Recovery of the cosmological peculiar velocity from the density field in the weakly non-linear regime

M. J. Chodorowski,[1] E. L. Łokas,[1] A. Pollo [1] and A. Nusser [2]

[1] *Copernicus Astronomical Center, Bartycka 18, 00–716 Warsaw, Poland*
[2] *Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, 85740 Garching, Germany*

2 October 2018

ABSTRACT
Using third-order perturbation theory, we derive a relation between the mean divergence of the peculiar velocity given density and the density itself. Our calculations assume Gaussian initial conditions and are valid for Gaussian filtering of the evolved density and velocity fields. The mean velocity divergence turns out to be a third-order polynomial in the density contrast. We test the power spectrum dependence of the coefficients of the polynomial for scale-free and standard CDM spectra and find it rather weak. Over scales larger than about $5\,h^{-1}$Mpc, the scatter in the relation is small compared to that introduced by random errors in the observed density and velocity fields. The relation can be useful for recovering the peculiar velocity from the associated density field, and also for non-linear analyses of the anisotropies of structure in redshift surveys.

Key words: cosmology: theory – galaxies: clustering – galaxies: formation – large–scale structure of the Universe

1 INTRODUCTION

In the gravitational instability paradigm for the formation of structure in the Universe, the peculiar motions (deviations from the Hubble flow) of galaxies are tightly related to the mass distribution in the Universe. Quantitative relations between the peculiar velocity field, $\boldsymbol{v}$, and the mass density contrast field, $\delta = \rho/\rho_b - 1$, where ρ_b is the background density, can be obtained from the equations of motion of a collisionless self-gravitating system. In the linear regime, the relation between the density and the velocity fields is

$$\delta(\boldsymbol{x}) = -f(\Omega)^{-1}\nabla \cdot \boldsymbol{v}(\boldsymbol{x})\,, \tag{1}$$

where $f(\Omega) \simeq \Omega^{0.6}$ (e.g. Peebles 1980) and we express distances in units of km s^{-1}. Given an assumed value for Ω, the above relation can be used to reconstruct the mass density field from the large-scale velocity field. The comparison of the reconstructed mass field with the observed large-scale galaxy density field serves as a method for estimating Ω (Dekel et al. 1993) and as a test of the gravitational instability hypothesis (but see Babul et al. 1994). The linear theory relation is applicable only when the density fluctuations are small compared to unity. However, the observed density fluctuations from current redshift surveys (e.g. Fisher et al. 1994) and from the POTENT reconstruction of density fields, slightly exceed the regime of applicability of linear theory. For example, the density contrast in regions like the Great attractor is about unity even when smoothed over scales of the order of 1000 km s^{-1} (Dekel et al. 1993). Fu-

ture redshift surveys and catalogs of peculiar velocities will provide reliable estimates of density and velocity fields on scales where nonlinear effects are certainly not negligible and need to be incorporated in analyzing the data. Various nonlinear relations between the velocity and the associated density field have been developed. One approach is to assume phenomenological parametric forms of these relations, which are to be calibrated with N-body simulations (Dekel 1994, Ganon et al. 1998). Another complementary approach is to analytically derive these relations based on various approximations to nonlinear dynamics. For the purpose of deriving the density from the observed velocities, the POTENT algorithm for example uses a nonlinear approximation based on the Zel'dovich approximation (Nusser et al. 1991). The inverse problem, in which one wishes to recover the velocity from a given density field is yet unsolved in the Zel'dovich approximation as it involves a set of nonlinear differential equations, which do not have analytic solutions. This paper aims at deriving a nonlinear relation which can be used to recover the velocity field from the density field. This is relevant to comparisons of observed velocities of galaxies, using for example the Tully-Fisher relation, with predicted velocity fields from the density field estimated from redshift surveys (Strauss & Davis 1988, Yahil 1988, Kaiser et al. 1991, Hudson 1994, Davis, Nusser & Willick 1996, Willick & Strauss 1998). Moreover, such a relation can be useful for estimating Ω from the distortions of clustering in redshift space (Kaiser 1987, Fisher & Nusser 1996, Taylor & Hamilton 1996, Hamil-

ton 1997) when measured, for example, from the anisotropies of the correlation functions or, equivalently, the power spectra. This is particularly important for measurements of redshift distortions in the Sloan Digital Sky Survey (Gunn & Knapp 1993) and the Anglo-Australian 2dF galaxy survey. Nonlinear analysis of these data might also aim at breaking the degeneracy between Ω and bias (Chodorowski & Łokas 1997; hereafter Paper I, Bernardeau, Chodorowski & Łokas, 1998).

One way to go in order to derive velocity from density is to assume that the divergence of the velocity field at any point in space can be approximated by an expansion in terms of the density contrast at that point. The coefficients of this expansion can then be determined using either N-body simulations (Nusser et al. 1991, Mancinelli et al. 1994) or perturbation theory (Bernardeau 1992) under the assumption of Gaussian initial conditions. The advantage of this approach is that it provides the velocity divergence directly in terms of the density. For irrotational flows, the velocity field can readily be recovered from its divergence given some boundary conditions at large distances. Indeed, if we define[⋆]

$$\theta(\boldsymbol{x}) \equiv -f(\Omega)^{-1} \nabla \cdot \boldsymbol{v}(\boldsymbol{x}), \tag{2}$$

then the velocity field is simply

$$\boldsymbol{v}(\boldsymbol{x}) = \frac{f(\Omega)}{4\pi} \int \mathrm{d}^3 x' \theta(\boldsymbol{x}') \frac{\boldsymbol{x}' - \boldsymbol{x}}{|\boldsymbol{x}' - \boldsymbol{x}|^3} . \tag{3}$$

One caveat to this approach is that the velocity divergence is not determined uniquely by the density field (Chodorowski 1997, Mancinelli & Yahil 1995, Catelan et al. 1995) and the scatter around the derived relation is expected to propagate into errors in the velocity field. However, in the weakly nonlinear regime, the values of $\delta(\boldsymbol{x})$ and $\theta(\boldsymbol{x})$ are still strongly correlated (Bernardeau 1992, Paper I) and the merits of this approach completely overwhelm this caveat given that the problem is not deterministically solved even in the simple Zel'dovich approximation.

In this paper we rigorously compute the mean $\theta(\boldsymbol{x})$ given $\delta(\boldsymbol{x})$, i.e., $\langle\theta\rangle|_\delta$, up to third order in perturbation theory, assuming Gaussian initial conditions. As we shall see later, the derivation is greatly simplified when following Paper I in which the mean of $\delta(\boldsymbol{x})$ given $\theta(\boldsymbol{x})$ was calculated. In Paper I we explicitly calculated the coefficients, a_i, appearing in the expansion

$$\Delta(\theta) \equiv \langle\delta\rangle|_\theta = a_1\theta + a_2(\theta^2 - \sigma_\theta^2) + a_3\theta^3 , \tag{4}$$

where σ_θ^2 is the variance of the velocity divergence field. The numerical values of a_i are found to be in good agreement with the results of N-body simulations by Chodorowski et al. (1998) and by Ganon et al. (1998). The relation (4) allows one to reconstruct a density field from the corresponding velocity field.

But why can we not invert the relation (4) to obtain the $\Theta(\delta) \equiv \langle\theta\rangle|_\delta$ to the relevant order in perturbation theory? The reason is simply that the scatter of δ around the mean value given by (4) introduces, in general, a bias in the estimate of $\Theta(\delta)$ obtained by straightforward inversion.

⋆ Note a slight difference from the commonly used definition, e.g. Bernardeau (1994)

The paper is organized as follows: in Section 2 we describe how the mean of the velocity divergence can be computed from the density using the results of Paper I. In Section 3 we calculate the numerical values of the coefficients entering the formula for $\Theta(\delta)$. Section 4 is devoted to the computation of the scatter in the θ–δ relation. Summary is given in Section 5.

2 THE FORMALISM

We now outline the derivation of the relation between the divergence of the velocity field and the density contrast. Let us express the density contrast as a sum of terms δ_i, each corresponding to the ith order in perturbation theory,

$$\delta = \delta_1 + \delta_2 + \delta_3 + \dots , \tag{5}$$

and, similarly, for the velocity divergence

$$\theta = \theta_1 + \theta_2 + \theta_3 + \dots . \tag{6}$$

In general, the ith order solution is of the order of $(\delta_1)^i$ (Fry 1984, Goroff et al. 1986). We assume here that θ and δ are well approximated by truncating the above expansion at the third order. The linear theory solution mentioned in Section 1 is simply the perturbative expansion truncated at the lowest, i.e. first order term, δ_1 and θ_1. (For explicit forms of higher-order solutions and other details see Paper I.) The coefficients, a_i, appearing in the expansion of the function Δ, equation (4), are combinations of the joint moments of the density and velocity divergence fields. According to Paper I

$$a_1 = 1 + \left[\Sigma_2 + \frac{(S_{3\delta} - S_{3\theta})\, S_{3\theta}}{3} - \frac{\Sigma_4}{2} \right] \sigma^2 , \tag{7}$$

$$a_2 = \frac{S_{3\delta} - S_{3\theta}}{6} , \tag{8}$$

$$a_3 = \frac{\Sigma_4 - (S_{3\delta} - S_{3\theta})\, S_{3\theta}}{6} . \tag{9}$$

In equation (7), σ^2 is the leading, *linear*, contribution to the variance of the density field, $\sigma^2 \equiv \langle\delta_1^2\rangle$. The quantity $S_{3\delta}$ denotes the skewness of the density field,

$$S_{3\delta} = \frac{\langle\delta^3\rangle}{\langle\delta^2\rangle^2} = \frac{3\langle\delta_1^2\delta_2\rangle}{\sigma^4} + \mathcal{O}(\sigma^2) , \tag{10}$$

and $S_{3\theta}$ is the skewness of the velocity divergence field defined in an analogous way. The quantities Σ_2 and Σ_4 are given by

$$\Sigma_2 = \frac{\langle\delta_2\theta_2\rangle_c - \langle\theta_2^2\rangle_c + \langle\delta_1\delta_3\rangle_c - \langle\theta_1\theta_3\rangle_c}{\sigma^4} \tag{11}$$

and

$$\Sigma_4 = \frac{3\langle\delta_1^2\delta_2\theta_2\rangle_c - 3\langle\theta_1^2\theta_2^2\rangle_c + \langle\delta_1^3\delta_3\rangle_c - \langle\theta_1^3\theta_3\rangle_c}{\sigma^6}. \tag{12}$$

In the expressions above, the symbol $\langle\cdot\rangle_c$ stands for the connected (reduced) part of the moments.

In the derivation of the coefficients a_i, no assumption was made about the particular forms of δ_j and θ_j, except the linear theory result $\delta_1 = \theta_1$. Consequently, the inverse relation for the mean of θ given δ, can immediately be written by exchanging the symbols δ and θ in equations (4) and (7)–(12). Thus if we express the function $\Theta \equiv \langle\theta\rangle|_\delta$ in the following form

$$\Theta(\delta) = r_1\delta + r_2(\delta^2 - \sigma_\delta^2) + r_3\delta^3, \tag{13}$$

where $\sigma_\delta^2 \equiv \langle\delta^2\rangle$ is the variance of the density field, then the coefficients r_i can be written as

$$r_1 = 1 + \left[\Sigma_2' + \frac{(S_{3\theta} - S_{3\delta})S_{3\delta}}{3} - \frac{\Sigma_4'}{2}\right]\sigma^2, \tag{14}$$

$$r_2 = \frac{S_{3\theta} - S_{3\delta}}{6}, \tag{15}$$

$$r_3 = \frac{\Sigma_4' - (S_{3\theta} - S_{3\delta})S_{3\delta}}{6}, \tag{16}$$

where

$$\Sigma_2' = \frac{\langle\theta_2\delta_2\rangle_c - \langle\delta_2^2\rangle_c + \langle\theta_1\theta_3\rangle_c - \langle\delta_1\delta_3\rangle_c}{\sigma^4} \tag{17}$$

and

$$\Sigma_4' = \frac{3\langle\theta_1^2\theta_2\delta_2\rangle_c - 3\langle\delta_1^2\delta_2^2\rangle_c + \langle\theta_1^3\theta_3\rangle_c - \langle\delta_1^3\delta_3\rangle_c}{\sigma^6}. \tag{18}$$

Equations (13)–(18) are, strictly speaking, valid only up to third order in σ_δ and for typical values of δ, i.e., away from rare peaks. The spatial average of θ is given by $\int\langle\theta\rangle|_\delta p(\delta)\,d\delta = \int\Theta(\delta)p(\delta)\,d\delta$. The mean of the first two terms in equation (13) is exactly zero. The mean of the third term is $r_3\langle\delta^3\rangle = r_3 S_{3\delta}\sigma_\delta^4 \sim \mathcal{O}(\sigma_\delta^4)$ (see eq. [10]). Hence, our third-order formula (13) fulfills the requirement that the average of the velocity divergence vanishes up to the terms cubic in σ_δ, as expected.

By comparing the coefficients r_i with the expressions (7)-(12) for the coefficients a_i, we see that the only new quantities to be computed are $\langle\delta_2^2\rangle_c$ and $\langle\delta_1^2\delta_2^2\rangle_c$, appearing in Σ_2' and Σ_4' respectively. Therefore, one can express r_i as combinations of a_i and some residuals D_i. We uniquely define the residuals D_i in the following way. Let us approximate the function Θ by the perturbative inversion of the function Δ, equation (4). The approximations of the coefficients r_i obtained in this manner will be denoted by n_i. The exact values of the coefficients are then given by a sum of the approximate values n_i and the residuals D_i,

$$r_i = n_i + D_i. \tag{19}$$

Straightforward calculation of the coefficients n_i yields

$$n_1 = 2 - a_1 - 2a_2^2\sigma^2, \tag{20}$$

$$n_2 = -a_2, \tag{21}$$

$$n_3 = -a_3 + 2a_2^2. \tag{22}$$

We will now transform the coefficients r_i into forms similar to the coefficients n_i. The case of r_2 is trivial:

$$r_2 = \frac{S_{3\theta} - S_{3\delta}}{6} = -\frac{S_{3\delta} - S_{3\theta}}{6} = -a_2, \tag{23}$$

so comparing with equation (21) we see that $D_2 = 0$. The calculation of r_1 and r_3 is more lengthy but straightforward. The transformation of the coefficient r_3 is simpler when we rewrite Σ_4' in a form consisting of explicitly asymmetric and symmetric parts,

$$\Sigma_4' = \frac{S_{4\theta} - S_{4\delta}}{4} - \frac{3\langle\theta_1^2(\delta_2 - \theta_2)^2\rangle_c}{2\sigma^6}. \tag{24}$$

Here, $S_{4\delta}$ is the kurtosis of the density field,

$$S_{4\delta} = \frac{6\langle\delta_1^2\delta_2^2\rangle_c + 4\langle\delta_1^3\delta_3\rangle_c}{\sigma^6} + \mathcal{O}(\sigma^2), \tag{25}$$

and $S_{4\theta}$ is the kurtosis of the velocity divergence field defined in an analogous way. The final result is

$$r_1 = 2 - a_1 - 2a_2^2\sigma^2 + \tilde{D}_1\sigma^2, \tag{26}$$

$$r_2 = -a_2, \tag{27}$$

$$r_3 = -a_3 + 2a_2^2 + D_3, \tag{28}$$

where

$$\tilde{D}_1 = \frac{3\langle\theta_1^2(\delta_2 - \theta_2)^2\rangle_c}{2\sigma^6} - \frac{\langle(\delta_2 - \theta_2)^2\rangle_c}{\sigma^4} - \frac{5(S_{3\delta} - S_{3\theta})^2}{18} \tag{29}$$

and

$$D_3 = \frac{(S_{3\delta} - S_{3\theta})^2}{9} - \frac{\langle\theta_1^2(\delta_2 - \theta_2)^2\rangle_c}{2\sigma^6}. \tag{30}$$

Note that we have introduced $\tilde{D}_1 = D_1/\sigma^2$ since this rescaled parameter, like D_3, does not depend on σ.

The expressions for Δ and Θ up to *second* order in perturbation theory can be derived by neglecting terms of order σ^3. The second order expression for the mean θ given δ is therefore

$$\langle\theta\rangle|_\delta = \delta - a_2(\delta^2 - \sigma_\delta^2) \tag{31}$$

(cf. Bernardeau 1992). This expression is identical to that obtained by direct inversion of the second order expression for $\langle\delta\rangle|_\theta$. Only when third order corrections are included, does the scatter around the mean introduce a bias in the relation when derived by direct inversion. It is interesting to note that although D_1 and D_3 introduce $\mathcal{O}(\sigma^3)$ corrections, they themselves are constructed exclusively from the first and the second order solutions, $\delta_1 = \theta_1$, δ_2 and θ_2.

3 NUMERICAL CALCULATIONS OF r_i

All of the numerical results presented in this section are performed for fields smoothed with a Gaussian filtering window of width R. All the terms appearing in the expressions for r_i have exactly the same mathematical structure as the terms appearing in the expressions for a_i. Therefore, the same arguments that we used in Paper I to argue that the parameters a_i are almost Ω-independent apply as well to the case of r_i. For example, $r_2 = -a_2$, and in Paper I we explicitly derived the Ω-dependence of a_2 and showed it to be extremely weak. Thus, the following calculations of r_i are performed for a flat Universe, but we expect r_i to be robust to the value of Ω.

The calculation of the terms $\langle\delta_2^2\rangle/\sigma^4$ and $\langle\delta_1^2\delta_2^2\rangle/\sigma^6$ which contribute, respectively, to Σ_2' and Σ_4' can be found in Łokas et al. (1995, 1996). The rest of the terms appearing in the expressions for r_i were calculated in Paper I.

3.1 Power law spectra

First we present results for the scale-free spectra of the form

$$P(k) = Ck^n, \quad -3 \le n \le 1. \tag{32}$$

The coefficients r_2 and r_3 are then independent of the normalization of the spectrum and the smoothing scale. The results are straightforward to obtain from equations (15)–(16) and they are given in Table 1 for various spectral indices.

By comparison of r_2 and r_3 given by equations (27)–(28) with the approximate values n_2 and n_3 given by equations

Table 1. The coefficients r_1, r_2, r_3 and n_3 as functions of the spectral index n for scale-free power spectra and Gaussian smoothing

n	r_1	r_2	r_3	n_3
-3.0	$\approx 1 + 0.3\sigma^2$	-0.190	0.0826	0.0826
-2.5	$1 + 0.202\sigma^2$	-0.192	0.0822	0.0822
-2.0	$1 + 0.077\sigma^2$	-0.196	0.0818	0.0821
-1.5	$1 - 0.296\sigma^2$	-0.203	0.0812	0.0822
-1.0	$-$	-0.213	0.0806	0.0835
-0.5	$-$	-0.227	0.0797	0.0865
0	$-$	-0.246	0.0783	0.0928
0.5	$-$	-0.270	0.0756	0.1051
1.0	$-$	-0.301	0.0707	0.1283

(21)–(22) we find that for $n > -3$, the residual $D_2 = 0$ but $D_3 \neq 0$. Only for $n = -3$ we find that r_3 exactly equals to n_3 and $D_3 = 0$. We present n_3 in the last column of Table 1. We see that the approximate values n_3 diverge from the exact ones, r_3, more significantly for higher spectral indices.

In order to obtain r_1 we use equation (14) and get for the unsmoothed fields

$$\Sigma_2' = -\frac{1681}{4410} - h(n) \approx -0.4 \tag{33}$$

where $h(n)$ is the part weakly dependent on n which contributes roughly 10% to the value of Σ_2'. Using this result we estimate the value of the coefficient r_1 in this case to be

$$r_1 \approx 1 + 0.3\sigma^2. \tag{34}$$

The value of the residual $\tilde{D}_1 = D_1/\sigma^2$ given by equation (29) is for the unsmoothed fields independent of n and equal to $-32/2205 = -0.0145$.

When smoothing is applied we are restricted to the spectral indices $n < -1$ as in the case of a_1 (see Paper I). The case of $n = -3$ with smoothing corresponds to the same case with no smoothing (because for such a spectrum the dominant contribution comes from the small wave-numbers at which the window function equals to unity), hence r_1 is given by equation (34). For $n = -2$ the calculations can be performed analytically and we find
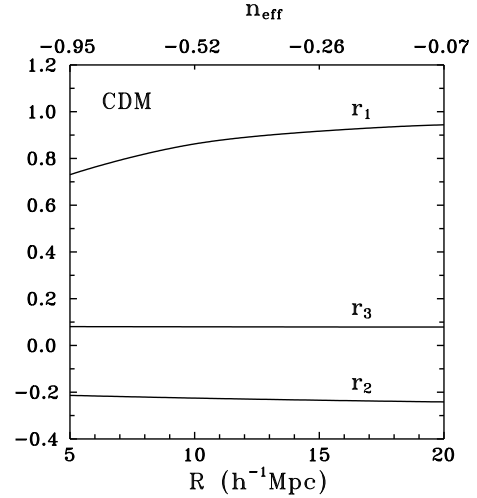
$$\Sigma_2' = -\frac{29}{196}\pi = -0.465 \tag{35}$$

hence

$$r_1 = 1 + 0.0768\sigma^2. \tag{36}$$

For half-integer values of spectral indices we calculated the values of r_1 numerically. The results are given in the second column of Table 1. Comparing equation (26) with equation (20) we obtain the values of $\tilde{D}_1 = D_1/\sigma^2$. We find that in this range of spectral indices the coefficient n_1 is a good approximation of r_1 that is the correction introduced by $\tilde{D}_1$ is very small. Therefore, in Table 1 we do not present the values of n_1.

3.2 The standard CDM spectrum

As an example of a scale-dependent power spectrum we consider the standard CDM (with coefficients given by Efstathiou, Bond & White 1992) normalized so that the linear rms fluctuation in spheres of radius $R = 8h^{-1}$ Mpc is equal

Figure 1. The coefficients r_1, r_2 and r_3 for the standard CDM spectrum normalized to (top hat) $\sigma_8 = 1$ in the weakly nonlinear range of Gaussian smoothing scales. The values of the effective index corresponding to each of the smoothing scales are also plotted.

to unity. Given a_2 the coefficient r_2 is obtained immediately from equation (27).

In calculating r_3 we use the effective index

$$n_{\text{eff}} = -\frac{R}{\sigma^2}\frac{d\sigma^2(R)}{dR} - 3 \tag{37}$$

which measures the slope of the power spectrum at smoothing scale R. Then the values of r_3 at a given scale are found by interpolating the values known for scale-free spectra (Table 1) at the effective index corresponding to that scale.

The values of Σ_2' that contribute to r_1 cannot however be calculated using the effective index method. We therefore calculate the correction to unity in r_1 by numerical integration of the relevant formulae (see Paper I). As in the case of a_1 we benefit from the fact that the slope of the CDM spectrum approaches -3 at large wave-numbers and thus the weakly nonlinear correction to the value of r_1 converges even for scales where the effective spectral index is close to $+1$.

The results for the CDM spectrum in the weakly nonlinear range of scales are presented in Figure 1. In the values of r_1 we incorporated the linear variances σ^2 determined by the normalization to $\sigma_8 = 1$. The Figure also shows the values of the effective index corresponding to each of the smoothing scales.

4   SCATTER IN THE δ–θ RELATION

In Paper I we have computed mean δ given θ, $\langle\delta\rangle|_\theta$. Analogous calculations lead to a formula for the conditional variance, $\sigma^2|_\theta \equiv \langle(\delta - \langle\delta\rangle|_\theta)^2\rangle|_\theta$; we present them in brief in Appendix A. The result is

$$\sigma^2|_\theta = b_0\sigma_\theta^4 + b_2\sigma_\theta^2\theta^2 + \mathcal{O}(\sigma_\theta^5), \tag{38}$$

where

$$b_0 = \frac{\langle(\delta_2 - \theta_2)^2\rangle}{\sigma^4} - \frac{\langle\theta_1^2(\delta_2 - \theta_2)^2\rangle}{2\sigma^6} + \frac{(\Delta S_3)^2}{18} \tag{39}$$

Table 2. The coefficients b_0 and b_2 as functions of the spectral index n for scale-free power spectra and Gaussian smoothing

n	b_0	b_2
-3.0	$\frac{32}{2205} \simeq 0.0145$	0
-2.5	0.0155	$3.61 \cdot 10^{-5}$
-2.0	0.0193	$2.66 \cdot 10^{-4}$
-1.5	0.0288	$1.02 \cdot 10^{-3}$
-1.0	0.0517	$2.86 \cdot 10^{-3}$
-0.5	0.116	$6.77 \cdot 10^{-3}$
0.0	0.378	$1.45 \cdot 10^{-2}$
0.5	–	$2.95 \cdot 10^{-2}$
1.0	–	$5.76 \cdot 10^{-2}$

and

$$b_2 = \frac{\langle \theta_1^2 (\delta_2 - \theta_2)^2 \rangle}{2\sigma^6} - \frac{(\Delta S_3)^2}{9} . \tag{40}$$

Here, $\Delta S_3 \equiv S_{3\delta} - S_{3\theta}$.

In the present paper we have computed mean θ given δ, $\langle \theta \rangle|_\delta$. The conditional variance in this case, $\sigma^2|_\delta \equiv \langle (\theta - \langle \theta \rangle|_\delta)^2 \rangle|_\delta$, may be similarly obtained from equations (38)–(40) by exchanging the symbols δ with θ. Unlike the coefficients of the mean trend, equations (7)–(12), the coefficients b_0 and b_2 are invariants with respect to this exchange. Therefore,

$$\sigma^2|_\delta = b_0 \sigma_\delta^4 + b_2 \sigma_\delta^2 \delta^2 + \mathcal{O}(\sigma_\delta^5), \tag{41}$$

with b_0 and b_2 given by equations (39) and (40) respectively.

The coefficients b_0 and b_2 may be written in terms of the residuals $\tilde{D}_1$ and D_3, equations (29) and (30):

$$b_0 = -\tilde{D}_1 - 2D_3 \tag{42}$$

and

$$b_2 = -D_3 . \tag{43}$$

Thus the presence of the residuals in equations (26)–(28) for the coefficients r_i is a direct consequence of the scatter in the relation between δ and θ. Were the relation entirely deterministic, it would be describable by one function, so then $\Theta = \Delta^{-1}$, $r_i = n_i$ and there would be no residuals. However, since in weakly nonlinear regime the values of pairs (δ, θ) form an elongated set of some scatter, it is not surprising that averaging along different coordinates gives different curves.

As typically $\delta \sim \theta \sim \sigma$, we have $\sigma^2 \delta^2 \sim \sigma^2 \theta^2 \sim \sigma^4$ and both terms entering the formulas (38) or (41) for the conditional variances are of the order of σ^4. The rms value of the scatter in the δ–θ relation – the square root of the conditional variance – is therefore of the order of σ^2. A good measure of the elongation of a two-dimensional set of points (δ, θ) is the ratio of its scatter to a typical value of δ or θ. In our case this ratio is $\sim \sigma$. When $\sigma \to 0$, the ratio thus tends to zero, as expected: in the limit of linear theory the relation between δ and θ is deterministic.

We present the values of b_0 and b_2 as functions of the spectral index n in Table 2. We see that b_0 and b_2 are always positive and the conditional variance is positive-definite, as required.

It is rather astonishing that the formulas for the coefficients b_0 and b_2 for the conditional variance are constructed exclusively from second-order perturbative contributions for δ and θ. In section 2 we have shown that up to second order

$\Theta = \Delta^{-1}$; the residuals D_i appear only when third-order corrections to the mean trend are included. Still, D_i's and b_i's contain only second-order terms. Why? The explanation of this apparent paradox is provided by Chodorowski (1997). In this work, for unsmoothed fields, the weakly non-linear density has been shown to be a local function of the two velocity scalars: the expansion (velocity divergence) and the shear,

$$\delta(\boldsymbol{x}) = \theta(\boldsymbol{x}) + \frac{4}{21} \left[\theta^2(\boldsymbol{x}) - \frac{3}{2} \Sigma^2(\boldsymbol{x}) \right] + \mathcal{O}(\sigma^3) \tag{44}$$

(cf. Mancinelli & Yahil 1995 and Catelan et al. 1995). Here, the shear scalar Σ is

$$\Sigma \equiv (\Sigma_{ij} \Sigma_{ij})^{1/2} , \tag{45}$$

where

$$\Sigma_{ij} \equiv \frac{1}{2} (v_{i,j} + v_{j,i}) - \frac{1}{3} \theta \delta_{ij} \tag{46}$$

and $v_{i,j}$ are velocity derivatives. Thus, the scatter in the δ–θ relation comes from the shear term in equation (44). The mean trend is

$$\langle \delta \rangle|_\theta = \theta + \frac{4}{21} \left[\theta^2 - \frac{3}{2} \langle \Sigma^2 \rangle|_\theta \right] + \mathcal{O}(\sigma^3) . \tag{47}$$

It reproduces up to second order equation (4) since at the lowest order θ and Σ are statistically independent, hence $\langle \Sigma^2 \rangle|_\theta = \langle \Sigma^2 \rangle = (2/3)\sigma^2$ (see Chodorowski 1997 for details). The conditional variance is

$$\begin{aligned} \sigma^2|_\theta &= \frac{4}{49} \left\langle \left(\Sigma^2 - \langle \Sigma^2 \rangle|_\theta + \mathcal{O}(\sigma^3) \right)^2 \right\rangle \Big|_\theta \\ &= \frac{4}{49} \left\langle \left(\Sigma^2 - \langle \Sigma^2 \rangle \right)^2 \right\rangle + \mathcal{O}(\sigma^5) . \end{aligned} \tag{48}$$

The linear term in equations (44) and (47), θ, has cancelled out in the above formula for the conditional variance. Therefore, third-order contributions to equation (44) result in contributions to the conditional variance that are already of the order σ^5. This is in contrast to the formula for weakly nonlinear growth of ordinary variance,

$$\langle \delta^2 \rangle - \langle \delta_1^2 \rangle = \langle \delta_2^2 \rangle + 2\langle \delta_1 \delta_3 \rangle , \tag{49}$$

which does involve a third-order contribution due to the presence of the linear term.

The above consideration is strictly valid only for unsmoothed fields, but the linear terms in the conditional variance will also cancel out for smoothed fields and the lowest order contribution to it will be constructed exclusively from second-order terms.

The ensemble average on the RHS of equation (48) can be performed: the result is

$$\sigma^2|_\theta = \frac{32}{2205} \sigma^4 + \mathcal{O}(\sigma^5) . \tag{50}$$

The case of unsmoothed fields corresponds to the case of smoothed fields with the spectral index $n = -3$. From Table 2 we see that for $n = -3$ the parameters $b_0 = 32/2205$, $b_2 = 0$ and equation (38) coincides with the above formula. Thus, equation (38) is an extension of the result of Chodorowski (1997) for the case of smoothed fields.

In this work we are mostly interested in the conditional variance $\langle (\theta - \langle \theta \rangle|_\delta)^2 \rangle|_\delta$. In Table 2 we see that $b_2 \ll b_0$ always. For mildly nonlinear fields ($\sigma_\delta \lesssim 1$) this implies that

in equation (41) for $\sigma^2|_\delta$ the term quadratic in δ is negligible with respect to the constant one. [The coefficient b_2 is *exactly* zero in the case of a top-hat smoothing (Chodorowski et al. 1998)]. Furthermore, the actual δ-dependence of the conditional variance must originate from the terms that are of order higher than σ_δ^4. In Appendix A we show that the next-order term is $\propto \sigma_\delta^4 \delta$ which is of the order of σ_δ^5; other terms are already $\mathcal{O}(\sigma_\delta^6)$. This is indeed observed in N-body simulations, where the estimated variance can be well described by the formula (Chodorowski et al. 1998)

$$\sigma^2|_\delta^{(NB)} = b_0^{(NB)} \sigma_\delta^4 + b_1^{(NB)} \sigma_\delta^4 \delta, \qquad (51)$$

with $b_1^{(NB)} \simeq 0.8 b_0^{(NB)}$. In principle, the calculation of the coefficient b_1 involves fourth-order perturbation theory. If density is a local function of the velocity divergence and the shear up to *third* order, similar arguments to these presented after equation (44) may be used to argue that the formula for b_1 should be composed from at most third-order perturbative contributions. Even if this is true, however, we find that the moments $\sigma^{-6}\langle\delta_2(\delta_2 - \theta_2)^2\rangle$ and $\sigma^{-6}\langle\theta_1(\delta_2 - \theta_2)(\delta_3 - \theta_3)\rangle$ will enter into this formula, of extra complexity in comparison to any moments computed here. Given all that, in this paper we will not try to predict the value of b_1.

5   SUMMARY

We have derived a local relation between the divergence of the weakly non-linear peculiar velocity field and the corresponding density contrast field. Specifically, we have computed the mean value of the velocity divergence given density up to third order in perturbation theory, assuming Gaussian initial conditions. Our perturbative calculation yields a third-order 'expansion' of the mean value of the divergence in terms of the density contrast. The coefficients of this expansion have been explicitly calculated for scale-free and standard CDM power spectra, for Gaussian smoothing of the fields in question. In the case of CDM, the coefficients depend weakly on the smoothing scale. Moreover, they are expected to depend very weakly on Ω. It is interesting that the value of the linear coefficient in the relation differs from unity – the value predicted in linear theory. The corrective term to this value is proportional to the variance of the density field. The same form of the corrective term is found in the relation between the mean velocity divergence and the density resulting from the Zel'dovich approximation (Chodorowski 1998).

In order to assess the tightness of the relation, we have computed the scatter in the velocity divergence around its mean value given density. We find that, at least up to third order in perturbation theory, the scatter is almost independent of the density contrast. The rms value of the scatter relative to the rms value, σ_θ, of the divergence is approximately $b_0^{1/2}\sigma_\theta$ where $b_0 = 0.015$–0.37 for scale free power spectra with n ranging from -3 to 0. Over scales larger than about $5\,h^{-1}$Mpc, this scatter is small relative to that introduced by observational errors in analyses of redshift surveys and catalogs of peculiar velocities, such as the density–density or velocity–velocity comparisons. The main source of these errors are the sparse and non-uniform sampling of the velocity data and the scatter in the distance indicators (e.g.

the Tully-Fisher relation). As an example, let us consider the *IRAS*–POTENT comparison (Sigad et al. 1998), employing Gaussian smoothing length of $12\,h^{-1}$Mpc. At this scale, σ_θ is approximately 0.3, hence the ratio of the rms value of the scatter around our relation to the rms value of the divergence lies in the range 0.04–0.18. In contrast, the rms value of the observational scatter is approximately 0.21 (estimated from the mock catalogs in the 'standard volume'), thus the observed ratio is as high as 0.7. N-body simulations (Nusser et al. 1991, Mancinelli et al. 1994, Chodorowski et al. 1998; cf. also Sigad et al. 1998) also demonstrate that the scatter around relations of this type is relatively small. The relation and the formalism presented here can be valuable for quantifying the redshift space anisotropies in the non-linear regime. This is particularly important for probing small scales in the planned SDSS and 2dF surveys. We intend to use the relation for estimating the amount of anisotropy of structure in redshift space.

REFERENCES

Babul A., Weinberg D. H., Dekel A., Ostriker J. P., 1994, ApJ, 427, 1
Bernardeau F., 1992, ApJ, 390, L61
Bernardeau F., 1994, ApJ, 433, 1
Bernardeau F., Kofman L., 1995, ApJ, 443, 479
Bernardeau F., Chodorowski M. J., Łokas E. L., 1998, in preparation
Catelan P., Lucchin F., Matarrese S., Moscardini L., 1995, MNRAS, 276, 39
Chodorowski M. J., 1997, MNRAS, 292, 695
Chodorowski M. J., 1998, preprint (astro-ph/9806087)
Chodorowski M. J., Łokas E. L., 1997, MNRAS, 287, 591
Chodorowski M. J., Bernardeau F., Łokas E. L., Stompor R., Kudlicki A., 1998, in preparation
Davis M., Nusser A., Willick J. A., 1996, ApJ, 473, 22
Dekel A., 1994, Annu. Rev. Astron. Astrophys., 32, 371
Dekel A., Bertschinger E., Yahil A., Strauss M., Davis M., Huchra J., 1993, ApJ, 412, 1
Efstathiou G. P., Bond J. R., White S. D. M., 1992, MNRAS, 258, 1P
Fisher K. B., Nusser A., 1996, MNRAS, 279, L1
Fisher K. B., Davis M., Strauss M. A., Yahil A., Huchra J. P., 1994, MNRAS, 267, 927
Fry J. N., 1984, ApJ, 279, 499
Ganon G., Dekel A., Mancinelli P. J., Yahil A., 1998, in preparation
Goroff M. H., Grinstein B., Rey S.-J., Wise M. B., 1986, ApJ, 311, 6
Gunn J. E., Knapp G., 1993, in Soifer B. T., ed, ASP Conf. Ser. Vol. 43, Sky Surveys, Astron. Soc. Pac., San Francisco, p. 267
Hamilton A. J. S., 1997, MNRAS, 289, 285
Hudson M., 1994, MNRAS, 266, 475
Juszkiewicz R., Weinberg D. H., Amsterdamski P., Chodorowski M. J., Bouchet F. R., 1995, ApJ, 442, 39
Kaiser N., 1987, MNRAS, 227, 1
Kaiser N., Efstathiou G., Saunders W., Ellis R., Frenk C., Lawrence A., Rowan-Robinson M., 1991, MNRAS, 252, 1

Longuet-Higgins M. S., 1963, J. Fluid Mech., 17, 459

Łokas E. L., Juszkiewicz R., Weinberg D. H., Bouchet F. R., 1995, MNRAS, 274, 730

Łokas E. L., Juszkiewicz R., Bouchet F. R., Hivon E., 1996, ApJ, 467, 1

Mancinelli P. J., Yahil A., Ganon G., Dekel A., 1994, in: Proceedings of the 9th IAP Astrophysics Meeting "Cosmic velocity fields", ed. F. R. Bouchet and M. Lachièze-Rey, (Gif-sur-Yvette: Editions Frontières), 215

Mancinelli P. J., Yahil A., 1995, ApJ, 452, 75

Nusser A., Dekel A., Bertschinger E., Blumenthal, G. R., 1991, ApJ, 379, 6

Peebles P. J. E., 1980, The Large-scale Structure of the Universe (Princeton: Princeton University Press)

Sigad Y., Eldar A., Dekel A., Strauss M. A., Yahil A., 1998, ApJ, 495, 516

Strauss M. A., Davis M., 1988, in Rubin V., Coyne G., eds, Large-scale motions in the Universe, Princeton Univ. Press, Princeton, p. 255

Taylor A. N., Hamilton A. J. S., 1996, MNRAS, 282, 767

Willick J. A., Strauss M. A., 1998, preprint (astro-ph/9801307)

Yahil A., 1988, in Rubin V., Coyne G., eds, Large-scale motions in the Universe, Princeton Univ. Press, Princeton, p. 219

APPENDIX A: CALCULATION OF THE CONDITIONAL VARIANCE

Here we describe the calculation of the variance $\sigma^2|_\theta = \langle\delta^2\rangle|_\theta - \langle\delta\rangle|_\theta^2$. The derivation is analogous to the derivation of the mean trend, presented in Paper I, therefore we will sketch it in brief. If $p(\delta, \theta)$ is the joint probability distribution function (PDF) for δ and θ, then mean δ^2 given θ is

$$\langle\delta^2\rangle|_\theta = \frac{\int \delta^2 p(\delta, \theta)\, d\delta}{p(\theta)}\,. \tag{A1}$$

The quantity $\mathcal{N} \equiv \int \delta^2 p(\delta, \theta)\, d\delta$ can be expressed as

$$\mathcal{N} = \frac{1}{2\pi}\int e^{-is\theta}\left.\frac{\partial^2}{\partial(it)^2}\Phi(it, is)\right|_{t=0} ds\,, \tag{A2}$$

where Φ is the characteristic function of the joint PDF. It is related to the cumulant generating function, $\mathcal{K}$, by the equation

$$\Phi(it, is) = \exp\left[\mathcal{K}(it, is)\right]. \tag{A3}$$

The cumulants, κ_{mn}, from which $\mathcal{K}$ is constructed,

$$\mathcal{K} = \sum_{\substack{(m,n)\neq(0,0)}}^{\infty} \frac{\kappa_{mn}}{m!n!}(it)^m(is)^n\,, \tag{A4}$$

are given by the *connected* part of the joint moments

$$\kappa_{mn} = \langle\delta^m\theta^n\rangle_c\,. \tag{A5}$$

Using equations (A3) and (A4) we obtain

$$\left.\frac{\partial^2}{\partial(it)^2}\Phi(it, is)\right|_{t=0} = \left[\sum_{n=0}^{\infty}\frac{\kappa_{2n}}{n!}(is)^n + \sum_{n,p=0}^{\infty}\frac{\kappa_{1n}\kappa_{1p}}{n!p!}(is)^{n+p}\right]\exp\left[\sum_{n=1}^{\infty}\frac{\kappa_{0n}}{n!}(is)^n\right]. \tag{A6}$$

By defining $z = \kappa_{02}^{1/2}s$, $\mu = \delta/\sigma_\delta = \delta/\kappa_{20}^{1/2}$, $\nu = \theta/\sigma_\theta = \theta/\kappa_{02}^{1/2}$, and the standard cumulants $\lambda_{mn} = \kappa_{mn}/(\kappa_{20}^{m/2}\kappa_{02}^{n/2})$, we find that

$$\mathcal{N} = \frac{1}{2\pi}\frac{\kappa_{20}}{\kappa_{02}^{1/2}}\int_{-\infty}^{\infty} dz\, e^{-\frac{1}{2}(z^2+2i\nu z)}\left[\sum_{n=0}^{\infty}\frac{\lambda_{2n}}{n!}(iz)^n + \sum_{n,p=1}^{\infty}\frac{\lambda_{1n}\lambda_{1p}}{n!p!}(iz)^{n+p}\right]\exp\left[\sum_{n=3}^{\infty}\frac{\lambda_{0n}}{n!}(iz)^n\right]. \tag{A7}$$

In weakly nonlinear regime the standard cumulants obey the following scaling hierarchy (Fry 1984, Bernardeau 1992)

$$\lambda_{mn} = S_{mn}\sigma^{m+n-2} + \mathcal{O}(\sigma^{m+n}) \tag{A8}$$

where σ is the linear variance of δ or, equivalently, of θ (recall that at linear order $\delta = \theta$). The series in equation (A7) are thus power series in a small parameter σ and we can truncate them at some order p neglecting contributions which are of order $> \sigma^p$. The leading-order formula for the conditional variance is obtained by keeping the terms up to the order of σ^2. Integrating the resulting expression yields

$$\begin{aligned}
\mathcal{N} =\ & \frac{1}{\sqrt{2\pi}}\frac{\kappa_{20}}{\kappa_{02}^{1/2}}e^{-\frac{1}{2}\nu^2} \times \\
& \left[H_0(\nu) + \lambda_{11}^2 H_2(\nu) + \lambda_{21}H_1(\nu) + \left(\lambda_{12} + \frac{\lambda_{03}}{6}\right)H_3(\nu) + \frac{\lambda_{03}}{6}H_5(\nu) + \right. \\
& \frac{\lambda_{22}}{2}H_2(\nu) + \left(\frac{\lambda_{13}}{3} + \frac{\lambda_{04}}{24} + \frac{\lambda_{21}\lambda_{03}}{6} + \frac{\lambda_{12}^2}{4}\right)H_4(\nu) + \\
& \left.\left(\frac{\lambda_{04}}{24} + \frac{\lambda_{12}\lambda_{03}}{6} + \frac{\lambda_{03}^2}{72}\right)H_6(\nu) + \frac{\lambda_{03}^2}{72}H_8(\nu)\right],
\end{aligned} \tag{A9}$$

where H_n are the n-th order Hermite polynomials.

We now turn to calculating the PDF $p(\theta)$ in equation (A1). Computing it in a similar way as $\mathcal{N}$ we *rederive* the so-called Edgeworth expansion for the variable ν (Longuet-Higgins 1963, Bernardeau & Kofman 1995, Juszkiewicz et al. 1995),

$$p(\nu) = \frac{1}{\sqrt{2\pi}}\, e^{-\nu^2/2}\left[1 + \frac{1}{6}\lambda_{03}H_3(\nu) + \frac{1}{24}\lambda_{04}H_4(\nu) + \frac{1}{72}\lambda_{03}^2 H_6(\nu)\right]\,. \tag{A10}$$

From equations (A9) and (A10), after some algebra we obtain

$$\langle\mu^2\rangle|_\nu = \mathcal{S}_0(\nu) + \mathcal{S}_1(\nu) + \mathcal{S}_2(\nu)\,, \tag{A11}$$

where

$$\mathcal{S}_0(\nu) = 1 - \lambda_{11}^2 + \lambda_{11}^2 \nu^2 \,, \tag{A12}$$

$$\mathcal{S}_1(\nu) = (\lambda_{21} - 3\lambda_{12} + 2\lambda_{03})\nu + (\lambda_{12} - \lambda_{03})\nu^3 \,, \tag{A13}$$

and

$$
\begin{aligned}
\mathcal{S}_2(\nu) \;=\; & -\frac{\lambda_{22} - 2\lambda_{13} + \lambda_{04}}{2} + \frac{2\lambda_{21}\lambda_{03} + 3\lambda_{12}^2 - 10\lambda_{12}\lambda_{03} + 5\lambda_{03}^2}{4} + \\
& \left(\frac{\lambda_{22} - 4\lambda_{13} + 3\lambda_{04}}{2} + \frac{-\lambda_{21}\lambda_{03} - 3\lambda_{12}^2 + 12\lambda_{12}\lambda_{03} - 8\lambda_{03}^2}{2} \right) \nu^2 + \\
& \left(\frac{\lambda_{13} - \lambda_{04}}{3} + \frac{\lambda_{12}^2 - 6\lambda_{12}\lambda_{03} + 5\lambda_{03}^2}{4} \right) \nu^4 \,.
\end{aligned}
\tag{A14}
$$

Note that $\mathcal{S}_k(\nu)$ are of the order of σ^k. To obtain $\langle \mu^2 \rangle|_\nu - \langle \mu \rangle|_\nu^2$ we use the expression for $\langle \mu \rangle|_\nu$ given explicitly in Paper I. The result is

$$
\begin{aligned}
\langle \mu^2 \rangle|_\nu - \langle \mu \rangle|_\nu^2 \;=\; & 1 - \lambda_{11}^2 - \frac{\lambda_{22} - 2\lambda_{13} + \lambda_{04}}{2} + \frac{\lambda_{21}\lambda_{03} + \lambda_{12}^2 - 4\lambda_{12}\lambda_{03} + 2\lambda_{03}^2}{2} + \\
& (\lambda_{21} - 2\lambda_{12} + \lambda_{03})\nu + \\
& \left(\frac{\lambda_{22} - 2\lambda_{13} + \lambda_{04}}{2} - \frac{\lambda_{21}\lambda_{03} + 2\lambda_{12}^2 - 6\lambda_{12}\lambda_{03} + 3\lambda_{03}^2}{2} \right) \nu^2 \,.
\end{aligned}
\tag{A15}
$$

Were the variables identical (to all orders), $\mu = \nu$, then the cumulants would be $\lambda_{11} = 1$, $\lambda_{mn} = \lambda_{(m+n)0}$, and the conditional variance would be zero, as expected. Returning to the 'physical' variables δ and θ, using the leading-order expressions for the standard cumulants, like

$$\lambda_{11}^2 = 1 - \frac{\langle (\delta_2 - \theta_2)^2 \rangle}{\sigma^2} + \mathcal{O}(\sigma^4) \,, \tag{A16}$$

and recalling that $\sigma_\theta = \sigma + \mathcal{O}(\sigma^3)$ we finally obtain

$$\langle \delta^2 \rangle|_\theta - \langle \delta \rangle|_\theta^2 = b_0 \sigma_\theta^4 + b_2 \sigma_\theta^2 \theta^2 + \mathcal{O}(\sigma_\theta^5) \,, \tag{A17}$$

where

$$b_0 = \frac{\langle (\delta_2 - \theta_2)^2 \rangle}{\sigma^4} - \frac{\langle \theta_1^2 (\delta_2 - \theta_2)^2 \rangle}{2\sigma^6} + \frac{(\Delta S_3)^2}{18} \tag{A18}$$

and

$$b_2 = \frac{\langle \theta_1^2 (\delta_2 - \theta_2)^2 \rangle}{2\sigma^6} - \frac{(\Delta S_3)^2}{9} \,. \tag{A19}$$

In equation (A15) higher-order corrections to the values of cumulants yield corrections to formula (A17) which are already of the order of σ_θ^6. The only exception is the term linear in $\nu = \theta/\sigma_\theta$ which yields a contribution of the order of σ_θ^5. Therefore, for mildly nonlinear fields, the first departure from the leading-order formula (A17) for the conditional variance $\sigma^2|_\theta$ should have the form $\propto \sigma_\theta^4 \theta$. Analogously, the expression for the conditional variance in the reverse case, $\sigma^2|_\delta = \langle \theta^2 \rangle|_\delta - \langle \theta \rangle|_\delta^2$, should contain a higher-order term of the form $\propto \sigma_\delta^4 \delta$.

This paper has been produced using the Royal Astronomical Society/Blackwell Science LaTeX style file.