Can superhorizon perturbations drive the acceleration of the Universe?

Éanna É. Flanagan

Laboratory for Elementary Particle Physics, Cornell University, Ithaca, NY 14853, USA.
(Dated: August 31, 2018)

It has recently been suggested that the acceleration of the Universe can be explained as the backreaction effect of superhorizon perturbations using second order perturbation theory. If this mechanism is correct, it should also apply to a hypothetical, gedanken universe in which the subhorizon perturbations are absent. In such a gedanken universe it is possible to compute the deceleration parameter \( q_0 \) measured by comoving observers using local covariant Taylor expansions rather than using second order perturbation theory. The result indicates that second order corrections to \( q_0 \) are present, but shows that if \( q_0 \) is negative then its magnitude is constrained to be less than or of the order of the square of the peculiar velocity on Hubble scales today. We argue that since this quantity is constrained by observations to be small compared to unity, superhorizon perturbations cannot be responsible for the acceleration of the Universe.

PACS numbers: 98.80.Cq

The observed acceleration of the expansion of the Universe \([1, 2]\) is a profound mystery. It is usually explained by positing a new form of matter with negative pressure, so-called dark energy \([4]\), or by a modification of general relativity at large distance scales \([2]\). Recently it has been suggested that the acceleration is instead driven by the backreaction effect of inflation-generated superhorizon perturbations via second order perturbation theory \([5, 6]\). The backreaction of perturbations had earlier been studied in other contexts by Brandenberger and collaborators \([5]\) and others \([8]\).

The basic idea is the following. As is well known, scalar perturbation modes whose wavelengths today are smaller than \( \lambda \sim 8 \text{ Mpc} \) have entered the nonlinear regime and are responsible for galaxies and galaxy clusters, while longer wavelength modes and in particular superhorizon modes (\( \lambda \gtrsim 3 \text{ Gpc} \)) are still in the linear regime today. That is, the fractional density perturbation due to these modes is small compared to unity, so each individual mode evolves with high accuracy according to the linearized equations. Nevertheless the net effect of the backreaction from all the superhorizon modes can still be significant. If we denote by \( \varepsilon = \delta \rho / \rho \sim 10^{-4} \) the fractional density perturbation on Hubble scales today, then one naively expects second order corrections to be of order \( \varepsilon^2 \sim 10^{-8} \) which is negligible. However, a more refined estimate for some particular second order effects obtained from integrating over all the superhorizon modes \([5]\) gives the scaling \( \varepsilon^2 F(k_{\text{min}}, k_{\text{max}}) \), where \( k_{\text{min}} \) and \( k_{\text{max}} \) are the minimum and maximum comoving wavenumbers of the scalar perturbation spectrum integrated over. Here \( F \) is a function for which \( F \to \infty \) as \( k_{\text{min}} \to 0 \). Thus, if the spectrum extends to sufficiently large wavelengths (as would be generated by a sufficiently long period of inflation), second order effects can become significant. In particular, Refs. \([5, 6]\) argue that such effects can give rise to an effective cosmological constant and drive the present acceleration of the Universe.

A potential difficulty with this idea is that it appears to be in conflict with the locality and causality of general relativity. Specifically, consider the finite spacelike hypersurface \( \mathcal{V} \) given by the interior of the intersection of our past lightcone with the spacelike hypersurface of some fixed redshift \( z = z_0 \). We will take \( z_0 = 5 \) say so that the observed supernova used as evidence of the universe’s acceleration are to the future of \( \mathcal{V} \). Then, the initial data within \( \mathcal{V} \) are sufficient to determine all the observations we make; in particular the observable effect of superhorizon perturbations must be encoded in this initial data \([1\)\]. In addition observations strongly constrain this initial data: we know that the geometry of \( \mathcal{V} \) on large scales can be accurately modeled as a Friedman-Robertson-Walker (FRW) background plus fractional perturbations of order \( \varepsilon \sim 10^{-4} \). Thus the issue is whether or not the space of initial data for the gravitational and matter fields on \( \mathcal{V} \) contains enough freedom, given the observational constraints, to mimic the effects of dark energy when no such dark energy is present in the matter stress energy tensor.

In this paper we shall argue that the freedom is insufficient. More precisely, we focus on the deceleration parameter \( q_0 = -a(t) \ddot{a}(t) / \dot{a}(t)^2 \), where \( a(t) \) is the scale factor, whose measured value is \( q_0 \sim -0.5 \). We will argue that (i) The local spatial curvature within \( \mathcal{V} \) is unconstrained and can be altered by second order effects, giving rise to changes in \( q_0 \) that could in principle be of order unity. However this effect cannot change a positive value of \( q_0 \) to a negative value of \( q_0 \). (ii) Non-isotropy and non-homogeneity of the initial data in \( \mathcal{V} \) on large scales can give rise to negative values of \( q_0 \), as suggested by Refs. \([8]\).

\[1\] Note that this initial data does not itself contain superhorizon modes: \( \mathcal{V} \) is a finite spherical region whose comoving radius is \( \sim 1.8H_0^{-1} \) for \( z_0 = 5 \). So for modes which are superhorizon today, \( \mathcal{V} \) extends across only a small fraction of the mode wavelength at \( z = z_0 \).
However the magnitude of this effect is constrained to be of order the square of the velocity perturbation on Hubble scales, which is constrained by observations of low order multipoles of the cosmic microwave background to be small compared to unity.

Method of analysis: If the superhorizon perturbation mechanism for driving acceleration is correct, it should apply not just to our Universe but also to other hypothetical universes. For ease of analysis, we will analyze a fictitious, gedanken universe which differs from ours only in that the perturbation spectrum at early times is modified to suppress the perturbations which are subhorizon today. In this universe the subhorizon perturbations are negligible today, while the superhorizon perturbations are taken to be the same as those used in Refs. [3, 4]. A demonstration that the superhorizon perturbation mechanism does not work in this context is fairly strong evidence that it cannot work in our Universe. The only possible loophole is the possibility that subhorizon perturbations somehow play an important role, which does not seem to be indicated by the analyses of Refs. [5, 6].

In this gedanken universe, the length and time scales over which the gravitational and matter fields are varying are all of order \( H_0^{-1} \) or larger. This allows us to perform an analysis in a local region using Taylor series expansions of the Einstein and hydrodynamic equations, which is much simpler than second order perturbation theory about a FRW background.

We model the matter source by the fluid stress energy tensor \( T_{\alpha\beta} = (\rho + p)u_{\alpha}u_{\beta} + pg_{\alpha\beta} \), where \( \rho \), \( p \), \( u^\alpha \) and \( g_{\alpha\beta} \) are the density, pressure, 4-velocity and metric. Consider a comoving observer at an event \( P \). Such an observer can measure the redshift \( z \) and luminosity distance \( L \) of nearby events, and thus measure the redshift-luminosity distance relation \( L = L(z, \theta, \phi) \). Here \( \theta \) and \( \phi \) are spherical polar coordinates in the observer’s local Lorentz frame. The dependence on these angles arises since we are allowing general local solutions of the Einstein equations; there is no requirement of isotropy. For small \( z \) this relation can be expanded as

\[
L = A(\theta, \phi)z + B(\theta, \phi)z^2 + O(z^3). \tag{1}
\]

We define the Hubble constant \( H_0 \) and deceleration parameter \( q_0 \) as measured by the observer by comparison with the conventional FRW relation \( L = H_0^{-1}z + H_0^{-1}(1 - q_0)z^2/2 + O(z^3) \), as in Ref. [3]:

\[
H_0 \equiv \langle A^{-1} \rangle, \quad q_0 \equiv 1 - 2H_0^{-2}\langle A^{-3}B \rangle. \tag{2}
\]

Here the angular brackets denote an average over the angles \( \theta, \phi \).

The particular prescription [2] for angular averaging is chosen for later convenience. Note that there is no unique prescription; one could for example use the definitions \( H_0^{-1} = \langle A \rangle \) and \( q_0 = 1 - 2H_0\langle B \rangle \). Observations that measure \( H_0 \) and \( q_0 \) typically assume isotropy and therefore effectively angle-average at some stage of the analysis, but the precise nature of the averaging is not usually discussed. We will argue below however that the effect of this ambiguity on the values of \( H_0 \) and \( q_0 \) is small.

Using local Taylor series expansions we can compute \( H_0 \) and \( q_0 \) in terms of the the density, 4-velocity and velocity gradients of the cosmological fluid evaluated at the observer’s location \( P \). We decompose the gradient in the usual way as

\[
\nabla_\alpha u_\beta = \frac{1}{3}\theta(g_{\alpha\beta} + u_\alpha u_\beta) + \sigma_{\alpha\beta} + \omega_{\alpha\beta} - u_\alpha a_\beta, \tag{3}
\]

where \( \theta, \sigma_{\alpha\beta}, \omega_{\alpha\beta} \) and \( a_\alpha \) are the expansion, shear, vorticity and 4-acceleration. For \( H_0 \) we find the well-known result

\[
H_0 = \frac{1}{3}\theta; \tag{4}
\]

the locally measured Hubble constant is just the expansion of the fluid. For \( q_0 \) we obtain

\[
q_0 = \frac{4\pi}{3H_0^2}(\rho + 3p) + \frac{1}{3H_0}\left[a_\alpha a^\alpha + \frac{7}{5}\sigma_{\alpha\beta}\sigma^{\alpha\beta} - \omega_{\alpha\beta}\omega^{\alpha\beta} - 2\nabla_\alpha a^\alpha\right]. \tag{5}
\]

Now to a good approximation in the present epoch we have \( p = 0 \) (assuming cold dark matter and baryons with no dark energy), which implies from \( \nabla_\alpha T^{\alpha\beta} = 0 \) that \( a_\alpha = 0 \). This yields

\[
q_0 = \frac{4\pi}{3H_0^2}p + \frac{1}{3H_0}\left[\frac{7}{5}\sigma_{\alpha\beta}\sigma^{\alpha\beta} - \omega_{\alpha\beta}\omega^{\alpha\beta}\right]. \tag{6}
\]

Discussion: Consider first the first term in Eq. (6). This reduces to the conventional value \( q_0 = 1/2 \) for a flat, matter dominated Universe when \( H_0^2 = 8\pi\rho/3 \). However in the present context the relation \( H_0^2 = 8\pi\rho/3 \) need not be satisfied; \( H_0 \) is instead given by Eq. (4). The deviation of this first term from 1/2 is related to the fact that the local analysis allows spatial curvature. If we define an effective local \( \Omega_k \) by \( \Omega_k = 1 - 8\pi\rho/(3H_0^2) \), then we obtain for the first term \( q_0 = (1-\Omega_k)/2 \), the conventional answer for a Universe with matter and spatial curvature.

The key point about the first term in (6) is that it is positive. Hence this term cannot drive an acceleration.

Consider next the second and third terms in Eq. (6), the squared shear and squared vorticity. These quantities have an unambiguous operational meaning; they can be measured by the observer in her local Lorentz frame. We can estimate the sizes of these terms as

---

2 In addition momentum conservation \( k \approx k_1 + k_2 \) rules out second order corrections to low spatial frequency observables \( (k \lesssim H_0) \) from interactions between very subhorizon modes \( k_1 \gg H_0 \) and superhorizon modes \( k_2 \lesssim H_0 \).
\[ \sigma_{ab} \sigma^{ab}, \omega_{b} \omega^{a} \sim (\delta v)^2/l^2, \] where \( \delta v \) is the typical scale of peculiar velocity (deviation from uniform Hubble flow), and \( l \) is the lengthscale over which the velocity varies. In the present context we have \( l \gtrsim H_0 \), by our assumption that subhorizon modes are negligible, which implies that the contribution from the second and third terms in \( \delta v \) to \( q_0 \) is of order \( \delta q_0 \sim (\delta v)^2 \sim \varepsilon \sim 10^{-4} \). We conclude that it is impossible in this model to achieve the measured value \( q_0 \sim -0.5 \) of the deceleration parameter.

Note that the key difference between our analysis and that of Refs. [5, 6] is one of interpretation. Refs. [5, 6] predict changes to \( q_0 \) that are quadratic in the first order perturbation variables, in agreement with our Eq. \( (8) \). The new information provided by our analysis is that the quadratic terms in Eq. \( (8) \) to \( q_0 \) are constrained to be small since observations constrain the total velocity perturbation \( \delta v \). Thus, while an order-unity renormalization of \( q_0 \) from second order effects is possible in principle, our analysis implies that such a renormalization would also require second order contributions to the fluid velocity that violate observational bounds.

**Details of derivation:** We use the local covariant expansion formalism based on bitensors [11]. We denote by \( x^\alpha \) the coordinates of the event \( P \) where the observer is making observations, and by \( x^\alpha \) the coordinates of an event \( Q \) in the observer’s vicinity. We shall mostly be interested in the case where \( Q \) is on the past lightcone of \( P \). We denote by \( \lambda \) an affine parameter along the geodesic \( x^\alpha = z^\alpha(\lambda) \) that joins \( Q \) and \( P \), chosen so that \( \lambda = 0 \) at \( Q \) and \( \lambda = 1 \) at \( P \). We define Synge’s world function (the squared geodesic interval) via

\[ \sigma(x, x') = \frac{1}{2} \int_0^1 d\\lambda \ g_{\alpha\beta}[z^\alpha(\lambda)] \frac{dz^\alpha}{d\\lambda}(\lambda) \frac{dz^\beta}{d\\lambda}(\lambda). \] (7)

Then \( \sigma_{\alpha\beta}(x, x') = \nabla_{\alpha}\sigma(x, x') \) is the tangent to the geodesic at \( P \). We define \( s(x, x') = -\sigma_{\alpha\beta}(x, x')u^\alpha(x) \); we will use \( s \) as our expansion parameter. We define the vector \( k^\alpha \) by

\[ \sigma_{\alpha\beta}(x, x') = s(x, x')k_\alpha(x); \] (8)

this is a future directed tangent to the geodesic which is normalized so that \( k_\alpha u^\alpha = -1 \) at \( P \). We define \( g_{\alpha\beta}(x, x') \) to be the linear operator of parallel transport from the tangent space at \( Q \) to the tangent space at \( P \), and we define \( \bar{u}^\alpha(x, x') = g^{\alpha\beta}(x, x')u_\beta(x') \). This quantity can be expanded in a local covariant Taylor series as

\[ \bar{u}^\alpha(x, x') = u^\alpha(x) + u^\beta(x)\sigma_{\alpha\beta}(x, x') + \frac{1}{2} u^\beta(x)\sigma_{\alpha\beta}(x, x') \sigma_{\gamma\delta}(x, x') + O(s^4). \] (9)

Using Eq. (B14) of Ref. [11] we can evaluate the coefficients to give \( u^\alpha(x) = -\nabla^\alpha(u^\beta) \) and \( u^\beta(x) = \nabla^\beta(u^\gamma) \).

The measured redshift of the event \( Q \) is given by the ratio of the inner products \( \bar{k} \cdot \bar{u} \) evaluated at \( Q \) and \( P \). Using the definition of \( \bar{u}^\alpha \) this can be written as

\[ 1 + z = \frac{\bar{u}^\alpha k_\alpha}{u^\alpha k_\alpha}. \] (10)

Using the expansion (10) and the definition (11) this can be written as

\[ z = u^\alpha k_\alpha k_\beta s - \frac{1}{2} u^\alpha k_\alpha k_\beta k_\gamma s^2 + O(s^3). \] (11)

Next we evaluate the luminosity distance \( L \). This is defined so that \( 4 \pi L^2 \) is the ratio between an energy emitted per unit time isotropically at \( Q \) and an energy per unit time per unit proper area received at \( P \):

\[ \frac{dE}{dt}(Q) = 4 \pi L^2 \frac{dE}{dt^2 A}(P). \] (12)

These quantities can be evaluated using the geometric optics approximation to the scalar (or Maxwell) wave equation [11]. The stress-tensor for the radiation field is \( T_{\alpha\beta} = A^2 l_\alpha l_\beta \), where \( l_\alpha \) is defined as being the set of null vectors at \( Q \) normalized according to \( l_\alpha u^\alpha = -1 \), and then extended along \( Q \)’s future light cone using the geodesic equation. We define the affine parameter \( \lambda \) by \( \dot{\lambda} = d/d\\lambda \). The normalization conditions for the vectors \( \bar{k} \) and \( \bar{l} \) imply that \( \dot{\bar{l}} = \bar{k}/(1 + z) \), and hence the affine parameters \( s \) and \( \lambda \) are related by \( s = \lambda/(1 + z) \). The amplitude \( A \) satisfies the differential equation

\[ d(ln A)/d\\lambda = -\bar{\theta}/2, \] (13)

where \( \bar{\theta} = \nabla_\alpha l^\alpha \) is the expansion. We choose the normalization of \( A \) so that \( A \approx 1/\lambda \) for \( \lambda \to 0 \) near \( Q \).

The energy flux at \( P \) can now be computed as \( dE/(dt^2 A) = T_{\alpha\beta} u^\alpha u^\beta = A^2 (k_\alpha u^\alpha)^2 = A^2 (1 + z)^2 \). The luminosity at \( Q \) can be evaluated by integrating the energy flux over a small sphere about \( Q \) of radius \( \lambda \); this gives \( dE/dt = A^2 (l_\alpha u^\alpha)^2 (4\pi \lambda^2) = 4\pi \). Combining these results yields \( L = (1 + z)/A \). Using the relation \( s = \lambda/(1 + z) \) we can rewrite this as

\[ L = (1 + z)^2 s A^2 \Delta(x, x')^{-1/2}, \] (14)

where we have defined \( \Delta = A^2 \lambda^2 \). This quantity satisfies \( \Delta \to 1 \) as \( Q \to P \) and also from Eq. (15) satisfies
the differential equation $d\ln \Delta/(d\lambda) = 2/\lambda - \ddot{\theta}$. By comparing with Eq. (32) of Ref. [12], we see that $\Delta(x, x')$ is the van Vleck determinant [10, 12]. Using the expansion $\Delta(x, x') = 1 + O(s^2)$ given in Ref. [12] and combining Eqs. (11) and (14) gives a relation between redshift $q$ and luminosity distance $L$ of the form (11), where the coefficients are

$$A(\theta, \varphi) = \frac{1}{(\nabla^\alpha u^\alpha)(\kappa_\alpha k_\beta)},$$

$$B(\theta, \varphi) = \frac{2}{(\nabla^\alpha u^\alpha)(\kappa_\alpha k_\beta)} + \frac{1}{2[(\nabla^\alpha u^\alpha)(\kappa_\alpha k_\beta)]^3}. \quad (16)$$

We now evaluate the averages over angles. Using the definition $H_0 = \langle A^{-1} \rangle$ together with $\langle k_\alpha k_\beta \rangle = (g_{\alpha\beta} + 4u_\alpha u_\beta)/3$ yields the result (14), using $(\nabla^\alpha u^\alpha)_{\alpha\beta} = 0$. Note that if we use the alternative angle-averaging definition $H_0^{-1} = \langle A \rangle$ we instead obtain $H_0^{-1} = \langle (\theta/3 + \sigma_{ij} n_i n_j + a_i a_j)^{-1} \rangle$, where $k^\alpha = (1, n^\alpha)$ in the local comoving frame at $P$. Evaluating this average treating the shear and acceleration as small compared to the expansion term, we obtain

$$H_0 = \theta/3 - 2\sigma_{ij} \sigma_{ij}/(5\theta) - a_i a_i/\theta + O(\sigma^2/\theta^3) + O(a_i a_i/\theta^3).$$

Thus the different averaging prescriptions give different answers. However the fractional differences are of order $\sigma^2/\theta^2$, which we have argued above is of order $\varepsilon$ and is small.

We now evaluate the angular average of the quantity $A^{-3}B$. We write this as $\langle A^{-3}B \rangle = I/2 + J$, where $I = \langle (\nabla^\alpha u^\alpha k_\alpha k_\beta) \rangle$ and $J = \langle (\nabla^\alpha u^\alpha k_\alpha k_\beta) \rangle^2$. For $I$ we obtain $J = 2\langle (\theta/3 + \sigma_{ij} n_i n_j + a_i a_j)^{-1} \rangle = 2\sigma^2/9 + 4\sigma_{ij} \sigma_{ij}/15 + 2a_i a_i/3$. Using the formula $\langle k_\alpha k_\beta k_\gamma \rangle = (g_{\alpha\beta} + g_{\gamma\beta} + g_{\beta\gamma} a_\alpha)/3 + 2u_\alpha u_\beta u_\gamma$, we obtain for $I$ the formula

$$I = \frac{1}{3} \nabla_\alpha u^\alpha - \frac{1}{3} \theta^2 - \sigma_{\alpha\beta} \sigma_{\alpha\beta} + \frac{1}{3} \omega_{\alpha\beta} \omega_{\alpha\beta} + \frac{5}{3} \omega_{\alpha\beta} a^\alpha a^\beta + 4\pi \rho \pi.$$  

Finally we note that using the alternative angular averaging definition $g_0 = 1 - 2H_0(B)$ would not change our conclusions. This average can be evaluated by expanding the denominators in Eq. (10) treating the last three terms in Eq. (9) as small compared to the expansion term, using the identity $\langle k_\alpha k_\beta k_\gamma k_\delta \rangle = 16u_\alpha u_\beta u_\gamma u_\delta u_\epsilon/3 + 16\sigma_{\beta\gamma} \sigma_{\beta\gamma} u_\epsilon u_\delta/3 + 8\omega_{\beta\gamma} \omega_{\beta\gamma} u_\epsilon u_\delta$, and performing manipulations similar to those used above.

This research was supported in part by NSF grant PHY-0140209. I thank David Chernoff, Saul Teukolsky, Henry Tye and Maxim Perelstein for helpful conversations.

* Electronic address: seet@cornell.edu

[1] A. G. Riess et al., Astron. J. 116, 1009 (1998); S. Perlmutter et al., Astrophys. J. 517, 565 (1999).
[2] C. L. Bennett et al., Astrophys. J. Suppl. 148, 1 (2003). For a review, see P. E. Peebles and B. Ratra, Rev. Mod. Phys. 75, 559 (2003).
[3] See, for example, S.M. Carroll, V. Duvvuri, M. Trodden and M. Turner, Phys. Rev. D 70, 043528 (2004); N. Arkani-Hamed et al., JHEP 0405, 074 (2004).
[4] E. W. Kolb, S. Matearrese, A. Notari and A. Riotto, Phys. Rev. D 71, 023524 (2004); ibid, astro-ph/0410541, ibid, hep-th/0503117.
[5] E. Barausse, S. Matearrese and A. Riotto, astro-ph/0511152.
[6] V.F. Mukhanov, L.R.W. Abramo and R.H. Brandenberger, Phys. Rev. Lett. 78, 1624 (1997); L.R.W. Abramo, astro-ph/9810049.
[7] L.R.W. Abramo, astro-ph/0310265.
[8] S. Räsänen, JCAP 0402, 003 (2004) (also astro-ph/0311257).
[9] Y. Nambu, Phys. Rev. D 62, 104010 (2000); ibid, Phys. Rev. D 65, 104013 (2002); C. Clarkson, Phys. Rev. D 70, 103524 (2004); K. Nakamura, gr-qc/0410241.
[10] K.A. Malik and W. D. Wands, Class. Quant. Grav. 21, L65 (2004); D. Langlois and F. Vernizzi, astro-ph/0503641.
[11] J. W. Moffat, astro-ph/0512110.
[12] J-F. Pascual-Sánchez, Class. Quant. Grav. 17, 4931 (2000).
[13] B. S. DeWitt and R. W. Brehme, Ann. Phys. (N.Y.) 9, 220 (1960). For a recent pedagogical review see E. Poisson, gr-qc/0306052.
[14] W.G. Anderson, E.E. Flanagan and A.C. Ottewill, Phys. Rev. D 71, 024036 (2005).
[15] M. Visser, Phys. Rev. D 47, 2395 (1993).