-Robertson-Walker (FRW) metric $\bar{g}_{\mu\nu}=a^2\gamma_{\mu\nu}$ with scale factor $a(t)$ normalized to unity today and constant comoving curvature in the spatial metric γ_{ij}. Here greek

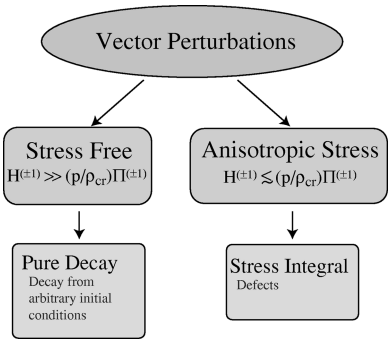

FIG. 4. Vector perturbations. Vector perturbations simply decay from their initial value in the stress-free limit. The integral solution in the presence of vector stress is given in Sec. VIII F and applies to defect models.

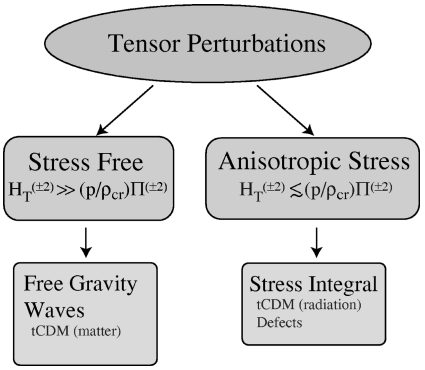

FIG. 5. Tensor perturbations. Tensor perturbations propagate as free gravity waves in the stress-free limit as is the case of a matter-dominated expansion. The integral solution in the presence of stresses is given in Sec. VIII G and may be applied to propagation during radiation domination as well as defect sources.

indices run from 0 to 3 while latin indices run over the spatial part of the metric: $i,j=1,2,3$. The component corresponding to conformal time

$$x^0 \equiv \eta = \int \frac{dt}{a(t)} \tag{1}$$

is $\gamma_{00}=-1$ and $\gamma_{0i}=\gamma_{i0}=0$. Unless otherwise specified, overdots represent derivatives with respect to conformal time and primes derivatives with respect to $\ln a$. $c=1$ throughout.

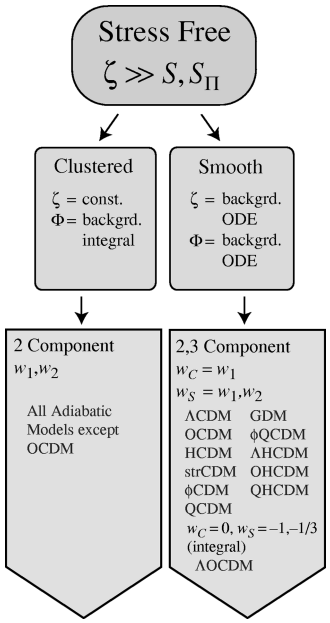

FIG. 6. Stress-free scalar perturbations. If all of the matter which drives the expansion participates in gravitational instability, the matter is said to be ''clustered''; otherwise, there exists a ''smooth'' component. In either case, the perturbations depend only on the background (''backgnd.'') equation of state w and can be tracked with simple techniques (middle row), which lead to exact solutions (bottom row) that describe behavior in a wide range of models.

The background curvature is given by $K=-H_0^2(1-\Omega_{\rm tot})$, where the Hubble constant is $H_0=100h\ {\rm km\,s^{-1}\ Mpc^{-1}}$.

The ten degrees of freedom for the perturbations in the symmetric metric tensor $g_{\mu\nu}$ can be parametrized as

$$g^{00}=-a^2(1-2A),$$
$$g^{0i}=-a^2B^i,$$
$$g^{ij}=a^2(\gamma^{ij}-2H_L\gamma^{ij}-2H_T^{ij}). \tag{2}$$

We refer to the lapse A as the potential, the three components of B_i as the metric shift, H_L as the curvature perturbation, and the five components of H_T^{ij} as the metric shear following the conventions of [1,2].

Likewise, the symmetric stress-energy tensor can be parametrized by ten components

$$T_0^0=-\rho-\delta\rho,$$
$$T_i^0=(\rho+p)(v_i-B_i),$$
$$T_0^i=-(\rho+p)v^i,$$
$$T_j^i=(p+\delta p)\delta_j^i+p\Pi_j^i, \tag{3}$$

i.e. the energy density and its perturbation $(\rho+\delta\rho)$, the isotropic stress (pressure) and its perturbation $(p+\delta p)$, the three components of the momentum density $(\rho+p)v_i$, and the five components of the anisotropic stress tensor Π_{ij}. Note that the metric shift B_i enters in T_i^0 but not T_0^i. Correspondingly, we shall see that B_i enters into the momentum but not the energy conservation equation.

By writing the metric and stress energy tensor in this form, we have maintained general covariance. As a result, the equations of motion that result below take the same *form* for any coordinate system where linear perturbation theory holds. We reserve the term gauge invariant refers for objects that have the same *value* in each frame.

B. Perturbation representation

While the perturbations are linear, they may be separated by their transformation properties under rotation without loss of generality (see Fig. 2). For covariant techniques that do not assume linear perturbations from the outset, see [27,28] and references therein. The five component metric shear H_T^{ij} and matter anisotropic stress Π^{ij} separate into one scalar, two vector, and two tensor components. The scalar stress generates potential flows $(\nabla\times\mathbf{v}=0)$, whereas the vector stress generates vorticity $(\nabla\cdot\mathbf{v}=0)$. The tensor stress generates tensor shear in the metric, which represents gravity waves in the transverse-traceless gauge. Therefore, the potential A, curvature H_L, density $\delta\rho$, and pressure δp perturbations are associated with scalar fluctuations alone, the metric shift B_i and velocity v_i with scalar and vector fluctuations, and the metric shear H_T^{ij} and anisotropic stress Π^{ij} with all three. Scalar, vector and tensor perturbations may be treated independently in linear perturbation theory.

Fluctuations can be decomposed into the normal modes of the Laplacian operator [1]

$$\nabla^2Q^{(0)}=-k^2Q^{(0)}\ \ {\rm S},$$
$$\nabla^2Q_i^{(\pm1)}=-k^2Q_i^{(\pm1)}\ \ {\rm V},$$
$$\nabla^2Q_{ij}^{(\pm2)}=-k^2Q_{ij}^{(\pm2)}\ \ {\rm T}, \tag{4}$$

where vector and tensor modes satisfy a divergenceless and transverse-traceless condition respectively:

$$\nabla^iQ_i^{(\pm1)}=0,\quad \gamma^{ij}Q_{ij}^{(\pm2)}=\nabla^iQ_{ij}^{(\pm2)}=0. \tag{5}$$

In flat space, these correspond to plane waves times a local angular basis for the vectors and tensors [29].

Vector and tensor objects can of course be built out of scalar and vector normal modes through covariant differentiation and the metric tensor [2]:

$$Q_i^{(0)}=-k^{-1}\nabla_iQ^{(0)},$$
$$Q_{ij}^{(0)}=\left(k^{-2}\nabla_i\nabla_j-\frac{1}{3}\gamma_{ij}\right)Q^{(0)},$$
$$Q_{ij}^{(\pm1)}=-\frac{1}{2k}[\nabla_iQ_j^{(\pm1)}+\nabla_jQ_i^{(\pm1)}]. \tag{6}$$

The perturbations in the kth eigenmode can now be written as

$$A=\hat{A}\,Q^{(0)},\quad H_L=\hat{H}_LQ^{(0)},$$
$$\delta\rho=\widehat{\delta\rho}Q^{(0)},\quad \delta p=\widehat{\delta p}Q^{(0)}, \tag{7}$$

which possess only scalar components

$$B_i=\sum_{m=-1}^{1}\hat{B}^{(m)}Q_i^{(m)},$$
$$v_i=\sum_{m=-1}^{1}\hat{v}^{(m)}Q_i^{(m)}, \tag{8}$$

which possess scalar and vector components, and

$$H_{Tij}=\sum_{m=-2}^{2}\hat{H}_T^{(m)}Q_{ij}^{(m)},$$
$$\Pi_{ij}=\sum_{m=-2}^{2}\hat{\Pi}^{(m)}Q_{ij}^{(m)}, \tag{9}$$

which possess all three types. Here, scalar perturbations are denoted with a superscript (0), which is elsewhere omitted. We will hereafter also omit the caret in the normal mode amplitudes since real-space objects will no longer appear. Thus $\hat{v}^{(0)}\equiv v$.

C. Gauge covariant equations

The equations of motion for the matter follow from the Einstein equations $G_{\mu\nu}=8\pi G T_{\mu\nu}$. Furthermore, the Bianchi identities guarantee $T_{\mu\nu}=0$ which represents covariant energy and momentum conservation.

For the background, energy conservation implies

$$\dot{\rho}=-3\frac{\dot{a}}{a}(1+w)\rho, \tag{10}$$

where $w=p/\rho$ determines the background equation of state. As a result of the isotropy of the background, momentum conservation yields no additional constraint.

The Einstein equation determine the evolution of the scale factor through

$$\left(\frac{\dot{a}}{a}\right)^2 \equiv \frac{8\pi G}{3}a^2\rho_{\rm cr}=\frac{8\pi G}{3}a^2(\rho+\rho_S). \tag{11}$$

Here we have divided contributions to the expansion rate into the ordinary density and an effective density component that does not participate in gravitational collapse and is hence labeled ''S'' for smooth. The curvature provides the only component that is smooth by fiat,

$$\rho_S=-\frac{3}{8\pi G a^2}K, \tag{12}$$

with $w_S=-1/3$. Even the cosmological constant is kept smooth simply by dynamics. However, this notation is convenient for considering components that are approximately smooth. By keeping a general ρ_S and ρ here, we avoid lengthy rederivation of the equations of motion for such cases. Note that $\rho+\rho_S=\rho_{\rm cr}$, the so-called critical density, and we can define a critical equation of state w_{cr} by

$$\dot{\rho}_{\rm cr}=-3\frac{\dot{a}}{a}(1+w_{\rm cr})\rho_{\rm cr}. \tag{13}$$

Scalar matter perturbations obey the continuity and Euler equations

$$\left[\frac{d}{d\eta}+3\frac{\dot{a}}{a}\right]\delta\rho+3\frac{\dot{a}}{a}\delta p=-(\rho+p)(kv+3\dot{H}_L), \tag{14}$$

$$\left[\frac{d}{d\eta}+4\frac{\dot{a}}{a}\right]\left[(\rho+p)\frac{(v-B)}{k}\right]=\delta p-\frac{2}{3}\left(1-3\frac{K}{k^2}\right)p\Pi +(\rho+p)A, \tag{15}$$

and place 2 constraints on the 4 scalar matter-variables.

The metric and matter are related by the Einstein equations

$$(k^2-3K)\left[H_L+\frac{1}{3}H_T+\frac{\dot{a}}{a}\frac{1}{k^2}(kB-\dot{H}_T)\right] =4\pi G a^2\left[\delta\rho+3\frac{\dot{a}}{a}(\rho+p)(v-B)\Big/k\right], \tag{16}$$

$$k^2\left(A+H_L+\frac{1}{3}H_T\right)+\left(\frac{d}{d\eta}+2\frac{\dot{a}}{a}\right)(kB-\dot{H}_T)=8\pi G a^2 p\Pi, \tag{17}$$

$$\frac{\dot{a}}{a}A-\dot{H}_L-\frac{1}{3}\dot{H}_T-\frac{K}{k^2}(kB-\dot{H}_T)=4\pi G a^2(\rho+p)(v-B)/k, \tag{18}$$

$$\left[2\frac{\ddot{a}}{a}-2\left(\frac{\dot{a}}{a}\right)^2+\frac{\dot{a}}{a}\frac{d}{d\eta}-\frac{k^2}{3}\right]A-\left[\frac{d}{d\eta}+\frac{\dot{a}}{a}\right]\left(\dot{H}_L+\frac{1}{3}kB\right) =4\pi G a^2\left(\delta p+\frac{1}{3}\delta\rho\right). \tag{19}$$

Only two of these equations are functionally independent. The combination of these equations that corresponds to the conservation equation $G^{\mu\nu}{}_{;\nu}=0$ is automatically satisfied by any choice of the 4 metric variables due to the Bianchi identities. The remaining degrees of freedom are related to gauge freedom as we shall see.

Momentum conservation for vector perturbations gives the Euler equation

$$\left[\frac{d}{d\eta}+4\frac{\dot{a}}{a}\right][(\rho+p)(v^{(\pm1)}-B^{(\pm1)})/k] =-\frac{1}{2}(1-2K/k^2)p\Pi^{(\pm1)}, \tag{20}$$

and the Einstein equations give

$$(1-2K/k^2)(kB^{(\pm1)}-\dot{H}_T^{(\pm1)}) =16\pi G a^2(\rho+p)(v^{(\pm1)}-B^{(\pm1)})/k, \tag{21}$$

$$\left[\frac{d}{d\eta}+2\frac{\dot{a}}{a}\right](kB^{(\pm1)}-\dot{H}_T^{(\pm1)})=-8\pi G a^2 p\Pi^{(\pm1)}. \tag{22}$$

Again the Bianchi identity reduces the number of independent equations to 2.

For the tensor modes, the Einstein equations reduce to a single relation

$$\left[\frac{d^2}{d\eta^2}+2\frac{\dot{a}}{a}\frac{d}{d\eta}+(k^2+2K)\right]H_T^{(\pm2)}=8\pi G a^2 p\Pi^{(\pm2)}. \tag{23}$$

Neither the conservation equations nor the Bianchi identity say anything about tensor perturbations.

Although these relations are exact, they do not provide a closed system. There are in general 10 equations for 20 variables in the background and perturbations separately. For the

background, homogeneity and isotropy brings this to 2 equations for the 3 variables (a), (ρ,p), where the grouping distinguishes metric and matter categories. For the perturbations, the general relations are broken up into 4 equations for the 8 variables (A,B,H_L,H_T), $(\delta\rho,\delta p,(\rho+p)v,p\Pi)$ for the scalar perturbations, 2 equations for 4 variables $(B^{(\pm1)}, H_T^{(\pm1)})$, $(v^{(\pm1)},\Pi^{(\pm1)})$ for each set of vector perturbations and 1 equation for 2 variables $(H_T^{(\pm2)})$, $(\Pi^{(\pm2)})$ for each set of tensor perturbations. We can express the remaining $1+10$ degrees of freedom as the ability to choose the equation of state for the background $w=p/\rho$, the 6 stress fluctuations $(\delta p,p\Pi,p\Pi^{(\pm1)},p\Pi^{(\pm2)})$ for the perturbations, and the gauge [the 4 quantities $(\delta\eta,\ \delta x_i)$ for an arbitrary coordinate shift].

D. Multicomponent generalization

The conservation equations (14), (15) and (20) are valid for each species whose stress-energy tensor is independently covariantly conserved. For example, they apply to the photon-baryon system and the dark sector which only interact through gravity. The Einstein equations (16)–(19), (21) and (23) of course still hold with the appropriate summation over components, e.g. $\rho=\Sigma_J\rho_J$. Note that we do not include the smooth component ρ_S in the multicomponent sum.

IV. COORDINATE CHOICE

A. Gauge transformations

The additional four component freedom in the Einstein equations is fixed by a choice of coordinates that relate the perturbations to the underlying smooth background. The most general coordinate transformation associated with the kth normal mode is [1]

$$\tau=\tilde{\tau}+TQ^{(0)},$$

$$x_i=\tilde{x}_i+LQ_i^{(0)}+L^{(1)}Q_i^{(1)}+L^{(-1)}Q_i^{(-1)}, \tag{24}$$

where T corresponds to a choice in time slicing and $(L,\ L^{(1)},\ L^{(-1)})$ a choice of spatial coordinates. Under the condition that metric distances be invariant, they transform the metric as [2]

$$A=\tilde{A}-\dot{T}-\frac{\dot{a}}{a}T,$$

$$B=\tilde{B}+\dot{L}+kT,$$

$$H_L=\tilde{H}_L-\frac{k}{3}L-\frac{\dot{a}}{a}T,$$

$$H_T=\tilde{H}_T+kL \tag{25}$$

for the scalar perturbations and

$$B^{(\pm1)}=\tilde{B}^{(\pm1)}+\dot{L}^{(\pm1)},$$

$$H_T^{(\pm1)}=\tilde{H}_T^{(\pm1)}+kL^{(\pm1)} \tag{26}$$

for the vector perturbations.

Similarly, they transform the components of the stress-energy tensor as [2]

$$\delta\rho_J=\widetilde{\delta\rho}_J-\dot{\rho}_JT,$$

$$\delta p_J=\widetilde{\delta p}_J-\dot{p}_JT,$$

$$v_J=\tilde{v}_J+\dot{L} \tag{27}$$

for the scalar perturbations and

$$v_J^{(\pm1)}=\tilde{v}_J^{(\pm1)}+\dot{L}^{(\pm1)} \tag{28}$$

for the vector perturbations. All other quantities in the metric and matter are gauge invariant. In particular, the tensor modes do not exhibit gauge freedom since the transverse-traceless condition on $Q^{(\pm2)}$ is sufficient to remove the gauge ambiguity. The gauge is thus fixed by conditions on the metric which fully specify the transformation $(T,\ L,\ L^{(\pm1)})$ from an arbitrary frame.

It is important to bear in mind that both the metric fluctuations (A,B,H_L,H_T) and the matter fluctuations $(\delta\rho,\delta p,\ [\rho+p]v)$ take on different numerical values in different frames even in this covariant notation. For example, if ρ evolves in time, a density perturbation $\delta\rho$ arises simply from the warping of the time hypersurface on which the perturbation is defined. Thus, a density perturbation differs negligibly only between frames separated by [see Eq. (10)]

$$T\ll\left[(1+w)\frac{\dot{a}}{a}\right]^{-1}\frac{\delta\rho_J}{\rho_J}. \tag{29}$$

The common gauge choices of the next section all agree on the density perturbation in the clustered component well inside the horizon.

B. Gauge choice

Gauge freedom can be used to simplify the equations of motion. Most commonly, it is employed to convert certain Einstein equations to algebraic relations and/or eliminate relativistic effects from the conservation equations.

1. Vector gauges

Let us first dispose of the vector degrees of freedom. There are two natural choices [1]: $H_T^{(\pm1)}=0$ which fixes the gauge completely and $B^{(\pm1)}=0$ which leaves an arbitrary constant offset in $H_T^{(\pm1)}$. The latter does not produce a dynamical effect and can always be eliminated by specifying an initial condition for $H_T^{(\pm1)}$.

2. Comoving (scalar) gauge

It is useful to consider a scalar gauge where the metric and matter fluctuations are simply related [1]. Inspection of the Euler and Einstein equations shows us that the coordinate

choice $B=v$ simplifies the equations of motion greatly. This fixes the time slicing through $T=(\tilde{v}-\tilde{B})/k$. The additional condition $H_T=0$ specifies that $L=\tilde{H}_T/k$ and fixes the gauge completely. We call this the *comoving* gauge since here the momentum density vanishes. The remaining metric variables are labeled $A=\xi$ and $H_L=\zeta$. This choice reduces the Euler equation to the algebraic relation for the potential,

$$(\rho+p)\xi=-\delta p+\frac{2}{3}(1-3K/k^2)p\Pi, \tag{30}$$

which is also simply related to the curvature through the Einstein equation (18):

$$\dot{\zeta}=\frac{\dot{a}}{a}\xi+4\pi Ga^2(\rho_S+p_S)v/k. \tag{31}$$

Here we have again rewritten the curvature component as a smooth density contribution as in Eq. (11) for easy generalization to approximately smooth cases.

The simple relation between the metric and stress perturbation of Eqs. (30) and (31) is what makes this gauge useful. A smooth component complicates these relations because of the difference between a frame that is comoving with v versus the total-momentum-weighted velocity $v\dot{\rho}/(\dot{\rho}+\dot{\rho}_S)$. With a constant comoving curvature, the continuity equation (14) reduces to an ordinary conservation equation since metric changes to the fiducial volume are absent.

3. Newtonian (scalar) gauge

Finally, the *Newtonian* gauge is defined by $B=H_T=0$ and labels $A=\Psi$ and $H_L=\Phi$. The gauge is completely specified through $T=-\tilde{B}/k+\dot{\tilde{H}}_T/k^2$ and $L=-\tilde{H}_T/k$. The Einstein equations are reduced to algebraic relations that generalize the Poisson equation of Newtonian gravity:

$$(k^2-3K)\Phi=4\pi Ga^2\left[\delta\rho+3\frac{\dot{a}}{a}(\rho+p)v\Big/k\right],$$

$$k^2(\Phi+\Psi)=-8\pi Ga^2p\Pi. \tag{32}$$

These algebraic relations and the fact that CMB anisotropies are simply related to Φ and Ψ make this gauge useful.

4. Gauge-covariant variables

It is often useful to speak of the variables of say the comoving gauge while in a Newtonian representation. Bardeen [1] introduced a so-called ''gauge-invariant'' language that achieves this. We denote such techniques as gauge covariant since they amount to introducing covariant expressions for objects that take on the desired meaning only in a specific frame. The only objects that cannot be made gauge covariant are those that are ill defined due to coordinate ambiguities.

To avoid confusion, we only use gauge-covariant variables to describe metric fluctuations (ζ,ξ,Φ,Ψ). Matter perturbations will always be represented in the comoving gauge unless otherwise specified. Note that v is the same in comoving and Newtonian gauges.

The comoving curvature and density can be usefully expressed in Newtonian variables

$$\zeta=\Phi+2(\Psi-\Phi')\frac{\rho_{\rm cr}}{\rho'}, \tag{33}$$

$$4\pi Ga^2\delta\rho=(k^2-3K)\Phi, \tag{34}$$

obtained through Eqs. (25) and (18). Likewise Eq. (31) can be rewritten as

$$\zeta'-\xi=(\Psi-\Phi')\frac{\rho_S'}{\rho'}. \tag{35}$$

Recall that primes represent derivatives with respect to $\ln a$.

As we shall see, employing both comoving and Newtonian metric variables in the covariant language allows us to exploit the simple relations to comoving stresses in the former and comoving density perturbations in the latter.

V. STRESS PHENOMENOLOGY

A. Stress representation

We have seen that the stresses of the matter components completely determine the evolution of perturbations (see Figs. 3, 4, and 5). The background stress is completely determined by the equation of state w. The scalar stress fluctuations are determined by functional relations between the pressure or isotropic stress perturbations δp, anisotropic stress perturbation $p\Pi$ and the density perturbation $\delta\rho$. These relations may also involve hidden internal degrees of freedom. A model may also possess background stress without stress perturbations and vice versa. We call the former a smooth stress and the latter a seed stress. Vector and tensor perturbations likewise depend on the components $p\Pi^{(\pm1)}$ and $p\Pi^{(\pm2)}$ of the anisotropic stress tensor.

Scalar stress perturbations control the basic elements of the structure formation history, and so we pay particular attention to categorizing their properties. As discussed in Sec. IV B, the isotropic scalar stress poses a special problem in that its value depends on the choice of coordinate or gauge. The comoving gauge (where the momentum density vanishes) provides a useful choice of gauge because of the simple relation between the metric and stress fluctuations. We will use these coordinates to define the *total* scalar stress

$$S=-\xi=\frac{\delta p}{\rho+p}-\frac{2}{3}(1-3K/k^2)\frac{p}{\rho+p}\Pi. \tag{36}$$

It is useful to isolate gauge invariant aspects of the stress perturbation. The *anisotropic* scalar stress

$$S_\Pi=-8\pi Ga^2p\Pi/k^2 \tag{37}$$

is gauge invariant by definition; this form of the anisotropic stress also enters into the Einstein equations separately from S.

An adiabatic stress perturbation obeys

$$\delta p_A=(\rho+p)S_A=\frac{p'}{\rho'}\delta\rho. \tag{38}$$

Although δp_A is not gauge invariant, the adiabatic sound speed is

$$c_s^2\equiv\frac{\delta p_A}{\delta\rho}=\frac{p'}{\rho'}. \tag{39}$$

This gauge invariance implies that there are no coordinate ambiguities when discussing the pressure support of adiabatic fluctuations.

The remaining gauge-invariant pressure perturbation is

$$S_\Gamma=\frac{1}{\rho+p}\left(\delta p-\frac{p'}{\rho'}\delta\rho\right). \tag{40}$$

Unlike adiabatic stresses, these may be present even when the comoving density perturbation $\delta\rho$ is negligible [see Eq. (34)]. Entropic stresses are the primary means of structure formation in most isocurvature models.

The comoving-gauge analogues of S_A and S_Γ are also useful. As long as $\delta p/\delta\rho$ is less than unity, superhorizon stresses are negligible compared with curvature fluctuations. We call the part of $\delta p/\delta\rho$ that is separable in time and space the *comoving* sound speed $c_C^2=f(k)g(\eta)$ and the accompanying stress *sonic*:

$$S_S=c_C^2\frac{\delta\rho}{\rho+p}. \tag{41}$$

Separability is not a gauge-invariant property, but this is not in itself a problem because the comoving frame is dynamically special.

We call the remaining isotropic stress the entropic stress

$$S_E=\frac{1}{\rho+p}(\delta p-c_C^2\delta\rho), \tag{42}$$

such that the total stress is

$$S=S_S+S_E+\frac{2}{3}\frac{k^2-3K}{8\pi Ga^2(\rho+p)}S_\Pi. \tag{43}$$

Note that if the comoving sound speed equals the adiabatic sound speed $c_C^2=c_s^2=p'/\rho'$, then $S_S=S_A$ and $S_E=S_\Gamma$.

B. Seed stress

Seed stresses provide a special case with unique properties. The effects of seed perturbations are in fact simpler to understand than those of the fluid type because the problem decouples completely. If the seeds do not interact directly with other types of matter, the conservation equations (14) and (15) imply that the metric perturbations only affect the seed perturbations at second order since bare ρ_s and p_s terms may be dropped. They become [30]

$$\left[\frac{d}{d\eta}+3\frac{\dot a}{a}\right]\delta\rho_s=-k(\rho_s+p_s)v_s-3\frac{\dot a}{a}\delta p_s, \tag{44}$$

$$\left[\frac{d}{d\eta}+4\frac{\dot a}{a}\right](\rho_s+p_s)\frac{v_s}{k}=\delta p_s-\frac{2}{3}(1-3K/k^2)p_s\Pi_s. \tag{45}$$

The basic principles of how stress fluctuations affect the gravitational potential still hold but here there is a stress contribution that is truly external to the system of metric fluctuations. It is again possible to have large-scale entropic stress perturbation in the absence of initial curvature or density perturbation.

The formal solutions to Eqs. (44) and (45) are

$$(\rho_s+p_s)v_s/k=a^{-4}\int d\eta a^4\left[\delta p_s-\frac{2}{3}(1-3K/k^2)p_s\Pi_s\right],$$

$$\delta\rho_s=-a^{-3}\int d\eta a^3\left[k(\rho_s+p_s)v_s+3\frac{\dot a}{a}\delta p_s\right]. \tag{46}$$

The task of understanding a seed model like defects reduces to understanding its stresses, but this is a formidable task in realistic seed models such as cosmological defects (e.g., [23]).

C. From stresses to curvature

These stresses are the fundamental sources and sinks of the metric perturbations. We therefore seek to express the comoving and Newtonian curvatures in terms of the stresses. Combining Eqs. (33), (35), and (36) with

$$\Psi=-\Phi+S_\Pi, \tag{47}$$

yields

$$\zeta'=-S+[\Phi'+\Phi-S_\Pi]\left(\frac{\rho_{\rm cr}'}{\rho'}-1\right), \tag{48}$$

and

$$\frac{\sqrt{\rho}}{a}\left[\frac{a}{\sqrt{\rho}}\Phi\right]'=-\frac{1}{2}\frac{\rho'}{\rho}\zeta+S_\Pi. \tag{49}$$

In a non-critical (non-flat, $\rho\neq\rho_{\rm cr}$) universe, we combine these to eliminate ζ from the evolution equation for the Newtonian curvature, giving

$$\Phi''+\left(1-\frac{\rho''}{\rho'}+\frac{1}{2}\frac{\rho_{\rm cr}'}{\rho_{\rm cr}}\right)\Phi'+\left(\frac{1}{2}\frac{\rho_{\rm cr}'+\rho'}{\rho_{\rm cr}}-\frac{\rho''}{\rho'}\right)\Phi$$
$$=\frac{1}{2}\frac{\rho'}{\rho_{\rm cr}}S+S_\Pi'+\left(\frac{1}{2}\frac{\rho_{\rm cr}'+\rho'}{\rho_{\rm cr}}-\frac{\rho''}{\rho'}\right)S_\Pi. \tag{50}$$

Recall that the Newtonian curvature is simply related to the comoving density perturbation through Eq. (34).

In a critical-density universe ($\rho=\rho_{\rm cr}$), Eqs. (48) and (49) can be formally solved as integrals over the stress fluctuations to yield

$$\zeta(a)=\zeta(0)-\int\frac{da}{a}S \tag{51}$$

and

$$\Phi=\zeta-\frac{\sqrt{\rho}}{a}\int\frac{da}{\sqrt{\rho}}[\zeta-S-S_{\rm II}]+C\frac{\sqrt{\rho}}{a}. \tag{52}$$

The last term is the decaying mode of Φ where $C=$const.

There are three general conclusions that we can draw from Eqs. (51) and (52). The first is that in the absence of an initial comoving curvature perturbation [$\zeta(0)=0$], a stress fluctuation will generate one of order $\zeta\sim -S$. The same goes for the Newtonian curvature $\Phi\sim -S$. The reason for this behavior is that a stress gradient $k\delta p$ generates a potential flow with $(\rho+p)v\sim(k\eta)\delta p$ which generates a density perturbation of $\delta\rho\sim-(k\eta)^2\delta p$ and hence a curvature perturbation of $\Phi\sim-\delta p/\rho_{\rm cr}$. Note that this intuitive argument fails for other gauge choices.

Second, starting with a curvature perturbation and assuming sonic stresses $\delta p=c_C^2\delta\rho$, it is clear that the same mechanism of generating flows will generate an opposing curvature fluctuation $\Delta\Phi\sim-c_C^2\delta\rho/\rho\sim-(c_Ck\eta)^2\Phi$ that will destroy the initial curvature fluctuation when $c_Ck\eta\sim1$. In physical terms this occurs because pressure support prevents perturbations from collapsing and hence causes the curvature perturbation to redshift away.

Finally, the anisotropic stress contributes to the total stress S and thus can both create and destroy comoving curvature fluctuations. Furthermore, it enters separately into the Newtonian curvature through Eq. (52). This is because the Newtonian frame, unlike the comoving frame, is defined to be globally shear free ($B=H_T=0$). The coordinate transformation that maps the comoving frame to the shear free frame depends on the anisotropic stress and hence aliases background evolution into contributions to the Newtonian curvature.

Unfortunately, despite their general appearance, Eqs. (51) and (52) are only formal solutions since the time evolution of the stress sources generally depends on the curvature fluctuation itself. We will use the special properties of smooth, anisotropic, entropic, and sonic stresses to address this problem in Secs. VI, VII, and VIII.

D. From curvature to observables

As discussed in Sec. II D, the Newtonian curvature is directly related to observables in the CMB and large scale structure. We are now in a position to quantify these relations. The Newtonian curvature Φ and potential Ψ encapsulate all observable properties of scalar fluctuations. The contribution of a given k-mode to the amplitude of the lth multipole moment of the CMB anisotropy is given by [10]

$$\frac{\Theta_l}{2l+1}\approx\int_0^{\eta_0}d\eta e^{-\tau}\{[\dot\Psi-\dot\Phi+\dot\tau(\Theta_0+\Psi)]\times j_l[k(\eta_0-\eta)]+\dot\tau v_b j_l'[k(\eta_0-\eta)]\}, \tag{53}$$

where we have dropped the small correction due to the polarization of the CMB. τ is the optical depth to Compton scattering between the present (η_0) and the epoch in question (η); Θ_0 is the photon temperature perturbation in Newtonian gauge, and v_b is the baryon velocity in Newtonian or comoving gauge. For an open universe, the spherical Bessel function j_l is replaced by the hyperspherical Bessel function. Note that the prime here and here only refers to a derivative with respect to the argument of the Bessel function. With the random phase assumption for the k-modes, the scalar contribution to power spectrum of the anisotropies is

$$C_l=\frac{2}{\pi}\int\frac{dk}{k}k^3\frac{\langle\Theta_l^*\Theta_l\rangle}{(2l+1)^2}. \tag{54}$$

In Eq. (53), the $\dot\Psi-\dot\Phi$ term leads to the so-called integrated Sachs-Wolfe (ISW) effect and contributes once the optical depth to scattering becomes small, i.e. after last scattering. The other terms, $\Theta_0+\Psi$ and v_b, are localized to the last scattering surface itself and represent the effective temperature of the distribution and the Doppler effect respectively. They are responsible for the so-called acoustic peaks in the CMB spectrum, the morphology of which can be directly read off of the metric driving terms [10].

The influence of time variability in Ψ and Φ depends on how the variation rate $1/\Delta\eta$ compares with the perturbation crossing time for sound (before last scattering) and light (after last scattering). For variations on a much shorter time scale ($k\Delta\eta\ll1$ or $kc_s\Delta\eta\ll1$), only the total change $\Delta(\Psi-\Phi)$ is observable since all photons suffer a uniform gravitational redshift. For variations on a much longer time scale the effects mainly cancel out as the photons either traverse many wavelengths of the fluctuation during the variation with redshifts and blueshifts cancelling or equivalently undergo many acoustic oscillations. Changes whose duration are synchronized with the oscillation period are the most effective. These general considerations apply to metric variations of the vector and tensor type as well. Two commonly encountered examples are the cancellation cutoff at high l for a uniform potential decay and the driving of acoustic oscillations from synchronized potential decay in the radiation dominated epoch.

A constant potential also leads to observable effects through the effective temperature term $\Theta_0+\Psi$. Since Θ_0 is the temperature perturbation in Newtonian gauge, it can be obtained by a gauge transformation from the comoving gauge by noting that in that gauge both the density perturbation and the potential ξ are negligible because stress perturbations must be negligible for the potential to remain constant. Equations (25) and (27) then imply

$$\Theta_0=\frac{\delta\rho_\gamma}{4\rho_\gamma}-\frac{\dot a}{a}v\Big/k,$$

$$\Psi = \dot{v}/k + \frac{\dot{a}}{a} v \Big/ k. \tag{55}$$

For adiabatic fluctuations, $\delta\rho_\gamma/\rho_\gamma \propto \delta\rho/\rho$ and is hence negligible outside the horizon by virtue of the Poisson equation (34). Since Ψ is a constant by assumption, these equations can be integrated to give

$$\Theta_0 = -\frac{2}{3(1+w)}\Psi, \tag{56}$$

and hence $\Theta_0 + \Psi = \Psi(1+3w)/(3w+3)$. This relation ultimately comes from the fact that Ψ represents a time shift and $a \propto t^{2/3(1+w)}$ [31]. For $w=0$, this reduces to the well-known result that the effective temperature is $\Psi/3$ in the adiabatic sCDM model [9]. Compared with this model, those that are dominated by the ISW term $\Delta(\Psi-\Phi) \approx 2\Psi - S_\Pi$ such as traditional isocurvature and smooth models potentially have up to 6 times the anisotropy for a given potential fluctuation.

The behavior of the density perturbations that underlies the large-scale structure is even more directly related to the Newtonian curvature perturbation. Inside the horizon, all reasonable choices of gauge agree on the density perturbation. In particular, the comoving gauge density perturbation is algebraically related to Φ by Eq. (34) and hence allows a simple translation of results for one to the other. Consequently, the relation between temperature and potential fluctuations discussed in the last paragraph translates directly into a relation between density perturbations and CMB anisotropies.

With this relation, the so-called transfer function of the density perturbations below the current horizon can be simply read off of the time history in the potential without the usual gauge ambiguity in defining the initial density perturbation,

$$T(k) = \frac{\Phi(\eta_0,k)}{\Phi(0,k)} \frac{\Phi(0,0)}{\Phi(\eta_0,0)}. \tag{57}$$

The power spectrum of density perturbations today is then proportional to $T(k)^2 P_{\rm initial}$. Any process that makes the potential decay relative to the $k=0$ mode will produce a downturn in $T(k)$ and a reduction of small-scale power.

VI. STRESS-FREE PERTURBATIONS

Adiabatic models for structure formation, those whose initial conditions contain true comoving curvature perturbations, all go through a period in which the scalar stress perturbation may be neglected. This is a direct consequence of causality. Stress gradients affect structure formation simply by the causal motion of matter. On scales larger than the horizon, the change in the comoving curvature perturbations due to causal motion can always be neglected.

We begin in Sec. VI A with adiabatic models that have no smooth components. Although the comoving curvature remains at the value set by the initial condition, the Newtonian curvature depends on the equation of state. We consider then in Sec. VI B the effect of a smooth component on the Newtonian curvature (and hence the density perturbation). For each case, we begin with a general description of the resultant phenomenology. We then illustrate the phenomenology with full solutions of the perturbation equations and discuss applications within the current generation of structure formation models of Sec. II E. We will follow this pattern throughout the paper. An overview of the results is given in Fig. 6.

A. Clustered case

We begin with the case in which there is no smooth component ($\rho_S = 0$) and stress that fluctuations are negligible compared with metric fluctuations ($S \ll \zeta$). Here, Eq. (48) simply implies

$$\zeta = \text{const.} \tag{58}$$

Equation (52) then gives

$$\Phi = \zeta\left(1 - \frac{\sqrt{\rho}}{a}\int \frac{da}{\sqrt{\rho}}\right) + \frac{\sqrt{\rho}}{a}\int \frac{da}{\sqrt{\rho}} S_\Pi + C\frac{\sqrt{\rho}}{a}. \tag{59}$$

We will focus only on the behavior of the first term. The second term may be neglected if $S_\Pi \ll \zeta$. However, unlike $S \ll \zeta$, $S_\Pi \ll \zeta$ is not a consequence of causality and is mildly violated in the case of free radiation [10] and can be strongly violated in defect models [32]. We return to consider its effects in Sec. VII A. The last term is a decaying mode that carries no comoving curvature ($\zeta = 0$). For a constant equation of state w, it scales as

$$\Phi \propto a^{-3(1+w)/2-1}. \tag{60}$$

It is apparent from the form of the first term that since ρ cannot grow with time ($w \geqslant -1$), the first integral in Eq. (59) goes to a constant between 0 and 1 with the two extremes representing $w \to \infty$ and $w = -1$ respectively. The integral is dominated by the most recent epoch; only the equation of state at the epoch in question matters. It is simple to show that during periods where w is approximately constant [33],

$$\frac{\Phi}{\zeta} \to \frac{3+3w}{5+3w}. \tag{61}$$

At each epoch where w decreases from w_1 to w_2, the Newtonian curvature decreases by

$$\frac{\Phi(w_1) - \Phi(w_2)}{\Phi(w_1)} = 2\frac{(w_1 - w_2)}{(1+w_1)(5+3w_2)}. \tag{62}$$

These results are easy to understand in the non-relativistic limit. In the absence of stresses, the gradients in the gravitational potential set up flows as $v \sim (k\eta)\Phi$. The divergence of this flow generates density perturbations $\delta \sim (1+w)(k\eta)^2\Phi$ for constant w. By the Poisson equation $\Phi \sim (k\eta)^{-2}\delta$, this is exactly the rate of growth needed to keep the potential constant. Here we have used the fact that $4\pi G a^2 \rho = 3(\dot{a}/a)^2/2 \sim \eta^{-2}$ if there are no smooth components. When w decreases from w_1 to w_2, the change affects

the proportionality constants but not the scaling between δ and Φ. This leads to a decrease in the Newtonian potential to a new constant.

Full solutions. If the matter is predominantly composed of two components with different but constant equations of state w_1 and w_2, we can solve Eq. (59) exactly provided $S_\Pi \ll \zeta$. The result is

$$\frac{\Phi}{\zeta} = 1 - \frac{2}{5+3w_1} F\left[1, \frac{1}{2}; 1+n; \frac{y}{1+y}\right] = 1 - \frac{1}{3}\frac{1}{w_1 - w_2}(1+y)^{1/2} y^{-n} B_{y/(1+y)}\left(n, \frac{1}{2} - n\right), \tag{63}$$

where $n \equiv (5+3w_1)/(6w_1 - 6w_2)$ and

$$y = \rho_2/\rho_1 \propto a^{3(w_1 - w_2)}. \tag{64}$$

Here, $B_x(p,q)$ is the incomplete beta function and $F(a,b;c;x)$ is Gauss's hypergeometric function (i.e., ${}_2F_1$). The combination $y/(1+y)$ often enters into such solutions as it is just $\rho_2/\rho_{\rm cr}$, the fractional density perturbation supplied by component 2 as a function of time.

Any case with $2n$ equal to an integer can be expressed in elementary form. In particular, if $n = N$ or $N + 1/2$ for an integer $N \geqslant 0$, then

$$F\left(1, \frac{1}{2}; n+1; x\right) = \frac{1}{x}\sum_{k=0}^{N-1} \frac{\Gamma(n+1)\Gamma(n-k-1/2)}{\Gamma(n-k)\Gamma(n+1/2)}\left(\frac{x-1}{x}\right)^k + \left(\frac{x-1}{x}\right)^N \frac{\Gamma(n+1)\Gamma(n-N+1/2)}{\Gamma(n-N+1)\Gamma(n+1/2)} f_n, \tag{65}$$

where

$$f_n = \begin{cases} 1/\sqrt{1-x}, & n = N, \\ \dfrac{1}{2\sqrt{x}} \log \dfrac{1+\sqrt{x}}{1-\sqrt{x}}, & n = N + \dfrac{1}{2}. \end{cases} \tag{66}$$

Applications. These solutions apply to any adiabatic model with a matter-dominated epoch and may be used to explore the behavior of perturbations entering and exiting the matter-dominated epoch. For example, the matter-radiation case reduces to [34]

$$\frac{\Phi}{\zeta} = \frac{3}{5} + \frac{2}{15y} - \frac{8}{15y^2} - \frac{16}{15y^3} + \frac{16\sqrt{1+y}}{15y^3}. \tag{67}$$

This solution only holds in the absence of anisotropic stress perturbations and does not strictly apply to the usual radiation components. The neutrinos carry anisotropic stress [10] as do the photons after recombination. We consider how such effects can be taken into account in Sec. VII A.

For adiabatic models where there exists a component with $w_2 < 0$ [14,15,17], these solutions describe the exit from the matter-dominated epoch. A comparison of the analytic solution with full numerical solutions in those cosmologies is given in Fig. 7. A special case is $w_1 = 0$ and $w_2 = -1$, the matter to cosmological constant transition, where [35,36]

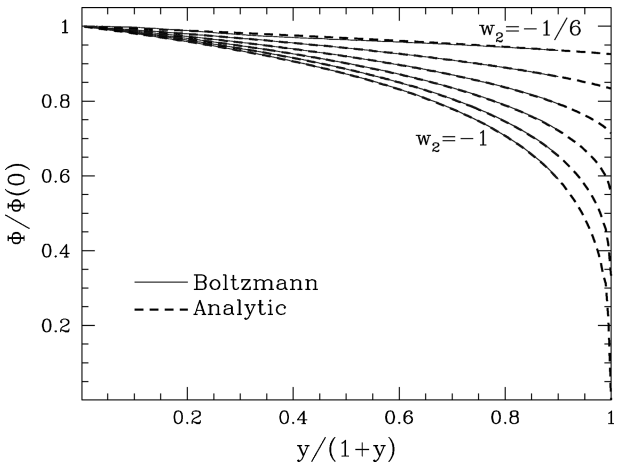

FIG. 7. Stress-free clustered case. In the case where all contributions to the expansion rate also cluster, the comoving curvature is constant but the Newtonian curvature changes with the equation of state w. For the case of the transition between $w_1 = 0$ and w_2 between -1 and $-1/6$, we show here the analytic solution of Eq. (63) compared with a full numerical Boltzmann solution of a QCDM model (including radiation) for a mode that is outside the horizon at the given epoch. Notice that only for $w_2 = -1$ (cosmological constant) does the potential decay to zero as $\rho_2/\rho_{\rm cr} = y/(1+y) \to 1$.

$$\frac{\Phi}{\zeta} = \frac{3}{5}\left[1 - \frac{1}{3}(1+y)^{1/2} y^{-5/6} B_{y/(1+y)}(5/6, -1/3)\right]. \tag{68}$$

As is evident from Eq. (61), the cosmological constant is the only case where the Newtonian curvature decays to zero. Note in particular that $w_2 = -1/3$ does not correspond to the behavior of spatial curvature in spite of the fact that the effect on the expansion rate is the same. The presence of fluctuations in the w_2 component prevents the gravitational potential from decaying to zero. Across the $w_1 = 0$ (matter-dominated) to $w_2 = -1/3$ transition, the Newtonian curvature goes from $3\zeta/5$ to $\zeta/2$. Thus, the string-dominated (strCDM) model which has $w_{\rm str} = -1/3$ does not behave like an open model on the large scales relevant for the CMB (cf. [16])

B. Smooth components

We next consider how smooth components affect the growth of structure. Although curvature is the only component that is smooth by definition and a cosmological constant the only one that is smooth by dynamics, under certain circumstances other components can be approximately smooth.

We define a component (S) to be smooth if its density fluctuations are small in comparison to those of the clustering components (C), i.e. $\delta\rho_S \ll \delta\rho_C$, regardless of whether $\rho_S < \rho_C$. To be maintained dynamically, the respective energy fluxes must also satisfy $(\rho_S + p_S)v_S \ll (\rho_C + p_C)v_C$. In this section, we add the subscript (C) on the remaining matter to remind the reader that part of the total matter density

has been designated as effectively smooth in the division of Eq. (11). Our analysis also applies to cases where $\delta\overline{\rho_S} \ll \delta\rho_C$ where the time-average is over the dynamical time of the C component (see Sec. VIII C).

One cannot demand that $\delta\rho_S$ be identically zero since the continuity equation (14) generates a density fluctuation as the metric curvature changes unless $w_S \equiv p_S/\rho_S = -1$ (a cosmological constant). The fractional density fluctuation is generically at least of order the curvature fluctuation Φ. Since this term exceeds the energy flux term outside the horizon due to causality, energy conservation forbids smooth components (with $w_S \neq 1$) on these scales.

Components that are smooth within the horizon are possible. For example, the C component may be driven to collapse by potential gradients while the S component is supported against collapse by stress gradients (see Sec. VIII C). Here $\delta\rho_S/\rho_S \sim \Phi \ll \delta\rho_C/\rho_C$. Since these stress gradients are set up exactly so as to keep the component smooth, one can replace such stress effects in Eq. (50) with an additional smooth density component [37]. The remaining perturbations can then be approximated as stress-free and generate curvature fluctuations as

$$\Phi'' + \left(1 - \frac{\rho_C''}{\rho_C'} + \frac{1}{2}\frac{\rho_{\rm cr}'}{\rho_{\rm cr}}\right)\Phi' + \left(\frac{1}{2}\frac{\rho_{\rm cr}' + \rho_C'}{\rho_{\rm cr}} - \frac{\rho_C''}{\rho_C'}\right)\Phi = 0. \tag{69}$$

Now even for constant w_C, $\Phi' = 0$ no longer solves the equation of motion. Mathematically, ρ_S adds to the expansion drag (Φ') terms but not the gravitational (Φ) terms. Physically, potential flows still create density perturbations as $\rho_C \sim (\rho_C + p_C)(k\eta)^2\Phi$ but the Poisson equation leads to a smaller potential $\Phi = (k\eta)^{-2}\delta\rho_C/(\rho_C + \rho_S)$. As this process continues, Φ decays away.

Equation (69) has simple solutions in the limit that $\rho_{\rm cr}' \gg \rho_C'$, as is usually the case when the smooth component dominates the expansion. In this case, the general solution to the equation is

$$\Phi = C_1 a^{-1} + C_2 a^{-1}\int d\ln a \frac{a\rho_C'}{\rho_{\rm cr}^{1/2}}. \tag{70}$$

Both terms represent decaying modes as long as $w_{\rm cr} - 2w_C < 1$.

Full solutions. The full solution to Eq. (69) can be obtained analytically for a clustering component with a constant equation of state and present-day fractional density contribution pair (w_1, Ω_{C1}), a smooth component with (w_1, Ω_{S1}), and/or a second smooth component with $(w2, \Omega_{S2})$:

$$\Phi_j = y^{1/3\Delta w}\left(\frac{y}{1+y}\right)^{\alpha_j} \times F\left(\alpha_j, \alpha_j + \frac{1}{2}; 2\alpha_j + 1 - \frac{3w_1+1}{6\Delta w}; \frac{y}{1+y}\right), \tag{71}$$

where $\Delta w = w_2 - w_1$ with

$$\alpha_j \equiv \frac{1+3w_1}{12\Delta w}\left[1 \mp \sqrt{1 + 24\frac{\Omega_{C1}}{\Omega_{C1}+\Omega_{S1}}\frac{1+w_1}{(1+3w_1)^2}}\right],$$

$$y = \rho_2/\rho_1 = \frac{\Omega_{S2}}{\Omega_{S1}+\Omega_{C1}}a^{-3\Delta w}, \tag{72}$$

where $j = 1,2$ for the growing and decaying modes corresponding to $-,+$ in the α_j equation. Note that negative contributions to the critical density, e.g. positive spatial curvature $K > 0$, are also covered by these solutions.

The hypergeometric solution allows one to identify elementary solutions more easily. In particular, cases where $\alpha_j = -|N_1/2|$, $(3w_1+1)/6\Delta w = N_1$, or $2\alpha_j - (3w_1+1)/6\Delta w + 1/2 = N_1$ can be expressed in terms of elementary functions. Cases of the form $F(a,b;b+N_1;x)$, $F(a,b;a+N_1;x)$, or $F(N_1/2,N_2/2;N_3/2;x)$ can be expressed in terms of elementary functions, incomplete beta functions, and/or complete elliptic integrals. Here, the N_j are integers and a and b are real numbers. For example, the case of $w_1 = 0$ and $w_2 = -1/6$ can be expressed in closed form for any $\Omega_{C1}/(\Omega_{C1}+\Omega_{S1})$:

$$\Phi_j \propto \frac{y^{\alpha_j}}{a}\frac{2\alpha_j + 1 + \sqrt{1+y}}{(1+\sqrt{1+y})^{2\alpha j+1}}. \tag{73}$$

For $w_1 = 0$, any case in which $2\alpha_j + 1/6w_2 + 1/2$ is an integer can be simplified. One can also simplify the growing mode of cases with $w_1 = \Omega_{S1} = 0$ and w_2^{-1} equal to an integer.

Finally for completeness, there is a well-known special case of Eq. (69) that is not completely covered by Eq. (71) but that does have an integral solution. This involves a clustering component $w = 0$ (CDM) and a smooth component composed of an arbitrary admixture of $w = -1/3$ (curvature) and $w = -1$ (Λ) pieces [38]:

$$\Phi \propto a^{-1}\rho_{\rm cr}^{1/2}\int \frac{d\ln a}{a^2\rho_{\rm cr}^{3/2}}, \quad \Phi \propto a^{-1}\rho_{\rm cr}^{1/2}, \tag{74}$$

for the growing and decaying modes respectively. Of course, if either the curvature or Λ component is negligible, the solutions have an analytic form described by Eq. (71). Unfortunately, this solution cannot be generalized to arbitrary combinations of smooth components since it relies on the fact that $w = -1/3$ does not accelerate the expansion and that $w = -1$ gives a constant density contribution.

Applications. Smooth components are widely found in adiabatic models and their description in terms of Eq. (71) are given in Table I.

In fact, all CDM variants (CDMv) go through a phase in the radiation-dominated epoch when the perturbations in the radiation are pressure supported leading to an essentially smooth component $\Omega_{\rm rad} = \Omega_{S2}$ with $w_2 = 1/3$ and also a smooth baryonic component $\Omega_b = \Omega_{S1}$ with $w_1 = 0$ that is held against collapse by Compton coupling to the photons (see Sec. VIII C and [39]). In this case, the clustered matter is the CDM ($\Omega_{\rm CDM} = \Omega_{C1}$), and the solution (71) describes the evolution of the CDM density perturbations via the Pois-

TABLE I. Correspondence between the analytic solution of Eq. (71) and models with smooth components. Ω_K is the fractional effective density supplied by the curvature component.

w_1	w_2	Ω_C	Ω_{S1}	Ω_{S2}	Model
0	1/3	$\Omega_{\rm cdm}$	$\Omega_{\rm b}$	$\Omega_{\rm rad}$	CDMv
0	–	$\Omega_{\rm cdm+b}$	Ω_ν	–	HCDM
0	-1	$\Omega_{\rm cdm+b}$	0	Ω_Λ	ΛCDM
0	-1	$\Omega_{\rm cdm+b}$	Ω_ν	Ω_Λ	ΛHCDM
0	$-1/3$	$\Omega_{\rm cdm+b}$	0	Ω_K	OCDM
0	$-1/3$	$\Omega_{\rm cdm+b}$	Ω_ν	Ω_K	OHCDM
0	$-1/3$	$\Omega_{\rm cdm+b}$	0	$\Omega_{\rm str}$	strCDM
0	–	$\Omega_{\rm cdm+b}$	Ω_ϕ	–	ϕCDM
0	w_Q	$\Omega_{\rm cdm+b}$	0	Ω_Q	QCDM, GDM
0	w_Q	$\Omega_{\rm cdm+b}$	Ω_ν	Ω_Q	QHCDM, GDM
0	w_Q	$\Omega_{\rm cdm+b}$	Ω_ϕ	Ω_Q	ϕQCDM, GDM

son equation. The behavior deep in the radiation domain is given by Eq. (70), which says that density perturbations actually grow logarithmically. The full solution (71) maps this logarithmic mode into a power-law growth across the matter-radiation transition and is useful for determining the amplitude of small-scale fluctuations as a function of the baryon content [39].

Smooth components that dominate at late times are also described by Eq. (71). Many such models have been proposed to reduce the amount of small-scale power in the standard CDM model. The prototypical examples are the ΛCDM model where $\Omega_{S2}=\Omega_\Lambda=1-\Omega_m$ ($w_2=-1$) and the open model OCDM $\Omega_{S2}=\Omega_K=1-\Omega_m$ ($w_2=-1/3$). The former case is special in that it may alternately be considered a clustered component (see Sec. VI A). The latter case must be considered as a smooth component, and the growing mode reduces to [40]

$$\Phi \propto \frac{1}{y}\left(1+\frac{3}{y}-\frac{3}{y}\sqrt{\frac{1+y}{y}}\tanh^{-1}\sqrt{\frac{y}{1+y}}\right), \qquad (75)$$

where $y \propto a$.

The obvious generalization of such models involves a smooth component with an equation of state that differs from the curvature or cosmological constant examples. The prototypical case is the HCDM model where a massive neutrino (or hot dark matter) component with $w_1 \approx 0$ remains smooth on small scales due to residual relativistic effects. Our general solution in fact allows an additional smooth component w_2, which could be curvature (OHCDM), a cosmological constant (ΛHCDM), or even quintessence with an equation of state $w_2=w_Q$ (QHCDM). In fact, in the case where there is no hot component, the latter solution describes the ''quintessence'' model (QCDM) that has recently received much attention [14,15]. Here, a scalar field supplies a density component that is smooth inside a sound horizon that corresponds to the particle horizon. A comparison of the analytic solution to numerical results for QCDM and ΛHCDM is given in Figs. 8 and 9. Likewise, the solution applies to the GDM generalization [17] where the sound horizon is allowed to be arbitrary.

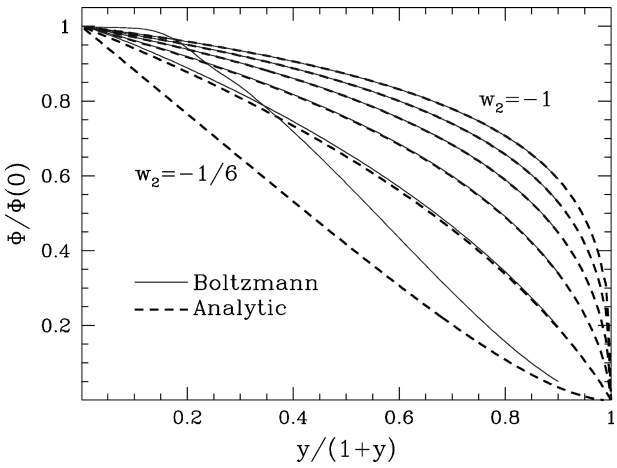

FIG. 8. Single smooth component. The same as Fig. 7 except that w_2 is taken to be a smooth component. Here Eq. (71) is compared to a QCDM model but for a mode that is well inside the horizon at $y=1$. Discrepancies at early times are due to radiation contributions in the QCDM models and are particularly pronounced for $w_2 \to 0$ since then the epoch when the Q and matter components are equal is driven into the radiation dominated era for this flat $\Omega_m=0.35$ model.

Finally, Eq. (74) describes the case of a CDM model with both curvature and cosmological constant contributions.

C. Vector perturbations

The behavior of vector modes is far simpler than that of scalar modes. Vector anisotropic stress can be neglected if $H_T^{(\pm 1)} \gg p\Pi^{(\pm 1)}/\rho$ in a gauge where $B^{(\pm 1)}=0$. In the absence of stress perturbations, vector modes simply decay with the expansion [see Eq. (20)] is solved by

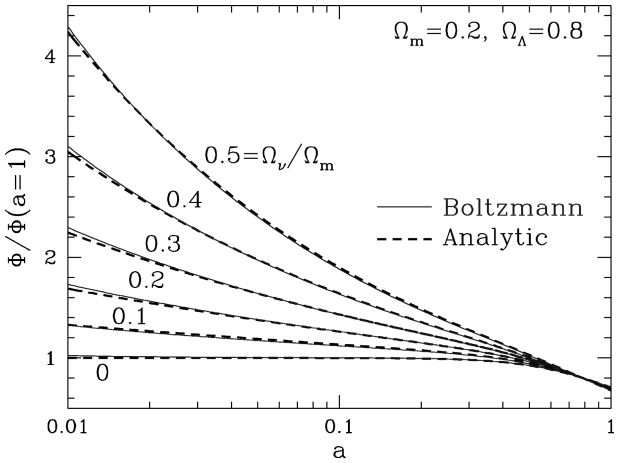

FIG. 9. Two smooth components. Equation (71) is compared with numerical solutions in a ΛHCDM which has smooth Λ and hot (H) dark matter Ω_ν components. Numerical solutions are for modes well below the Jeans scale of the hot dark matter and discrepancies at early times are due to radiation contributions in the model.

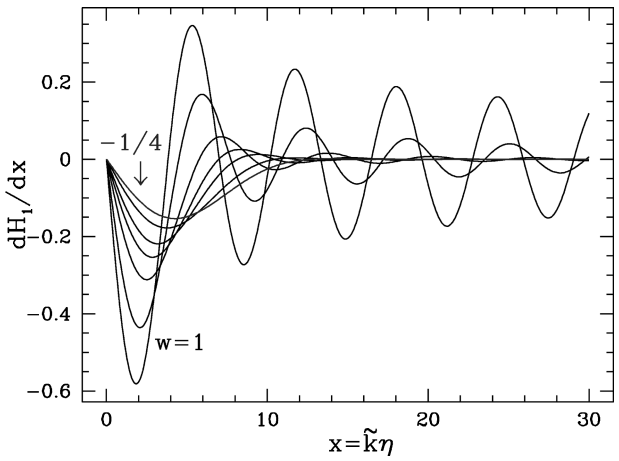

FIG. 10. Gravity wave modes. The amplitude of free gravity waves remains constant outside the horizon and oscillates and decays inside in a manner dependent on the equation of state of the background. Shown here is the derivative of the mode since it acts as the source of radiation anisotropies. Note that as w increases the oscillatory phase begins sooner relative to horizon crossing.

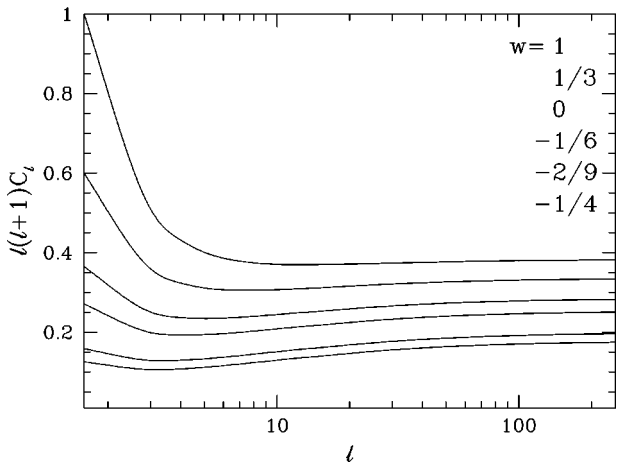

FIG. 11. Gravity wave anisotropies generated in free radiation. Numerical solutions are from a Boltzmann code with scattering sources removed. Notice that as w increases so do the anisotropies. This is directly related to the behavior of the gravity wave modes in Fig. 10.

$$(\rho+p)(v^{(\pm1)}-B^{(\pm1)})\propto a^{-4},$$

$$kB^{(\pm1)}-\dot{H}_T^{(\pm1)}\propto a^{-2}. \tag{76}$$

The metric source is related algebraically to this quantity by Eq. (21).

D. Tensor perturbations

The tensor anisotropic stress is negligible if $H_T^{(\pm2)} \gg p\Pi^{(\pm2)}/\rho_{\rm cr}$. In this limit, tensor metric fluctuations remain constant outside the horizon regardless of the expansion rate and propagate as free gravity waves inside of it. They are described by Eq. (23) with the sources set to zero.

Full solutions. If the expansion is dominated by a component with constant equation of state $w_{\rm cr}>-1/3$, the fundamental modes of gravity waves are

$$H_1=\frac{2^{m+1}\Gamma(m+3/2)}{\sqrt{\pi}}x^{-m}j_m(x),$$

$$H_2=\frac{2^{m+1}\Gamma(m+3/2)}{\sqrt{\pi}}x^{-m}n_m(x), \tag{77}$$

with $x=k\eta$, $m=(1-3w)/(1+3w)$ and $\tilde{k}=\sqrt{k^2-2K}$. Note we have normalized the modes so that $H_1(0)=1$; this mode drops from unity into damped oscillations when the wavelength reaches some fraction of the horizon that decreases with m and hence increases with w. In Fig. 10, we plot the derivative of these modes, as that determines its effect on the radiation through gravitational redshifts.

For $w<-1/3$, the universe accelerates and the horizon stops growing with the scale factor. This implies that gravity waves will also freeze out at some finite value related to their amplitude when the universe began accelerating. The solutions for the gravity wave behavior relative to the epoch of freeze-out follow the form of Eq. (77) with $x=k\int_\eta^\infty d\eta$. The H_1 mode then has damped oscillations as x increases from negative values but freezes in to a finite value as $x\to0$ in the infinite future.

Applications. The solutions above help us understand the phenomenology of tensor anisotropies in the CMB. The tensor analogue to Eq. (53) is

$$\frac{\Theta_l}{2l+1}\approx\sqrt{\frac{3}{8}\frac{(l+2)!}{(l-2)!}}\int_0^{\eta_0}d\eta e^{-\tau}\dot{H}_T^{(\pm2)}\frac{j_l(k\Delta\eta)}{(k\Delta\eta)^2}, \tag{78}$$

where $\Delta\eta=\eta_0-\eta$. The power spectrum is again given by Eq. (54). The appearance of the $e^{-\tau}$ damping reflects the fact that anisotropic stress cannot be supported in the optically thick limit; we take $\tau\to0$ in the examples here to consistently neglect all anisotropic stress effects. This in fact is a good approximation for tensor anisotropies in the neutrino background radiation.

For a flat universe with constant $w>-1/3$, the results are shown in Fig. 11 for the same scale-invariant initial spectrum of gravity waves, $k^3|H_T^{(\pm2)}|^2=$const. Notice that the anisotropies decrease as w is decreased. Like the scalar ISW effect discussed in Sec. V D, the contribution of a decaying tensor mode to the anisotropy depends on how long the gravity wave takes to decay relative to the light travel time across the perturbation. In the limit that it decays before horizon crossing, the photons experience the full gravitational effect. If it decays well after horizon crossing, then the effect suffers cancellation as the photons traverse many wavelengths of the perturbation. Thus, the relative contribution to the anisotropy can be read off the behavior of the normal mode in Fig. 10. As w increases, changes in H_T occur at smaller times relative to horizon crossing. The anisotropy contribution accordingly goes up. The effect is most dramatic for the quadrupole since all k-modes above the horizon contribute to the quadrupole. This effect explains why the tensor spectrum in the usual

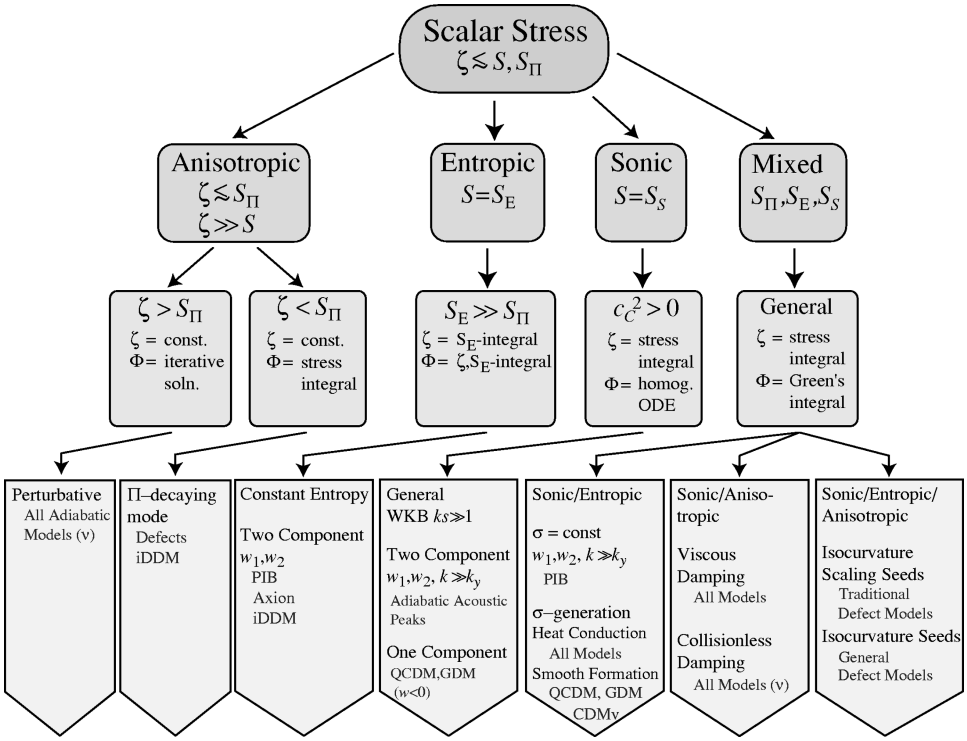

FIG. 12. Scalar stress perturbations. Scalar stress perturbations are divided into three pure classes: anisotropic, entropic, and sonic. If one of these is dominant, then general techniques (third row) can be applied to analyze the resultant behavior of fluctuations. Many models (fourth row) do go through phases where the stresses are dominated by one of the pure stresses. The designations ''all models'' and ''all adiabatic models'' assume a neutrino background and fluctuations that were present before last scattering. Several cases with analytic solutions illustrate the range of behaviors.

matter-radiation universe shows an upturn in the spectrum as one goes from modes that crossed the horizon in the matter dominated epoch to those that crossed in the radiation dominated epoch (see Fig. 17).

For $w<-1/3$ the gravity waves freeze out. Since CMB anisotropies are driven by changes in the gravity-wave amplitude, the addition of a $w<-1/3$ component should suppress anisotropies; this prediction is in agreement with the effect found in ΛCDM and QCDM models [41,42].

VII. PURE STRESSES

We have shown in the previous section that in certain regimes stress perturbations can be ignored. However, to provide a complete history of structure formation, one must track the perturbations across all scales and time. We will first consider the pure stress cases in which the dominant stress contribution is anisotropic, entropic, or sonic, as defined in Sec. V A. These prototypical cases have analytic solutions and are the starting point for the general cases discussed in Sec. VIII. They also have direct application in many models. The anisotropic and sonic solutions are applicable to all adiabatic variants of the CDM model (CDMv). The entropic solutions show how prototypical isocurvature models such as the baryon (PIB) or axion (AXI) isocurvature models form structure. We outline these results in Fig. 12.

A. Anisotropic stress

Anisotropic stresses play a special role because they enter directly into the Newtonian metric through S_Π, as opposed to other stresses which only contribute through the causal motion of the matter. This is despite the fact that the comoving curvature ζ depends only on S. In an isocurvature model, only the comoving curvature need vanish initially, not the Newtonian curvature.

Similarly, even though ζ remains constant above the horizon in adiabatic models, the Newtonian curvature Φ evolves under anisotropic stresses [see Eq. (52)]. Recall that

$$S_\Pi = -8\pi G a^2 p \Pi/k^2 = -3\left(\frac{\dot{a}}{a}\right)^2 \frac{p}{\rho_{\rm cr}}\Pi/k^2, \qquad (79)$$

such that its effect is enhanced by $(\rho'/\rho_{\rm cr})(k\eta)^{-2}$ relative to S. Note, however, that once the universe enters a period when $(p/\rho_{\rm cr})\Pi \ll (k\eta)^2\zeta$, all traces of the anisotropic stress from any previous period vanish in the relation between Φ and ζ. Its effect does not vanish from the CMB, however, since anisotropies record a time-integrated history of the gravitational potentials.

Full solutions. To close this system of equations, we need a relation between S_Π and ζ. We will consider two limiting cases: when $S_\Pi < \zeta$ as in the case of stresses from radiation backgrounds and when $S_\Pi \gg \zeta$ as is possible with models involving active sources such as defects.

In the former case, anisotropic stress is generally created as a by-product of gravitational instability. Its anisotropic nature suggests that it can be created from shear in the velocity and metric, and its coordinate transformation proper-

ties demand that its source be gauge invariant. The linear combination of these sources that satisfies these requirements is $(kv-H_T)$. The anisotropic stress can also have a dissipation time scale η_Π. Together these considerations imply an evolution of the form

$$\dot{\Pi}+\eta_\Pi^{-1}\Pi=4(kv-H_T)\alpha. \tag{80}$$

In this section, we are interested in large-scale effects and hence want the longest time scale for the dissipation; this is set by the expansion time. We will consider cases where the time scale is much smaller than the expansion time in Sec. VIII D.

If we take $\eta_\Pi^{-1}=3\dot{a}/a$, we recover the phenomenological parametrization of[1] [17]

$$\dot{\Pi}+3\frac{\dot{a}}{a}\Pi=4(kv-H_T)\alpha, \tag{81}$$

which has the formal solution

$$\Pi=4a^{-3}\int d\eta a^3(kv-H_T)\alpha. \tag{82}$$

From Eq. (18), we find

$$\Pi=-\frac{8}{3}a^{-3}\int d\eta a^3\alpha\frac{k^2}{1+w}\frac{\rho_{\rm cr}}{\rho}\left(\frac{\dot{a}}{a}\right)^{-2}\left(\dot{\Phi}-\frac{\dot{a}}{a}\Psi\right), \tag{83}$$

for modes well under the curvature scale. Employing Eq. (61) for the zeroth-order potentials and assuming constant α and $\rho=\rho_{\rm cr}$ yields

$$\frac{\Pi}{\zeta}=-2\alpha(k\eta)^2\frac{(1+3w)^2}{(4+3w)(5+3w)}, \tag{84}$$

for constant w. With this source in Eq. (52), the Newtonian potential becomes

$$\frac{\Phi}{\zeta}=\frac{3+3w}{5+3w}(1+\beta), \tag{85}$$

with

$$\beta=16\frac{w}{1+w}\frac{1}{(4+3w)(5+3w)}\alpha, \tag{86}$$

if the curvature contributes negligibly to the expansion rate. If $\beta<1$, this process may be repeated to obtain the desired accuracy. For example, employing the second order form for the Newtonian potential

$$\frac{\Psi}{\zeta}=-\frac{3+3w}{5+3w}\left\{1-\frac{3}{2}\beta\left[1-\frac{3}{2}(1+w)\beta\right](1+w)\right\}, \tag{87}$$

one can build the third order relation for the curvature

$$\frac{\Phi}{\zeta}=\frac{3+3w}{5+3w}\left(1+\beta\left\{1-\frac{3}{2}\beta\left[1-\frac{3}{2}(1+w)\beta\right](1+w)\right\}\right). \tag{88}$$

In the opposite limit that the anisotropic stress contribution is large compared with the other perturbations ($S_\Pi\gg\zeta$, $S_\Pi\gg S$), the integral solutions for the Newtonian metric reduce to

$$\Phi=\frac{\sqrt{\rho}}{a}\int\frac{da}{\sqrt{\rho}}S_\Pi,$$

$$\Psi=-\Phi+S_\Pi. \tag{89}$$

Notice that the integrals remain finite as long as S_Π diverges at zero no faster than $a^{-5/2-3w/2}$ and the prefactor $\sqrt{\rho}/a$ is simply the decaying mode of the Newtonian curvature [see Eq. (59)].

Applications. Collisionless radiation provides an application for these results. The anisotropic stress of radiation is related to the quadrupole moment as defined in Eq. (53) by

$$\Pi_\gamma=\frac{12}{5}\Theta_2 \tag{90}$$

on scales much smaller than the curvature radius. Equation (53) also implicitly gives the equation of motion of the lth multipole as

$$\dot{\Theta}_l=k\left[\frac{l}{2l-1}\Theta_{l-1}-\frac{l+1}{2l+3}\Theta_{l+1}\right]. \tag{91}$$

This is an infinite set of coupled differential equations representing the fact that radiative stress depends on internal degrees of freedom of the distribution. However, [43] introduced an approximation based on a solution of these equations in the absence of sources

$$\frac{\Theta_l}{2l+1}=Cj_l(k\eta), \tag{92}$$

which allows us to express

$$\Theta_{l+1}=(2l+3)(k\eta)^{-1}\Theta_l-\frac{2l+3}{2l-1}\Theta_l. \tag{93}$$

Applying this closure relation to the quadrupole, noting that $(\dot{a}/a)=1/\eta$ in the radiation-dominated era, and rewriting the dipole in covariant form $\Theta_1=v_\gamma-H_T$ shows that the anisotropic stress of free radiation in the radiation-dominated era obeys Eq. (81) with $\alpha=p_\gamma/p=\rho_\gamma/\rho_{\rm rad}$.

The photons actually do not behave in this manner before last scattering since their coupling to the baryons destroys any quadrupole moment in the distribution (see Sec. VIII D). However, the same analysis applies to the massless neutrinos, whose anisotropic stress can be approximately parametrized by $\alpha=\rho_\nu/\rho_{\rm rad}$. Equation (88) shows that the change in Φ is enhanced by a factor $1+(2/15)(\rho_\nu/\rho_{\rm rad})$. With Eq. (56), we find the effective temperature of the CMB to be

[1] $\alpha=c_{\rm vis}^2/w_g$ in the notation of [17].

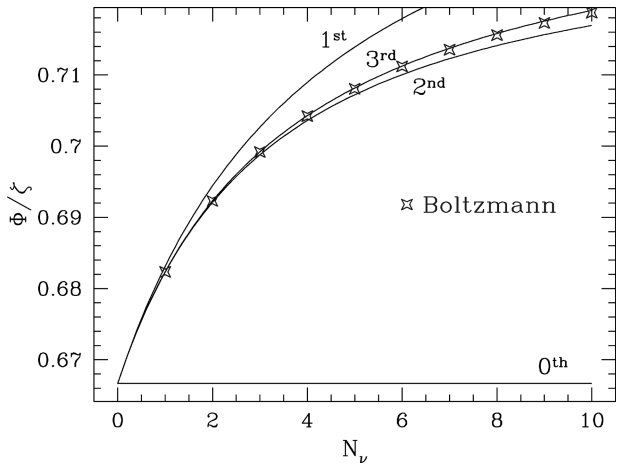

FIG. 13. Anisotropic stress at large scales. The presence of anisotropic stress affects the Newtonian (Φ) but not comoving (ζ) curvature. If $S_\Pi < \zeta$, solutions for Φ can be constructed by iteration. Here we show an example where the anisotropic stress is produced by massless neutrinos and compare numerical results from a Boltzmann code with the iterative solution. To first order, $S_\Pi/\zeta = 4N_\nu/15(4.4+N_\nu)$, where N_ν is the number of massless neutrino species.

$$\Theta_0 + \Psi = -\frac{1}{2}\Psi = -\frac{1}{3}\zeta(0)\left(1 - \frac{4}{15}\frac{\rho_\nu}{\rho_{\rm rad}}\right) \tag{94}$$

and the ISW combination to be

$$\Psi - \Phi = -\frac{4}{3}\zeta(0)\left(1 - \frac{1}{15}\frac{\rho_\nu}{\rho_{\rm rad}}\right) \tag{95}$$

to first order. These results are equivalent to but more physically transparent than those of [10]. In addition, Eq. (88) introduces second and third order corrections that are important for models with additional neutrino species or higher neutrino temperatures T_ν,

$$\frac{\rho_\nu}{\rho_\gamma} = 0.681\frac{N_\nu}{3}\left(\frac{1.401T_\nu}{T_\gamma}\right), \tag{96}$$

as shown in Fig. 13.

Equations (94) and (95) imply that the neutrinos have significant but opposite effects on the effective temperature and ISW terms in the CMB anisotropy equation (53). The degeneracy is broken as the fluctuation enters the horizon, as we shall see in Sec. VIII D, leading to detectable effects in the acoustic peaks [44].

Finally, seed defect sources generally have large anisotropic stress contributions outside the horizon. Causality dictates that $p_s\Pi_s$ behaves as white noise above the horizon which together with the so-called scaling ansatz leads to an $\eta^{-1/2}$ temporal behavior. This implies that $S_\Pi \propto \eta^{-5/2} \propto a^{-5(1+3w)/4}$. Since this diverges slower than $a^{-5/2-3w/2}$, it does not imply a divergence of any observable [32] and contributes mainly to the decaying mode of the curvature [1]. Note that these considerations assume vanishing spatial curvature.

B. Entropic stress

We next consider the case in which entropic stress dominates the other stress components $S_E \gg S_S, S_\Pi$. In this limit S_E does not depend on ζ and Eqs. (51) and (52) are more than simply a formal solution: they are the integral solutions for an arbitrary source in the absence of smooth components. For instance, if the entropic stress has a power-law behavior $S_E \propto a^n$, then the comoving curvature will have the same behavior $\zeta = -\int d\ln a S_E \propto a^n$ assuming $\zeta(0) = 0$.

It is important to note that once S_E turns off, ζ will remain constant. Thus entropic stresses that act for only a fixed amount of time generate curvature fluctuations out of isocurvature initial conditions that then behave in the same manner as initial curvature fluctuations.

A natural way of establishing an entropic stress is to have two components whose sound speeds differ at some point in the evolution. It is useful to introduce the ''entropy''

$$\sigma = \frac{\delta\rho_2}{\rho_2 + p_2} - \frac{\delta\rho_1}{\rho_1 + p_1}, \tag{97}$$

since its equation of motion is simply

$$\dot\sigma = -k(v_2 - v_1) \tag{98}$$

by virtue of Eq. (14), assuming no direct energy exchange between species. The equivalence principle guarantees that under purely gravitational evolution the velocity differences vanish so that a constant σ is a good approximation outside the horizon. Tight coupling between the components also implies $\dot\sigma = 0$.

Let us assume that the density perturbations are accompanied by sonic stresses in the rest frame of each component,

$$\Delta\rho_J = \delta\rho_J - \dot\rho_J(v_J - v)/k,$$

$$\Delta p_J = \delta p_J - \dot p_J(v_J - v)/k, \tag{99}$$

with $c_J^2 = \Delta p_J/\Delta\rho_J$. Under the constant entropy assumption $v_J = v$ and with the definition of the combined sonic stress,

$$c_C^2 = \frac{\rho_1' c_1^2 + \rho_2' c_2^2}{\rho'}, \tag{100}$$

the entropic stress becomes

$$\frac{S_E}{\sigma} = \left(\frac{\rho_1'\rho_2'}{\rho''}\right)(c_2^2 - c_1^2). \tag{101}$$

Consequently, entropic stresses are generated when the sound speeds of the two components differ. Furthermore, they can be much larger than the sonic stresses if $\delta\rho/\rho \ll \sigma$ as is the case for isocurvature models.

Full solutions. If the sonic stresses are also adiabatic ($c_J^2 = p_J'/\rho_J'$), then $S_E = S_\Gamma$ and $S_\Gamma/\sigma = (\rho_1'/\rho')'/3$. Equation (51) then yields

$$\zeta(a)=\frac{\sigma}{3}\frac{\rho_2'}{\rho'}\bigg|_0^a+\zeta(0). \tag{102}$$

The Newtonian potential follows from Eq. (52).

For example, in the constant w_1 and w_2 case

$$\frac{S_\Gamma}{\sigma}=\frac{(1+w_1)(1+w_2)\Delta w}{[1+w_1+(1+w_2)y]^2}y, \tag{103}$$

$$\frac{\zeta}{\sigma}=\frac{1}{3}\frac{(1+w_2)y}{1+w_1+(1+w_2)y}, \tag{104}$$

where recall that $\Delta w=w_2-w_1$ and $y=\rho_2/\rho_1$ and we have assumed isocurvature initial conditions $[\zeta(0)=0]$. Once ρ_2 dominates the energy density, the entropic stress has been converted into a constant comoving curvature $\zeta=\sigma/3$. This solution may then be substituted into Eq. (52) to obtain the integral solution

$$\frac{\Phi}{\sigma}=\frac{1}{3}\frac{y}{r+y}+\frac{1}{9\Delta w}\frac{\sqrt{1+y}}{y^n}\int_0^y dy\,\frac{y^{n-1}}{\sqrt{1+y}}\times\left[1-\frac{r^2+ry(3\Delta w+1)}{(r+y)^2}\right], \tag{105}$$

where $r=(1+w_1)/(1+w_2)$ and $n=-(5+3w_1)/6\Delta w$.

The Newtonian curvature has the limits

$$\frac{\Phi}{\sigma}=\begin{cases}\dfrac{1+w_2}{5+9w_1-6w_2}y & (y\ll 1),\\[2ex] \dfrac{1+w_2}{5+3w_2} & (y\gg 1).\end{cases} \tag{106}$$

The relation between the Newtonian and comoving curvature in the $y\gg 1$ limit is exactly the same as the stress-free clustered case of Sec. VI A.

Another interesting case is when the two components are initially both radiation $w_1(0)=w_2(0)=1/3$ but the second component w_2 becomes non-relativistic. Equation (102) implies that by the time $\rho_2\gg\rho_1$, a curvature fluctuation

$$\zeta=\frac{\sigma}{3}\frac{\rho_1}{\rho}\bigg|_0 \tag{107}$$

is generated from isocurvature conditions.

The more general solution (101) covers multicomponent models with non-adiabatic sonic stresses in one or both of the components and for example may be applied to scalar field models.

Applications. Entropy perturbations between the matter and the radiation are the basis of the prototypical isocurvature models, i.e. the baryon isocurvature model (PIB) and the axion isocurvature model. In this case, the integral in Eq. (105) can be explicitly solved to obtain [34]

$$\frac{\Phi}{\sigma}=\frac{16+8y-2y^2+y^3-16\sqrt{1+y}}{5y^3}, \tag{108}$$

where $y\propto a$. Thus, the curvature grows as a in the radiation-dominated era only to freeze out at the amplitude reached at matter-radiation equality.

Motivated by our study, one of us constructed an iDDM model that utilizes the mechanism described above where the density fluctuations in two radiation species are initially balanced [22]. The resulting constant superhorizon curvature fluctuation ζ is a property generally associated with adiabatic models. Another interesting consequence is that the ratio of large-scale CMB temperature anisotropies to the Newtonian potential is neither 1/3 nor 2. This is because the photons possess initial perturbations. These can be chosen to make the ratio greater than or less than 2; in particular it can be arranged to be close to the adiabatic 1/3 relation.

C. Sonic stress

Sonic stresses provide the final pure case. Here

$$S=S_S=c_C^2\frac{\delta\rho}{\rho+p}; \tag{109}$$

recall that adiabatic stresses are a special case of a sonic stress where $c_C^2=c_s^2\equiv p'/\rho'$. Because the stress fluctuation is related to the comoving density fluctuation by Eq. (109) and in turn to the Newtonian curvature via the hybrid Poisson equation (34), the evolution equation (50) for the Newtonian curvature becomes a homogeneous second order differential equation

$$\frac{\rho'}{a\rho_{\rm cr}^{1/2}}\left(\frac{a\rho_{\rm cr}^{1/2}}{\rho'}\Phi'\right)'+\left(\frac{1}{2}\frac{\rho_{\rm cr}'+\rho'}{\rho_{\rm cr}}-\frac{\rho''}{\rho'}+k^2s'^2\right)\Phi=0, \tag{110}$$

where the sound horizon is defined as

$$s=\int d\eta c_C. \tag{111}$$

Here, we have assumed that the wavelength is much smaller than the curvature scale, but all results are applicable to the general case with the replacement $k\to\sqrt{k^2-3K}$.

For $ks'\gg 1$, we can approximate this as an oscillator equation with an effective mass:

$$(m_{\rm eff}\Phi')'+k^2s'^2m_{\rm eff}\Phi=0, \tag{112}$$

with

$$m_{\rm eff}=\frac{a\rho_{\rm cr}^{1/2}}{-\rho'}. \tag{113}$$

Under the WKB assumption that the effective mass is slowly varying compared with the frequency of oscillation, the fundamental solutions to this equation are

$$\Phi_1=\left(\frac{-\rho'}{c_C}\right)^{1/2}\cos(ks),$$

$$\Phi_2 = \left(\frac{-\rho'}{c_C}\right)^{1/2} \sin(ks). \tag{114}$$

This equation says that once the fluctuation passes inside the sound horizon, the Newtonian potential oscillates and decays, reflecting analogous behavior in the density perturbation through the Poisson equation.

Full solutions. Equation (114) suggests that a change of variables to

$$Q = \left(\frac{c_C}{-\rho'}\right)^{1/2} \Phi \tag{115}$$

should simplify the equations of motion. Indeed, Eq. (110) becomes

$$\frac{d^2Q}{d(ks)^2} + \left(1 - \frac{F}{k^2 s^2}\right) Q = 0, \tag{116}$$

where

$$F = -\frac{s^2}{2s'^2}\left\{\frac{\rho'_{\rm cr}+\rho'}{\rho_{\rm cr}} - \frac{\rho''}{\rho'} + \frac{\rho'''}{\rho'} - \frac{3\rho''^2}{2\rho'^2} + \frac{\rho'_{\rm cr}}{2\rho_{\rm cr}}\left(\frac{\rho''}{\rho'} - \frac{c'_C}{c_C}\right) - \frac{c'_C}{c_C} - \frac{c''_C}{c_C} + \frac{3c'^2_C}{4c^2_C}\right\}. \tag{117}$$

If F is constant, then Eq. (116) is a variant of the Bessel equation, and the solutions Q will approach $\sin(ks)$ (up to a phase) at large ks. Furthermore, when $ks \gg |F|^{1/2}$, the term with F may simply be neglected, and again the solutions will be $\sin(ks)$ with arbitrary phase. Because the latter solutions will match trivially onto the former, we reach the conclusion that if, for a given mode, F is constant until $s \gg |F|^{1/2}/k$, then the appropriate Bessel function solution will hold for *all* times, regardless of how F varies once $ks \gg |F|^{1/2}$. If c_C and ρ vary on the Hubble time scale, then F is order unity, and the solution describes modes that are well inside the horizon before F begins to vary.

The simplest way to arrange F to be constant is to have a constant equation of state w_1 and constant sound speed c_C. For constant w_1, s is only defined for $w_1 > -1/3$, and we therefore set $\rho = \rho_{\rm cr}$. Then

$$F = 6\frac{1+w_1}{(1+3w_1)^2}, \tag{118}$$

independent of the sound speed. The growing and decaying solutions for Φ are then

$$\Phi_1 \propto \left(\frac{-\rho' ks}{c_C}\right)^{1/2} J_\nu(ks),$$

$$\Phi_2 \propto \left(\frac{-\rho' ks}{c_C}\right)^{1/2} N_\nu(ks), \tag{119}$$

where $\nu = (5+3w_1)/(2+6w_1)$. Again, once a given mode is well inside the sound horizon, the condition that w_1 and c_C be constant can be relaxed.

Applications. Adiabatic stress perturbations in a baryon-photon universe provide the prototypical example. In this case, the tight coupling between the photons and baryons through Compton scattering prevents the generation of entropy through the motion of matter (see Sec. VIII C). The growing mode of Eq. (119) becomes [34]

$$\frac{\Phi}{\zeta(0)} = \frac{3}{y^2}\left(\frac{k_y}{k}\right)^2 \left(1 + \frac{3}{4}y\right)^{3/4}\left[\frac{\sin(ks)}{ks} - \cos(ks)\right], \tag{120}$$

with $k_y = (\dot{a}/a)_{y=1}$ and

$$s = \frac{4\sqrt{2}}{3}\frac{1}{k_y}\ln\frac{\sqrt{4+3y}+\sqrt{3y+3}}{2+\sqrt{3}}. \tag{121}$$

We have chosen the normalization to match the stress-free solution $\Phi = 2\zeta/3$ initially. In this case, the acoustic oscillations of the photons are directly related to behavior of the potential via the Poisson equation

$$\Theta_0 = \frac{1}{3}\left(\frac{k}{k_y}\right)^2 \frac{y^2}{1+3y/4}\Phi = (1+3y/4)^{-1/4}\zeta(0)\cos(ks) \quad (ks \gg 1). \tag{122}$$

Since the acoustic oscillations are responsible for the acoustic peaks in the CMB from Eq. (53), this determines the morphology of the acoustic peaks in this simple adiabatic photon-baryon universe. The important result is that the acoustic oscillations follow a $\cos(ks)$ pattern in phase with an amplitude that is enhanced by the decay of the initial curvature perturbation. These results are in fact generic to adiabatic models due to similar evolution in the driving potentials [11].

A second example is provided by the QCDM and GDM models where the equation of state can go negative ($w < 0$) while the comoving sound speed remains real ($c_C^2 > 0$). In the QCDM model, $c_C^2 \equiv 1$ by virtue of the scalar-field equations of motion. In the GDM model, it is allowed to have any positive value. The solutions are strictly valid only after the exotic component dominates the expansion rate and fluctuations. To test this solution, we show an extreme example where the dark matter is all in a ''Q'' component (no CDM) with $w_Q = 0$ (see Fig. 14). The small departures at early and late times are due to the radiation and baryonic contributions.

VIII. MIXED STRESSES

Although the purely anisotropic, entropic and sonic stress cases of the last section illustrate many aspects of stress phenomenology, when all three are present they can interact and create a diverse range of behavior. In this section, we study a few typical cases: sonic stress generation from entropic stress, entropic stress generation from sonic stress, and anisotropic stress dissipation of sonic stress. The first process is

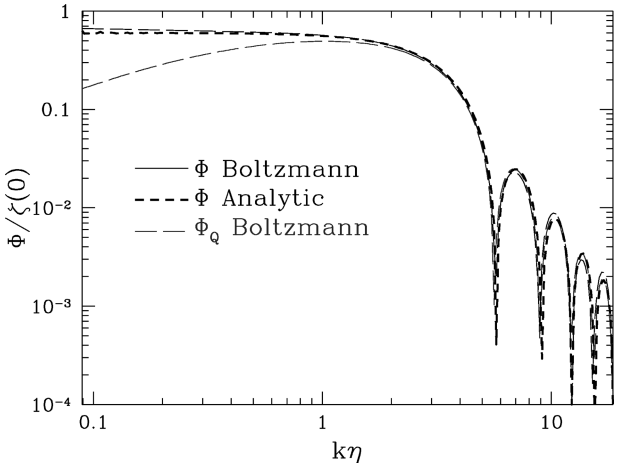

FIG. 14. Sonic stresses in a scalar-field dominated universe. Sonic stresses cause scalar-field perturbations to stop growing and hence lead to a decay in the Newtonian curvature in a scalar-field dominated universe. Here we compare a QCDM model with $w_Q = 0$ and no CDM with the analytic prediction of Eq. (119). The discrepancy at early times is due to radiation contributions in the QCDM model and those at late times from the baryons.

responsible for generating acoustic phenomena in all isocurvature models. The second process is responsible for the heat conduction in fluids and the generation of smooth components. The last process is responsible for viscous dissipation of acoustic phenomena and also provides an alternate means of generating a smooth component. We briefly discuss seed (defect) stresses, which carry not only all three types of scalar stress but also vector and tensor stresses as well. Finally, we consider the effect of tensor anisotropic stress on gravitational waves and CMB anisotropies in passive models.

A. Formal solution

If the equations of motion for a sonic stress perturbation are known to be solved by Φ_1 and Φ_2, the full solution can be written in integral form as

$$\Phi = A\Phi_1(a) + B\Phi_2(a) + \int_0^a \frac{d\tilde{a}}{\tilde{a}} \frac{\Phi_1(\tilde{a})\Phi_2(a) - \Phi_1(a)\Phi_2(\tilde{a})}{\Phi_1(\tilde{a})\Phi_2'(\tilde{a}) - \Phi_1'(\tilde{a})\Phi_2(\tilde{a})} \times \left[\frac{1}{2}\frac{\rho'}{\rho_{\rm cr}} S_E + S_\Pi{}' + \left(\frac{1}{2}\frac{\rho_{\rm cr}' + \rho'}{\rho_{\rm cr}} - \frac{\rho''}{\rho'} \right) S_\Pi \right], \quad (123)$$

where A and B are arbitrary constants. This solution is merely formal unless the remaining stress sources can be specified independently of Φ. We now consider some special cases.

B. Entropic and sonic stress

A model that begins with isocurvature initial conditions will generate adiabatic density fluctuations through the action of entropic stresses, as discussed in Sec. VII B. These density perturbations carry with them sonic perturbations through Eq. (39) that stop the further growth of density perturbations through the same pressure support mechanism discussed in Sec. VII C. Therefore, entropic growth of curvature fluctuations generally ceases once the fluctuations cross the sound horizon of the dominant species.

Full solutions. The solutions for Φ in Eq. (119) along with the entropic stress from Eq. (101) in Eq. (123) allow us to construct the full solution in the presence of a constant entropy and $\rho = \rho_{\rm cr}$:

$$\frac{\Phi}{\sigma} = -\frac{2}{3}\pi^2 G \left(\frac{-\rho' s}{c_C} \right)^{1/2} \int d\tilde{\eta} Y(\tilde{\eta}) \times [J_\nu(k\tilde{s}) N_\nu(ks) - J_\nu(ks) N_\nu(k\tilde{s})], \quad (124)$$

where

$$Y(\eta) = a^2 \left(\frac{-\rho' s}{c_C} \right)^{1/2} \left(\frac{\rho_1' \rho_2'}{\rho''} \right) (c_2^2 - c_1^2). \quad (125)$$

Recall that the combined and component sound speeds were defined in Eqs. (100) and (99) respectively. This solution is valid for $k/k_y \gg 1$ and $k \gg |K|^{1/2}$ (see Sec. VII C).

Applications. The case of $w_1 = 1/3$ and $w_2 = 0$ is of special interest because it corresponds to baryon-photon entropy perturbations, as in the PIB model before last scattering. The result is that on small scales where $ks \gg 1$ and $k/k_y \gg 1$, the curvature behaves as [34]

$$\frac{\Phi}{\sigma} = \frac{3}{4y} \left(\frac{k_y}{k} \right)^2 \left[1 - \frac{\sqrt{3}}{2}\frac{k_y}{k} \frac{(4+3y)^{3/4}}{y} \sin(ks) \right], \quad (126)$$

which is again directly related to the temperature oscillations by Eq. (122). The first term in brackets is due to density perturbations in the baryons remaining from the constant entropy condition $\sigma = \delta_b - \frac{3}{4}\delta_\gamma \approx \delta_b$. The second term represents decaying acoustic oscillations from the adiabatic pressure. The extra factor of k_y/k reflects the fact that the curvature grows as a until sound horizon crossing at $a_H \propto (k_y/k)$ [45].

An interesting result is that the acoustic oscillation follows a $\sin(ks)$ relation, implying the opposite phase in the acoustic peaks compared with the $\cos(ks)$ adiabatic case [10]. This result is rather generic to isocurvature models again due to the similar behavior of the driving potentials. For example, axionic isocurvature models where the entropy is between the radiation and the CDM also follows this pattern.

For this reason, isocurvature seed pressure also tends to generate this type of acoustic pattern [11]. These stresses are found in topological defect models; indeed the dominant scalar modes of strings, monopoles, textures, and nontopological textures do behave in this manner [23]. However, only in the latter two are the other modes sufficiently small at the first few peaks to yield clean acoustic peaks even for the scalar perturbations alone. Defect models generically have vector and tensor stresses that generate comparable levels of anisotropy and further obscure acoustic phenomena.

It is possible to construct isocurvature models with an ''adiabatic'' pattern of acoustic peaks. The simplest way to

arrange this is to create constant comoving curvature perturbations ζ through entropic stresses that turn off well before the perturbation crosses the horizon. A concrete example of this kind of mechanism is given in [22]. As is clear from Sec. VII B and originally pointed out by [11], for an isocurvature model to generate *scale-invariant* curvature perturbations requires that the entropic stress have superhorizon scale correlations which cannot be generated causally without an inflationary epoch.

C. Sonic and entropic stress

Likewise, an adiabatic or sonic fluctuation will not generally remain so as it crosses the horizon. Inside the horizon, the fact that components with different equations of state have different pressure responses to gravitational compression will cause the species to move independently. The generation of entropic stresses is in fact a primary mechanism for creating the smooth components of matter discussed in Sec. VI B.

Full solutions. The case of two components with constant w_1 and w_2 again provides an instructive example. The equation of motion for the entropy, Eq. (98), is constructed out of combining the continuity equations (14) of the two species and has the formal solution

$$\sigma = \sigma(0) - k\int d\eta (v_2 - v_1). \tag{127}$$

Recall that the entropy is related to the entropic stress by Eq. (101). Note that entropy leads to entropic stress only if the two components differ in their equation of states ($\Delta w \neq 0$).

Since we are interested in the generation of entropy, we will assume that the initial entropy perturbation $\sigma(0)$ vanishes. Two interesting cases are when the entropy generation time scale η_Γ is small compared to the expansion time and when it is comparable to the expansion time scale.

In the former case, the entropy will lead to dissipative behavior if

$$\sigma \approx -k\eta_\Gamma v, \tag{128}$$

where η_Γ is some characteristic time scale for entropy generation. While this form is not gauge invariant even though σ is, the ambiguity vanishes inside the horizon. Hence, Eq. (128) is a good approximation in the desired case $\eta_\Gamma \ll \eta$. Substituting Eq. (128) into Eq. (101) and then Eq. (15) in Newtonian gauge and assuming a solution of the form $v \propto \exp(i\int \omega d\eta)$, we obtain the dispersion relation

$$\omega = \pm k c_C - i\frac{1}{2}\frac{\rho_1'\rho_2'}{\rho'^2}(c_2^2 - c_1^2)k^2\eta_\Gamma, \tag{129}$$

where recall c_C was defined in Eq. (100). Note that we have assumed $\eta_\Gamma/\eta \ll 1$ in order to replace rest-frame sound speeds with comoving sound speeds.

This describes an exponential damping of sound waves as the wavelength passes the ''diffusion'' scale $k = (\eta_\Gamma \eta)^{-1}$.

In the opposite regime, where the entropy is generated on the horizon scale, the sonic nature of the total system breaks down before acoustic oscillations even start. In this case, components can decouple completely and form subsystems where sonic, anisotropic, or smooth effects can occur separately. Here it is simpler to describe the behavior of individual subsystems through variables that are comoving with respect to the individual species J. The sound speed c_J and the contribution to the Newtonian potential are useful [see Eq. (99)]:

$$\Phi_J = 4\pi G \Delta\rho_J/(k^2 - 3K), \tag{130}$$

which implies $\Phi = \Sigma_J \Phi_J$. If the density fluctuations in the other species are still below their own sound horizon, then the oscillating components can become effectively smooth in comparison. From that time forward, we can treat the system as having a smooth component, and the curvature fluctuation contributed by the other species is governed by Eq. (69).

Applications. The case of a short entropic timescale $\eta_\Gamma \ll \eta$ is realized in the photon-baryon fluid before recombination and is relevant for considering the damping of acoustic phenomena in the CMB for all models of structure formation [46,47]. Here the entropic time scale is derived from the baryon Euler equation

$$v_b - v_\gamma \approx \dot{\tau}^{-1} R \dot{v}, \tag{131}$$

where $R = 3\rho_b/4\rho_\gamma$. Inserting this into Eq. (127), we obtain $\eta_\Gamma = \dot{\tau}^{-1} R$ under the rapid oscillation assumption. The dispersion relation then becomes [47]

$$\omega = \pm k c_s + i\frac{1}{6}k^2\dot{\tau}^{-1}\frac{R^2}{(1+R)^2}. \tag{132}$$

These entropic pressure techniques provide a simpler and more transparent derivation of this well-known result than exists in the literature. Note that heat-conduction damping is suppressed by R in the photon-dominated epoch, and we shall see that in that case viscous damping from the anisotropic stress is more important.

The opposite limit of an entropic timescale on order the expansion time $\eta_\Gamma \sim \eta$ is applicable to all models with CDM. The CDM never participates in the acoustic oscillations of the baryon-photon system even during the radiation-dominated era. In Fig. 15, we show an example of a perturbation deep in the radiation-dominated era in a CDM model with the usual neutrino content. Before horizon crossing $k\eta \ll 1$, the perturbations are adiabatic and the total potential Φ is constant with contributions from the photons Φ_γ and the neutrinos Φ_ν. The sound horizon for the radiation is $s = \eta/\sqrt{3}$, and once $k\eta \gg \sqrt{3}$, the potential contribution of the photons starts to decay as in Eq. (120). The neutrino contributions decay even faster due to their anisotropic stress, as we will discuss in the next section. The CDM contribution $\Phi_{\rm CDM}$ then turns over and behaves as $\Phi_{\rm CDM} \propto \ln(Ca)/a$ where C is some constant [39]. This is exactly the behavior predicted by Eq. (69) assuming that the radiation component is smooth. In actuality, the photon component is not smooth compared with the CDM component but oscillates sufficiently rapidly that its time-average is negligible.

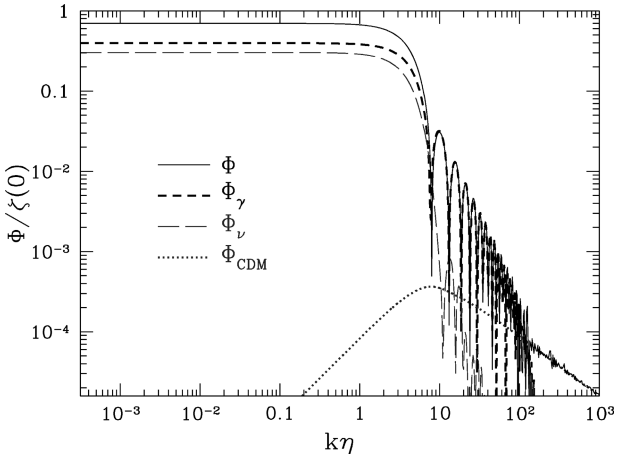

FIG. 15. Creation of a smooth radiation background. Numerical results in a ΛCDM model are shown here. Notice that upon crossing the sound horizon, the photon contribution Φ_γ to the total curvature Φ damps and oscillates, while the neutrino contributions damp much more rapidly due to collisionless damping from its anisotropic stress. The oscillating photon perturbations yield little time-averaged effect, and the CDM evolves under Eq. (71), leading to the well-known logarithmic tail in the CDM transfer function [48].

These considerations also apply to models in which an additional component with $w_{\rm GDM}<0$ comes to dominate at late times. Quintessence and GDM models (see Fig. 16) are examples thereof. In Sec. VII C, we considered the case where the GDM completely dominated the expansion. Here we consider the effect of adding a component of CDM.

The critical parameter here is the comoving sound speed of the GDM $c_{\rm GDM}$. Slowly rolling scalar fields found in the QCDM subcategory have $c_{\rm GDM}=1$, but components with lower sound speeds are possible in principle. The sound speed tells us how long after horizon crossing the GDM component stabilizes due to pressure support. If the perturbation is already within the sound horizon by the time the GDM comes to dominate the expansion rate ($y=1$), then the total potential will behave under Eq. (69) as if the potential is smooth regardless of the exact value of the sound speed. As the sound speed is lowered such that crossing occurs near ($y=1$), we see effects from the finite sound speed. In the limit where the perturbation remains above the sound horizon until the present, the solution returns to the clustered case of Eq. (59).

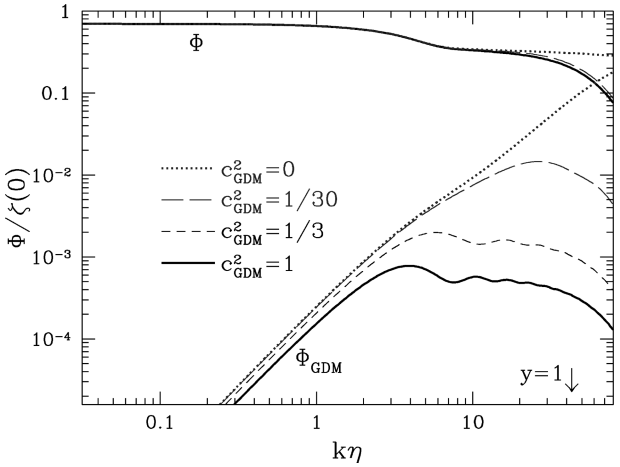

FIG. 16. Creation of a smooth quintessence or GDM component with $w_{\rm GDM}=-1/3$. As shown in Fig. 14, sonic stresses in the GDM component prevent perturbation growth inside the sound horizon $c_{\rm GDM}k\eta\sim 1$. If this occurs well before GDM domination at $y=1$, then the total Newtonian curvature will evolve as if the component were always smooth [i.e. under Eq. (71)]. If this occurs well after, the Newtonian curvature will behave as if all components were fully clustered [i.e. under Eq. (63)]. For a given scale, this depends on the comoving sound speed $c_{\rm GDM}$; for quintessence $c_{\rm GDM}=1$.

D. Sonic and anisotropic

We have seen in Sec. VII A how anisotropic stress can generate Newtonian curvature perturbations but, like adiabatic stress, it can also destroy them. In this context, anisotropic stress represents the ''frictional force'' set up in response to the non-uniform bulk flow of matter and shear in the metric. As we shall see, this sort of behavior is not confined to fluids. It represents another mechanism for generating a smooth component.

Full solutions. The phenomenological parametrization of anisotropic stress in Eq. (80) yields two interesting limits. The short time scale limit ($k\eta_\Pi\ll 1, \eta_\Pi/\eta\ll 1$) leads to viscous or collisional damping. Here, the anisotropic stress is algebraically related to the velocity in shear-free frames

$$\Pi = 4\alpha\eta_\Pi(kv - H_T). \tag{133}$$

Under the same assumptions used to derive the dispersion relation for heat conduction, Eq. (129), we find the viscous dispersion relation

$$\omega = \pm kc_C + i\frac{4}{3}k^2\alpha\eta_\Pi\frac{p}{p+\rho}\alpha, \tag{134}$$

implying dissipation at a characteristic scale $k\sim 1/\sqrt{\eta\eta_\Pi}$.

If the dissipation time scale is comparable to the expansion time scale, the damping occurs at horizon crossing but is more gradual. The formal solution in this limit is given in Eq. (82) and approximates the effects of collisionless damping.

Applications. Collisional damping occurs in the photon-baryon fluid before last scattering and is the primary dissipation mechanism for acoustic oscillations in the CMB [46]. Equation (133) then describes the anisotropic stress of the photons Π_γ with $\alpha=2/5$ and $\eta_\Pi=\frac{4}{3}\dot\tau^{-1}$. Repeating the calculation leading to Eq. (132), the dispersion relation for the oscillations becomes [46,49]

$$\omega = \pm kc_C + i\frac{8}{45}k^2\dot\tau^{-1}\frac{1}{1+R}. \tag{135}$$

In comparison to the heat conduction dissipation of Eq. (132), viscous dissipation is more effective if $R<1$, which is the case for the baryon content implied by big bang nucleosynthesis (e.g. [50]).

Collisionless damping occurs in free radiation. Free radiation behaves in this ''frictional'' manner because gradients in the potential flow are dissipated as radiation streams from one part of the flow to the next. In terms of the multipole moments, power in the dipole gets transferred to the quadrupole and so on through the hierarchy equations (91). As such, the anisotropic stress acts as the gateway for anisotropy generation in the CMB. It also leads to a damping of density and velocity perturbations inside the horizon for all species of free radiation and is the reason that in Fig. 15 the neutrino contributions damp more rapidly than the photon contributions. More generally, it is responsible for smoothing out components when entropic stresses cannot be generated, e.g. when $c_2^2 - c_1^2 = 0$.

These dissipational terms may also be important in stabilizing other forms of matter. One way to generate a smooth density component is to introduce an effective viscous rather than sonic stress [17]; a mechanism of this type is thought to be involved in stabilizing the strCDM model [16]. Turok [51] suggested an extreme example of this sort, in which the comoving sound speed of the seeds is imaginary but the anisotropic stress is perfectly balanced to counter the otherwise exponential growth of density fluctuations.

E. Sonic, entropic, and anisotropic seed stress

The seed stresses of topological defect models provide an example where all types of stresses coexist. While the consequences of a given seed stress for structure formation are straightforward to work out, the behavior of the seed stress itself is more difficult and requires simulations to work out in detail.

Full solutions. The two-point statistics of complicated defect models can be accurately modeled as the incoherent (quadrature) sum of a relatively small number (~ 10–20) of simple seed stress histories [51,23]. Each individual source may then be determined by the techniques above [see Eqs. (44) and (45)].

Applications. The simplest defect models typically have two other properties that have important phenomenological consequences. Defect models are causal in the classical sense. The stress perturbations δp_s and $p_s \Pi_s$ must fall off at least as white noise (k^0) outside the horizon and the initial curvature must vanish [$\zeta(0) = 0$]. The traditional string and texture models also obey a scaling relation that states that the stress histories depend on wavelength only through the combination $k\eta$. This ensures self-similarity of the structure at horizon crossing and leads to nearly scale-invariant CMB anisotropies. Because the simplest versions of these models run into difficulties when CMB anisotropies are compared with large-scale structure, phenomenological models that do not obey the scaling relation have recently received some attention [25].

F. Vector stress

We next consider the effect of vector anisotropic stresses. Recall that in the absence of vector stress, the vector perturbation decays. In order to generate an observable effect, vector perturbations must be continuously generated by vector anisotropic stresses. This implies that vector modes are unique to the active stress models, of which defects are an example.

Full solutions. The solutions in the presence of stress sources can be constructed via Green's function techniques from the stress-free solutions of Eq. (76). For the vector modes, the solution becomes

$$v^{(\pm 1)} - B^{(\pm 1)} = a^{-4}(\rho + p)^{-1}\left[C - \frac{1}{2}(1 - 2K/k^2) \times k \int_0^\eta d\tilde{\eta}\, a^4 p \Pi^{(\pm 1)} \right], \tag{136}$$

where $C = \text{const}$ represents the decaying mode. The remaining metric perturbation $kB^{(\pm 1)} - \dot{H}_T^{(\pm 1)}$ is related algebraically to Eq. (136) through Eq. (21).

Applications. The discussion of scalar seed stress in defect models in Sec. VIII E also applies to vector stresses with a few additional considerations. Defect models generally have comparable scalar, vector, and tensor anisotropic stress sources above the horizon [52]. Since CMB anisotropies are primarily generated at horizon crossing, these sources tend to yield comparable anisotropy contributions for modes that cross after last scattering. Vector modes that cross before last scattering do not contribute due to suppression of tensor anisotropies in the CMB from scattering. Thus, vector modes can obscure the first few scalar acoustic features in defect models. On the other hand, vector modes have a special signature in the CMB polarization that may assist in their isolation [29,53].

G. Tensor stresses

Tensor anisotropic stresses provide sources and sinks for gravity waves. It is well known that quadrupolar stresses generate gravity waves. Furthermore, a passing gravity wave will also impart some energy to the radiation backgrounds via differential gravitational redshifts and thereby decay.

Full solutions. An integral solution may be constructed out of the homogeneous solutions of Sec. VI D as

$$H_T^{(\pm 2)}(\eta) = C_1 H_1(\eta) + C_2 H_2(\eta) + \int_0^\eta d\tilde{\eta} \frac{H_1(\tilde{\eta}) H_2(\eta) - H_1(\eta) H_2(\tilde{\eta})}{H_1(\tilde{\eta}) H_2'(\tilde{\eta}) - H_1'(\tilde{\eta}) H_2(\tilde{\eta})} \times 8\pi G a^2 p \Pi^{(\pm 2)}, \tag{137}$$

where C_1 and C_2 are constants associated with the initial conditions.

Applications. We show the damping effect of anisotropic stress in the radiation backgrounds in Fig. 17. It is generally not included in standard Boltzmann codes that solve CMB anisotropies [54]. The reason is that it is negligible at large angles since the corresponding modes entered the horizon well into matter domination when the anisotropic stresses are negligible. Since this feedback effect is typically small, Eq. (137) may in principle be used to iterate to a solution from

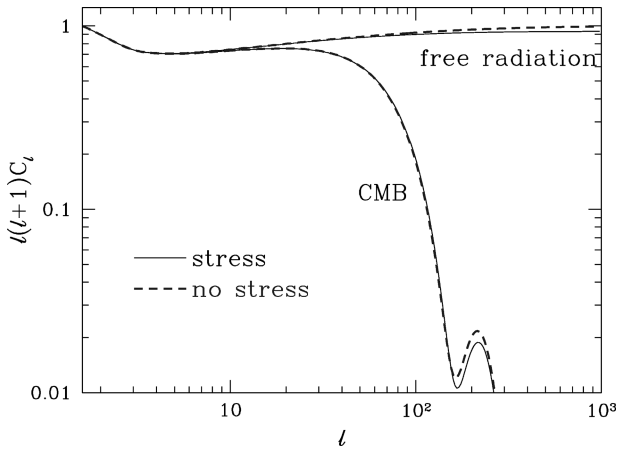

FIG. 17. The effect of tensor anisotropic stress on tensor anisotropies. Gravity waves generate tensor anisotropic stress in radiation that absorbs energy and damps the gravity wave. For free radiation like the neutrino background radiation, this reduces the anisotropies generated on small scales. For the CMB, the elimination of tensor anisotropic stresses by scattering cuts off small-scale contributions as well.

the free-gravity wave case of Eq. (78). However, the evolution of the anisotropic stress has a more important effect. Just as scalar anisotropic stresses in the photons are destroyed by scattering before last scattering (see Sec. VIII D), tensor anisotropic stresses are destroyed as $e^{-\tau}$. This cuts off the tensor contributions to the anisotropies before last scattering as indicated in Eq. (78) and shown in Fig. 17. The feedback effect still exists, but the level of the anisotropy itself makes it too low to be observable for reasonable tensor to scalar ratios.

The Green's function solution of Eq. (137) is more directly applicable to the seed stress case where $p\Pi^{\pm 2} = p_s \Pi_s$. As in the vector stress case, the phenomenological result is that defect models tend to have significant tensor contributions above the angle the horizon subtends at last scattering. As in the passive models, their contribution is cut off below this scale due to scattering.

IX. DESIGNER APPLICATION

With this general study in hand, we are now in a position to discuss prospects for reverse engineering the model for structure formation. Obviously, the specific route the inversion takes will depend on the results of ongoing experiments. Currently, the data favor a model with phenomenology like the ΛCDM model, e.g. the shape of the large-scale structure power spectrum [55,56], relative high to low redshift cluster abundances [57,58], supernova luminosity distances to redshift of $z \sim 0.5$ [59–61], and degree scale CMB anisotropies (e.g. [62,63]). If agreement between the ΛCDM model and future precision tests is good, can we say purely on phenomenological grounds that we have proved the existence of a cosmological constant? If the model varies from the data, can we use the methods developed here to modify the stress history, restore agreement, and in the process learn new information about the dark sector?

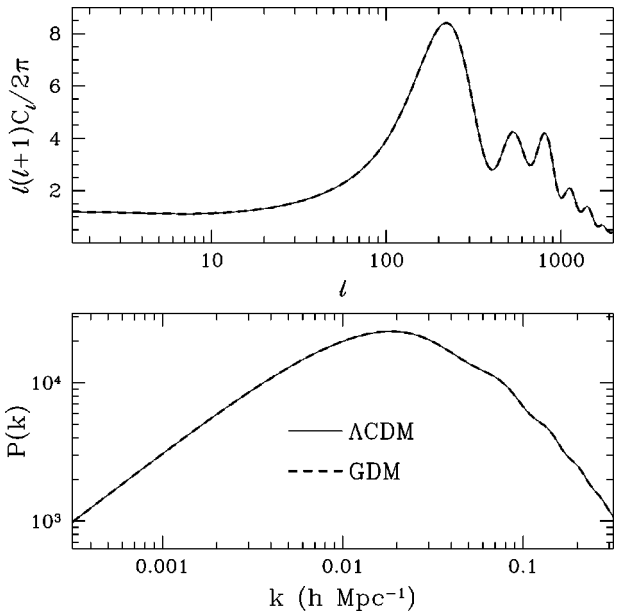

FIG. 18. Mimicking ΛCDM. Models with the same stress histories in the dark sector have the same observable consequences regardless of differences in how that sector is composed. Here numerical solutions for the CMB anisotropy and large-scale structure power spectra in a ΛCDM universe and a single-component critical GDM dominated universe of Eq. (139) are shown to be identical. However, variations in c_C, $\Delta\Phi$, and $\Delta\eta$ away from this tuned stress history do have observable consequences.

Let us address the first question. We know that the evolution of structure is completely defined by the stress history of the matter. Since the equations of state of the ordinary matter are known, the remaining element is the dark sector. To test the uniqueness of the ΛCDM model, we should look for alternate means of reproducing its stress history. The background stress history of the dark sector in a Λ model is given by

$$w_{\Lambda\mathrm{CDM}} = \frac{-a^3}{a^3 + \Omega_{\mathrm{CDM}}/\Omega_\Lambda}, \tag{138}$$

and its stress perturbations vanish. This suggests that a generalized dark matter component of the type introduced by [17] and parameters

$$w_{\mathrm{GDM}} = w_{\Lambda\mathrm{CDM}}, \quad c^2_{\mathrm{GDM}} = 0 \tag{139}$$

should reproduce the phenomenology of the ΛCDM model. Recall that c_{GDM} is the sound speed in the frame comoving with the GDM (see Sec. VIII C). To the extent that the comoving and GDM-comoving frames coincide, this form of dark matter exactly reproduces any mixture of Λ and CDM. We show an example in Fig. 18. Note that all classical cosmological and linear theory tests will return the same answer for the two models despite the fact that the GDM model is a single dark matter component model in a critical density universe. Non-linear effects are also identical if the same stress history is maintained throughout.

There are two lessons here. The first is that on purely phenomenological grounds we can do no more than measure the global properties of the dark sector. A multicomponent model and a single component model with the same stress history are formally identical.

The second is that reproducing the phenomenology of a ΛCDM model is rather simple: it requires an equation of state that varies from $w=0$ at high redshift to $w\approx -2/3$ today and a form of matter that is free of large-scale stress gradients in its comoving frame. Thus, there exists a wide class of such single component models that fit the current data as well as the ΛCDM model.

On the other hand, variations on the conditions in Eq. (139) have observable consequences for future measurements. Relaxing the stress-free perturbation condition by raising the sound speed reduces the small-scale power in the model and delays the formation of high redshift objects. The remaining freedom in the stress history is associated with the equation of state w. We can quantify this by recalling that the gravitational potential depends only on the quantity [see Eq. (59)]

$$\Delta\Phi \propto \frac{\sqrt{\rho}}{a}\int_0^a \frac{da}{\sqrt{\rho}}. \tag{140}$$

Large-scale structure constrains the value of this integral at the present ($a=1$). CMB anisotropies from the ISW effect are sensitive to variations in this function that occur on the order of the light-travel time across a wavelength (see Sec. V D). Since CMB anisotropies potentially probe nearly three orders of magnitude from the current horizon, variations at a fraction of a percent of the current expansion time are potentially visible in the CMB.

Similarly, distance measures such as the angular diameter distance to the last scattering surface ($a\sim 10^{-3}$) and the luminosity distance to high redshift objects ($a\sim 2/3$) probe the combination

$$\Delta\eta(a) \propto \int_a^1 da\, \frac{1}{a^2\rho^{1/2}}. \tag{141}$$

Current measures of $\Delta\Phi(a)$ and $\Delta\eta(a)$ are crude at best as they only proble their values at discrete epochs or are averaged over long time scales. A sharp test of the Λ + CDM hypothesis that should be possible with future measures involves reconstructing the time evolution of the equation of state $w_{\rm cr}$ through measures of $\Delta\Phi(a)$ and/or $\Delta\eta(a)$. Any combination of a cosmological constant and CDM will obey

$$w'_{\rm cr} = 3w_{\rm cr}(1+w_{\rm cr}). \tag{142}$$

The physical implication of this relation is that the background pressure is constant in time ($p'=0$) so that the adiabatic sound speed vanishes ($c_s^2 = p'/\rho' = 0$). If this condition is violated, we will have proved the existence of a new form of matter. If $w'_{\rm cr} > 3w_{\rm cr}(1+w_{\rm cr})$, then it can be supported by adiabatic stresses [64]; if not, then a form of matter with non-adiabatic stresses is required. Non-adiabatic stresses imply that the relation between the pressure and density perturbations does not follow that of their spatial averages.

The simplicity of the requirements of Eq. (139) raises the possibility of a unified description of the dark sector. A concrete but somewhat trivial example is a scalar field that rapidly oscillates around a non-zero potential minimum. The rapid oscillations average away all large-scale pressure effects save from the vacuum pressure of the potential minimum [65]. Unfortunately, the relationship between the constant and quadratic pieces of the scalar potential is left unexplained and so is no better than an explicit Λ + CDM model. Nonetheless, there may be more complicated examples, perhaps involving multiple fields, in which the mimic conditions (139) are approximately satisfied and do unify the two behaviors in a true sense.

This discussion shows that a reverse-engineered model for structure formation will in general not be unique. On the other hand, the observables can be translated into constraints on the stress histories and phenomenological models of the dark sector. These in turn can assist in the search for compelling physical candidates to compose the dark sector.

X. DISCUSSION

Without any assumptions other than the validity of general relativity and nearly homogeneous and isotropic initial conditions, the evolution of structure is completely determined by the stress history of matter. We have studied the means by which stresses, both in the background and the fluctuations, can alter the observable properties of the model.

We have examined the effects of smooth, anisotropic, sonic, and entropic stresses in structure formation, including their interactions and ability to generate effectively smooth density components. We have illustrated these behaviors with analytic solutions for systems with multiple components of differing background equations of state, which can themselves be time dependent in several important cases. These solutions have applications to nearly all of the current models for structure formation and are substantially more general than those existing in the literature.

Although this study is not exhaustive, we have made explicit all of the assumption required to arrive at specific models and their accompanying phenomenology. In the process, we have exposed the limitations of traditional categorization schemes like that in Fig. 1. These distinctions can be blurred in cases where the usual assumptions do not apply. We summarize several notable cases here.

A. Initial conditions

Isocurvature initial conditions imply a growing comoving curvature outside the horizon on scales relevant to large-scale structure and degree-scale anisotropies. The comoving curvature grows outside the horizon only by the action of stress perturbations. Once stress perturbations are turned off, the curvature remains constant until horizon crossing or curvature domination. These considerations provide a means for mimicking the phenomenology of adiabatic models [22] and

are important for interpreting the implications of CMB acoustic peak phenomenology; however, they do not invalidate the conclusions of [11].

The Newtonian curvature Φ *is simply proportional to the comoving curvature if the background equation of state is constant.* The Newtonian curvature admits a decaying mode whereas the comoving curvature does not. The decaying mode can be stimulated by anisotropic stress perturbations outside the horizon but has observable consequences only through the contribution remaining at horizon crossing [1].

If the Newtonian potential Ψ *is constant from last scattering to the present, the observed temperature perturbation depends only on the equation of state and* Ψ. The assumption here is that the comoving temperature perturbation is negligible and is only true if stress perturbations are also negligible compared with the comoving curvature for all time. The axion isocurvature model provides a simple counterexample. On the other hand, no assumptions about the anisotropic stress are necessary even when $\Phi=-\Psi$ no longer holds.

Isocurvature initial conditions predict an observed temperature perturbation of 2Ψ *on scales larger than the horizon at last scattering.* The assumption here is that the initial temperature perturbation in Newtonian gauge vanishes. This is not the case for models where the isocurvature conditions are established by balancing the photon density perturbations off another species of radiation. Furthermore, changes in the potential that are slowly varying compared with the light travel time across a perturbation do not affect the observed temperature.

B. Clustering properties

The smoothness of a component is gauge-dependent and hence has no physical meaning. The gauge dependence of a smooth component is not a problem *per se* as certain frames, e.g. the frame where the momentum of the component vanishs, are dynamically special. A smooth contribution to the density with $w_S\neq-1$ does violate covariant energy conservation in any coordinate system where the spatial curvature changes. Since the very presence of a smooth density component requires that the comoving curvature perturbation decay, there can be no identically smooth contributions in those coordinates except in the trivial zero curvature perturbation case. A component can be smooth relative to another species inside the horizon where the relativistic effects of curvature variation are negligible [15].

The behavior of a smooth component depends only on its equation of state w_S. Since all components except Λ and curvature are clustered outside the horizon, the manner in which a component becomes smooth is observable. The presence of sonic (supportive) and anisotropic (dissipative) stresses are two possibilities, but others exist in principle.

A ''clustered'' model like standard CDM has no dynamically important smooth stresses or density contributions. The radiation backgrounds and baryons are effectively smooth well inside the horizon prior to recombination. Smooth contributions are generic to structure formation models.

Smooth component behavior implies small density perturbations $\delta\rho_S<\delta\rho_C$. A component can be effectively smooth even while possessing large density fluctuations. The crucial assumption is that their time-average density over the dynamical time of the clustered species is smooth. Density perturbations in a component that vary rapidly with the expansion time generally lead to no effect on the growth of structure. The radiation backgrounds mentioned above provide a familiar example.

The missing mass (clustered dark matter) and missing energy (smooth $w_S<-1/3$ *dark matter) are separate problems.* The stress history of the dark sector completely defines its properties for classical cosmology and structure formation. Any combination of components that produces the same stress history will produce the same phenomenology. As an example, we have constructed a toy model that that exactly reproduces the ΛCDM phenomenology but employs a single component of dark matter in a critical density universe. Variations in the stress history produce models that satisfy the current constraints equally well but are potentially distinguishable from ΛCDM.

C. Perturbation type

Scale-invariant gravity waves preferentially enhance the low-order multipoles of CMB anisotropies. Enhancement only occurs if the gravity wave amplitude decays close to horizon crossing and is eliminated as the equation of state of the background drops.

There is a sharp distinction between active and passive models for structure formation. Models that have been labeled ''passive'' in the literature are those in which the stress perturbations are simply related to density, velocity, and metric fluctuations by equations of state, sound speeds, viscosity parameters, etc. Models in which there is a component whose stresses have no fixed relation to density and metric perturbations have been labeled ''active.'' The issue is the number of internal degrees of freedom that act to specify the stresses. For example, scaling defect stress histories are typically approximated by tens of parameters and those of particle dark matter by three or fewer. In principle, there is a spectrum of possibilities between the two and a corresponding spectrum of phenomenological consequences.

Our study is useful even if the current evidence supporting ΛCDM-type phenomenology holds up. Even the ΛCDM model itself does not have an entirely trivial stress history as its dark sector includes the neutrino background radiation. The observability of the neutrino stress history has been addressed numerically in [44]; we have examined its physical origin here. Furthermore, the dark sector could contain exotic features that produce more or less ordinary phenomenology, and one needs to construct sharp tests against alternatives. For example, a combination of cold dark matter and a cosmological constant must obey $w'_{\rm cr}=3w_{\rm cr}(1+w_{\rm cr})$. This relation also acts as the dividing line between models with exotic and ordinary matter. For exotic matter, the stress and density perturbations obey a relation that opposes that between the background stress and density; scalar fields are one example of exotic matter.

In summary, a purely phenomenological reverse-engineering of the model for structure formation will require the reconstruction of the time-averaged stress history of the dark sector. This inversion is generally not unique. Nonetheless, if the observed phenomenology remains close to that of our simplest models, our study of stress phenomenology should provide the means for constructing viable models. If the observations require more radical departures, our study should be useful in identifying the assumptions that are incorrect and assist in the search for the correct phenomenological model for structure formation.

ACKNOWLEDGMENTS

We would like to thank D.N. Spergel and P.J. Steinhardt for useful conversations. W.H. is supported by the Keck Foundation and Sloan Foundation, D.J.E. by Frank and Peggy Taplin, D.J.E. and W.H. by NSF-9513835.

[1] J. M. Bardeen, Phys. Rev. D **22**, 1882 (1980).
[2] H. Kodama and M. Sasaki, Prog. Theor. Phys. Suppl. **78**, 1 (1984).
[3] V. F. Mukhanov, H. A. Feldman, and R. H. Brandenberger, Phys. Rep. **215**, 203 (1992).
[4] J. M. Bardeen, P. J. Steinhardt, and M. S. Turner, Phys. Rev. D **28**, 679 (1983).
[5] T. J. Allen, B. Grinstein, and M. B. Wise, Phys. Lett. B **197**, 66 (1987).
[6] L. Moscardini, S. Mataresse, F. Lucchin, and A. Messina, Mon. Not. R. Astron. Soc. **248**, 425 (1991).
[7] T. W. B. Kibble, Nucl. Phys. **B262**, 227 (1985).
[8] D. Lyth, Phys. Rev. Lett. **78**, 1861 (1997).
[9] R. K. Sachs and A. M. Wolfe, Astrophys. J. **147**, 73 (1967).
[10] W. Hu and N. Sugiyama, Astrophys. J. **444**, 489 (1995); Phys. Rev. D **51**, 2509 (1995).
[11] W. Hu and M. White, Astrophys. J. **471**, 30 (1996); W. Hu, D. N. Spergel, and M. White, Phys. Rev. D **55**, 3288 (1997).
[12] J. A. Holtzman and J. R. Primack, Astrophys. J. **405**, 428 (1993).
[13] P. G. Ferreira and M. Joyce, Phys. Rev. Lett. **79**, 4740 (1997).
[14] J. A. Frieman, C. T. Hill, A. Stebbins, and I. Waga, Phys. Rev. Lett. **75**, 2077 (1995); K. Coble, S. Dodelson, and J. Frieman, Phys. Rev. D **55**, 1851 (1997); J. A. Frieman and I. Waga, *ibid.* **57**, 4642 (1998).
[15] R. R. Caldwell, R. Dave, and P. J. Steinhardt, Phys. Rev. Lett. **80**, 1582 (1998).
[16] D. N. Spergel and U.-L. Pen, Astrophys. J. Lett. **491**, L67 (1997).
[17] W. Hu, Astrophys. J. **506**, 485 (1998).
[18] G. Efstathiou and J. R. Bond, Mon. Not. R. Astron. Soc. **218**, 103 (1986).
[19] P. J. E. Peebles, Nature (London) **327**, 210 (1987).
[20] P. J. E. Peebles, Astrophys. J. Lett. **483**, L1 (1997).
[21] W. Hu, E. Bunn, and N. Sugiyama, Astrophys. J. Lett. **447**, L59 (1995).
[22] W. Hu, Phys. Rev. D **59**, 021301 (1998).
[23] U.-L. Pen, U. Seljak, and N. Turok, Phys. Rev. Lett. **79**, 1611 (1997).
[24] A. Albrecht, R. A. Battye, and J. Robinson, Phys. Rev. Lett. **79**, 4736 (1997).
[25] A. Albrecht, R. A. Battye, and J. Robinson, Phys. Rev. D **59**, 023508 (1999).
[26] C. Contaldi, M. Hindmarsh, and J. Magueijo, astro-ph/9809053.
[27] T. Gebbie and G. F. R. Ellis, astro-ph/9804316.
[28] A. Challinor and A. Lasenby, Phys. Rev. D **58**, 023001 (1998).
[29] W. Hu and M. White, Phys. Rev. D **56**, 596 (1997).
[30] S. Veeraraghavan and A. Stebbins, Astrophys. J. **365**, 37 (1990).
[31] M. White and W. Hu, Astron. Astrophys. **321**, 8 (1997).
[32] R. Durrer and M. Kunz, Phys. Rev. D **57**, 3199 (1998).
[33] D. Lyth, Phys. Rev. D **31**, 1792 (1985).
[34] H. Kodama and M. Sasaki, Int. J. Mod. Phys. A **1**, 265 (1986).
[35] S. Bildhauer, T. Buchert, and M. Kasai, Astron. Astrophys. **263**, 23 (1992).
[36] T. Matsubara, Prog. Theor. Phys. **94**, 1151 (1995).
[37] M. S. Turner and M. White, Phys. Rev. D **56**, 4439 (1997).
[38] D. Heath, Mon. Not. R. Astron. Soc. **179**, 351 (1977).
[39] W. Hu and N. Sugiyama, Astrophys. J. **471**, 542 (1996).
[40] E. J. Groth and P. J. E. Peebles, Astron. Astrophys. **41**, 143 (1975).
[41] L. Knox, Phys. Rev. D **52**, 4307 (1995).
[42] R. R. Caldwell and P. J. Steinhardt, Phys. Rev. D **57**, 6057 (1998).
[43] W. Hu, D. Scott, N. Sugiyama, and M. White, Phys. Rev. D **52**, 5498 (1995).
[44] W. Hu, D. J. Eisenstein, M. Tegmark, and M. White, Phys. Rev. D **59**, 023512 (1998).
[45] W. H. Press and E. T. Vishniac, Astrophys. J. **239**, 1 (1980).
[46] J. Silk, Astrophys. J. **151**, 459 (1968).
[47] P. J. E. Peebles and J. T. Yu, Astrophys. J. **162**, 815 (1970).
[48] J. M. Bardeen, J. R. Bond, N. Kaiser, and A. S. Szalay, Astrophys. J. **304**, 15 (1986).
[49] N. Kaiser, Mon. Not. R. Astron. Soc. **202**, 1169 (1983).
[50] D. N. Schramm and M. S. Turner, Rev. Mod. Phys. **70**, 303 (1998).
[51] N. G. Turok, Phys. Rev. D **54**, 3686 (1996); Phys. Rev. Lett. **77**, 4138 (1997).
[52] N. Turok, U.-L. Pen, and U. Seljak, Phys. Rev. D **58**, 023506 (1998).
[53] U. Seljak, U.-L. Pen, and N. Turok, Phys. Rev. Lett. **79**, 1615 (1997).
[54] U. Seljak and M. Zaldarriaga, Astrophys. J. **469**, 437 (1996).
[55] J. P. Ostriker and P. J. Steinhardt, Nature (London) **377**, 600 (1995).
[56] A. R. Liddle, D. H. Lyth, P. T. P. Viana, and M. White, Mon. Not. R. Astron. Soc. **282**, 281 (1996).
[57] R. G. Carlberg, S. L. Morris, H. K. C. Yee, and E. Ellingson, Astrophys. J. Lett. **479**, L19 (1997).

[58] N. A. Bahcall, X. Fan, and R. Cen, Astrophys. J. Lett. **485**, L53 (1997).
[59] A. G. Riess *et al.*, Astron. J. **116**, 1009 (1998).
[60] S. Perlmutter *et al.*, Nature (London) **391**, 51 (1998).
[61] P. Garnavich *et al.*, Astrophys. J. **509**, 74 (1998).
[62] C. H. Lineweaver, Astrophys. J. **505**, 69 (1998).
[63] M. Tegmark, astro-ph/9809201.
[64] T. Chiba, N. Sugiyama, and T. Nakamura, Mon. Not. R. Astron. Soc. **289**, 5 (1997); **301**, 72 (1998).
[65] M. Yu. Khlopov, B. A. Malomed, and Ya. B. Zel'dovich, Mon. Not. R. Astron. Soc. **215**, 575 (1985).