MPI-PhT/94-61

REAL-SPACE COSMIC FIELDS FROM REDSHIFT-SPACE DISTRIBUTIONS: A GREEN FUNCTION APPROACH[1]

Max Tegmark

Max-Planck-Institut für Physik, Föhringer Ring 6, D-80805 München
max@mppmu.mpg.de

B. C. Bromley

Theoretical Astrophysics, MS B288, Los Alamos National Laboratory,
Los Alamos, NM 87545
bromley@eagle.lanl.gov

Abstract

We present a new method for reconstructing the cosmological density, peculiar velocity and peculiar gravitational potential on large scales from redshift data. We remove the distorting effects of line-of-sight peculiar motions by using the linear theory of gravitational instability, in which the potential is the solution to a linear partial differential equation first derived by Kaiser. We solve this equation by deriving its Green function; the Green functions for the peculiar velocity and the real-space density follow directly. Reconstruction of the cosmic fields is thus reduced to integration over redshift space with the appropriate Green function kernel. Our algorithm has a single input parameter, $\beta \equiv \Omega^{0.6}/b$, where b is the linear bias factor and Ω is the cosmological density parameter. We discuss the virtues of this method for error control, for estimating β, and for constraining the bias mechanism.

[1]Published in *ApJ*, **453**, 533 (November 10, 1995). Submitted September 20, 1994. Available from *http://astro.berkeley.edu/~max/zspace.html* (faster from the U.S.) and *http://www.mpa-garching.mpg.de/~max/zspace.html* (faster from Europe).

1 Introduction

Redshift surveys are testimony to the richness of structure in the universe on large scales (see, for instance, Giovanelli & Haynes 1991). The observed mass distribution in "redshift space", the space defined by two sky coordinates and a radial recession velocity, depends on both the real-space mass density and the peculiar velocity field, each of which are of great importance in cosmology. The purpose of this paper is to present a new method for evaluating both the density and velocity fields, as well as the peculiar gravitational potential, directly from redshift data.

While quantifying the redshift-space distortions is difficult in general, Kaiser (1987) has solved the problem for large scales ($> 10\,\mathrm{Mpc}$) in the linear regime of gravitational instability theory (Peebles 1993, §5): The effect of linear peculiar flows on the observed redshift-space density field is, loosely speaking, to enhance the amplitude of fluctuation modes along the observer's line of sight by a factor of

$$\beta = \Omega^{0.6}/b\,, \tag{1}$$

where the parameter b admits the possibility that there is a linear bias in the fluctuations of luminous matter relative to the total mass distribution. This dependence on Ω has been exploited to estimate the cosmic mass density from redshift data (Hamilton 1992, 1993; Fisher, Scharf & Lahav 1994; Cole, Fisher & Weinberg 1994; Bromley 1994). For the purposes of removing redshift space distortions, we will assume that β is known or that it can be estimated in some self-consistent way from the data.

The problem of reconstructing cosmic fields from redshift data has attracted considerable interest in the last few years. Yahil et al. (1991), Taylor & Rowan-Robinson (1993) and Gramann, Cen & Gott (1994) have implemented iterative algorithms for real-space density and peculiar velocity reconstructions in the linear or quasi-linear regimes. In rough terms, the iterative strategy is to assume that the redshift data approximate the real-space density, calculate the peculiar flows, then correct the density accordingly. The procedure is repeated until the real-space distribution and the velocity field converge to a solution that is consistent with the observed redshifts. A different strategy has been to expand all fields in spherical harmonics, and then estimate each multipole individually (Fisher, Scharf & Lahav 1994; Nusser & Davis 1994).

Here we start with Kaiser's equation for redshift space distortions and solve for the Green functions that give the real-space density, the peculiar velocity field, and the peculiar gravitational potential. Thus, the reconstruction algorithm simply involves integration over the redshift-space density with a Green function kernel. The method is designed to hold on scales in the linear regime under the assumption that light traces mass or that luminous matter is a linearly biased sample of the total density field. The technique is valid for any survey geometry. In particular, we do not make Kaiser's "plane-parallel" approximation wherein the redshift survey is assumed to lie in a compact region at a large distance from the observer (see Zaroubi & Hoffman 1994).

The Green function method has distinct advantages over previous techniques: it is a noniterative, linear algorithm applicable to any survey geometry, and the work is done entirely in real space, not in the Fourier domain or with spherical harmonic coefficients. Hence the error bars due to edge effects can be calculated in a straightforward and intuitive way.

This paper is organized as follows. After reviewing Kaiser's result and establishing some notation in Section 2, we derive the Green function for the peculiar gravitational potential in Section 3, from which the kernels for peculiar velocity and real-space density follow directly. In Section 4, we specialize to the simple case known as the "distant survey approximation". In Section 5, we discuss how the technique can be applied to real-world situations where one faces nonlinear clustering, shot noise and finite volume. In Section 6, we give a numerical example of density reconstruction from artificial redshift-space data. We conclude by summarizing possible applications of the method for extracting Ω and constraining the biasing mechanism.

2 The partial differential equation

The first step is to obtain an expression, analogous to a Poisson equation, which relates the real-space peculiar gravitational potential ϕ to the density fluctuation field δ_S in redshift space (we denote redshift-space quantities with a subscript-S while subscript-R indicates real space). This was first done by Kaiser (1987), but we give a brief derivation here for completeness.

The fluctuation field is defined in the usual way,

$$\delta = \left[\rho(\mathbf{r}) - \overline{\rho}\right]/\overline{\rho}\,, \tag{2}$$

where ρ is the mass density and $\overline{\rho}$ is its mean. Here, the density fields are associated with luminous matter, and we assume that fluctuations in light are proportional to fluctuations in total mass on the scales of interest. We use time units such that the Hubble constant $H_0 = 1$.

Denoting the radial part of the peculiar velocity field $\mathbf{v}$ at a real-space position $\mathbf{r}$ by

$$U(\mathbf{r}) \equiv \mathbf{v}(\mathbf{r}) \cdot \widehat{\mathbf{r}}\,, \tag{3}$$

and the radial peculiar velocity in the observer's frame by

$$\Delta U(\mathbf{r}) \equiv [\mathbf{v}(\mathbf{r}) - \mathbf{v}(\mathbf{0})] \cdot \widehat{\mathbf{r}}\,, \tag{4}$$

the apparent position $\mathbf{s}$ of a galaxy in redshift space is

$$\mathbf{s}(\mathbf{r}) \equiv \mathbf{r} + \Delta U(\mathbf{r})\widehat{\mathbf{r}}\,. \tag{5}$$

As long as the perturbations are sufficiently small that this mapping is one-to-one, the overdensities in redshift space and in real space are related by

$$(1+\delta_R)d^3r = (1+\delta_S)d^3s\,. \tag{6}$$

Evaluating the Jacobian of the mapping (5) yields

$$d^3s = \left(1 + \frac{\Delta U}{r}\right)^2 (1 + \partial_r U)d^3r\,, \tag{7}$$

where ∂_r denotes the radial derivative $\partial/\partial r$. Thus, to first order in $\mathbf{v}$ and δ, and in the absence of selection effects and survey limits, the equation relating the densities in redshift space and in real space is

$$\delta_S = \delta_R - \partial_r U - 2\frac{U}{r} + 2\frac{\mathbf{v}(0)\cdot\widehat{\mathbf{r}}}{r}\,. \tag{8}$$

We can express the first three terms on the right hand side in terms of the peculiar gravitational potential, ϕ. The density δ_R and ϕ are related through the Poisson equation,

$$\nabla^2\phi = A\delta_R, \tag{9}$$

where (cf. Peebles 1993, eq. [5.107]) we have introduced the constant

$$A = 4\pi G\rho_0 = \frac{3}{2}\Omega. \tag{10}$$

In linear theory, the peculiar velocity is simply proportional to the gravitational force field, and is given by (cf. Peebles 1993, eq. [5.117])

$$\mathbf{v} = -\frac{\beta}{A}\nabla\phi. \tag{11}$$

Thus the velocity terms are

$$\partial_r U = -\frac{\beta}{A}\partial_r^2\phi\,; \qquad \frac{U}{r} = -\frac{\beta}{Ar}\partial_r\phi\,. \tag{12}$$

Substituting this into equation (8) yields what we will refer to as the *Kaiser equation*,

$$D\phi = A\delta'_S\,, \tag{13}$$

where the differential operator D is defined as

$$D \equiv \nabla^2 + \beta\partial_r^2 + 2\beta\frac{1}{r}\partial_r\,. \tag{14}$$

Here we have eliminated the last term in equation (8) by defining δ'_S as the redshift space density observed in the cosmic rest frame, the frame relative to which $\mathbf{v}$ is defined. From the observed dipole anisotropy in the microwave background radiation, we know that the peculiar velocity of our solar system is (Smoot *et al.* 1992)

$$\mathbf{v}(0) = 370 \pm 10 \text{ km/s} \tag{15}$$

in the direction $\alpha = 11.2^h$, $\delta = -7°$. The cosmic variance in this dipole is merely a few km/s, *i.e.*, considerably smaller than the errors quoted above. δ'_S can be directly computed from the redshift-space density we observe, δ_S, either by using the first order correction derived above,

$$\delta'_S(\mathbf{r}) = \delta_S(\mathbf{r}) - 2\frac{\mathbf{v}(0)\cdot\widehat{\mathbf{r}}}{r}, \tag{16}$$

or by simply adjusting the observed redshifts (radial galaxy positions):

$$s' = s + \mathbf{v}(0)\cdot\widehat{\mathbf{r}}. \tag{17}$$

In what follows, we will drop the prime and let δ_S denote the corrected field. It should be noted that this distinction is of only marginal importance. The last two terms in equation (8), are usually negligible, and have been dropped by most authors. We leave them in because they preserve the self-adjointness of the operator D. It turns out that the approximation of dropping them gives a more complicated solution, not a simpler one.

3 The Green functions

We now solve the Kaiser equation (13). Comparing the differential operator D with the expression for ∇^2 in spherical coordinates shows that it is simply the Laplacian with the radial part amplified by a factor $(1+\beta)$. It is self-adjoint for all β, which means that its eigenfunctions are orthogonal, and that expanding the fields in these functions (spherical Bessel functions of non-integer order) provides a straightforward way of solving the equation. A related formal solution is given by Nusser & Davis (1994), but for a slightly different operator yielding the solution in redshift space, not real space. Here we will instead give a more direct solution, in real space, which renders orthogonal expansions unnecessary.

By the linearity of the problem, we can always write down a formal solution as

$$\phi(\mathbf{r}) = A\int G_\phi(\mathbf{r},\mathbf{r}')\delta_S(\mathbf{r}')d^3r', \tag{18}$$

where the Green function (integral kernel) G_ϕ satisfies

$$(DG_\phi)(\mathbf{r},\mathbf{r}') = \delta^3(\mathbf{r}'-\mathbf{r}) \tag{19}$$

and the Dirac delta function is distinguished from the fluctuation fields δ_R and δ_S by the absence of a subscript. The constant A in Eq. (18) was inserted to allow us to think of G_ϕ as the operator inverse of D. Although G_ϕ is a function of six variables, there are in fact only two quantities upon which the dependence is non-trivial. Since the Kaiser equation is invariant under rotation and rescaling, it is easy to see that we can write

$$G_\phi(\mathbf{r},\mathbf{r}') = \frac{g\left(\mu, r/r'\right)}{\sqrt{rr'}} \tag{20}$$

for some function g, with the angle cosine $\mu \equiv \widehat{\mathbf{r}}\cdot\widehat{\mathbf{r}}'$. Let us utilize this property by expanding G_ϕ in Legendre polynomials as

$$G_\phi(\mathbf{r},\mathbf{r}') = \frac{1}{\sqrt{rr'}}\sum_{\ell=0}^{\infty}(2\ell+1)P_\ell(\mu)a_\ell(r/r'). \tag{21}$$

The Kaiser equation implies that the radial functions $a_\ell(x)$ satisfy

$$a_\ell'' + \frac{1}{x}a_\ell' - \frac{1}{x^2}\left[\frac{1}{4}+\frac{l(l+1)}{1+\beta}\right]a_\ell = \frac{\delta(x-1)}{4\pi(1+\beta)}. \tag{22}$$

Requiring a_ℓ to be continuous and imposing the boundary conditions that $a_\ell(x)$ remain finite both as $x \to 0$ and as $x \to \infty$, this equation has the solution

$$a_\ell(x) = -\frac{x^{\pm\alpha_\ell}}{8\pi(1+\beta)\alpha_\ell}, \tag{23}$$

where the exponent is $+\alpha_\ell$ if $x \leq 1$ and $-\alpha_\ell$ if $x \geq 1$, with

$$\alpha_\ell = \frac{1}{2}\sqrt{1 + \frac{4\ell(\ell+1)}{1+\beta}}\,. \tag{24}$$

Substitution back into equation (21) gives the solution

$$G_\phi(\mathbf{r},\mathbf{r}') = -\frac{1}{8\pi(1+\beta)\sqrt{rr'}}\sum_{\ell=0}^{\infty}\left(\frac{2\ell+1}{\alpha_\ell}\right)P_\ell(\mu)\left(\frac{r_<}{r_>}\right)^{\alpha_\ell}, \tag{25}$$

where $r_<$ and $r_>$ denote the smaller and larger of the numbers r and r', respectively. If we let $\beta \to 0$, the Kaiser equation approaches the Poisson equation, $(2\ell+1)/\alpha_\ell \to 2$ and equation (25) reduces to the familiar

$$G_\phi(\mathbf{r},\mathbf{r}') = -\frac{1}{4\pi|\mathbf{r}'-\mathbf{r}|}. \tag{26}$$

We now consider the peculiar velocity field $\mathbf{v}$, which according to equation (11) is proportional to the gradient of the peculiar gravitational potential ϕ. Rather than computing $\nabla\phi$ numerically, it is convenient to define a Green function $\mathbf{G_v}$ expressly for the velocity field which satisfies

$$\mathbf{v}(\mathbf{r}) = \int \mathbf{G_v}(\mathbf{r},\mathbf{r}')\delta_S(\mathbf{r}')d^3r'\,, \tag{27}$$

where, evidently,

$$\mathbf{G_v} = \beta\nabla G_\phi\,. \tag{28}$$

Taking the gradient of equation (25) with respect to $\mathbf{r}$ gives

$$\begin{aligned}\mathbf{G_v}(\mathbf{r},\mathbf{r}') \;&=\; -\frac{\beta}{8\pi(1+\beta)r'^2}\sum_{\ell=0}^{\infty}\left(\frac{2\ell+1}{\alpha_\ell}\right)\left(\frac{r}{r'}\right)^{\gamma_\ell-1} \\ &\times\; \{\gamma_\ell P_\ell(\mu)\widehat{e}_r + \ell\,[\mu P_\ell(\mu) - P_{\ell-1}(\mu)]\,\widehat{e}_\theta\}\,,\end{aligned} \tag{29}$$

where $\gamma_\ell \equiv -1/2+\alpha_l$ if $r < r'$ and $\gamma_\ell \equiv -1/2-\alpha_l$ if $r > r'$. The coordinate θ refers to a spherical coordinate system with the zenith in the $\mathbf{r}'$-direction, *i.e.*, where $\widehat{\mathbf{r}}\cdot\widehat{\mathbf{r}}' = \cos\theta$.

For completeness, we finally construct the Green function for the real-space density, G_δ, such that

$$\delta_R(\mathbf{r}) = \int G_\delta(\mathbf{r},\mathbf{r}')\delta_S(\mathbf{r}')d^3r' . \tag{30}$$

Since the Poisson equation (9) gives

$$G_\delta = \nabla^2 G_\phi , \tag{31}$$

taking the Laplacian of equation (25) yields

$$\begin{aligned} G_\delta(\mathbf{r},\mathbf{r}') &= \frac{1}{8\pi(1+\beta)r'^3}\sum_{\ell=0}^{\infty}\left(\frac{2\ell+1}{\alpha_\ell}\right)[\gamma(\gamma+1)-\ell(\ell+1)]P_\ell(\mu)\left(\frac{r}{r'}\right)^{\gamma-2} \\ &+ \left(\frac{3+2\beta}{3+3\beta}\right)\delta^3(\mathbf{r}'-\mathbf{r}) , \end{aligned} \tag{32}$$

where the delta-function term represents the full contribution of G_δ at $\mathbf{r}=\mathbf{r}'$. As expected,

$$G_\delta(\mathbf{r},\mathbf{r}') \to \delta^3(\mathbf{r}'-\mathbf{r}) \tag{33}$$

in the limit $\beta \to 0$. Contour plots of G_δ and the Green function giving the potential, with $\beta = 1$, appear in Figure 1. Note that $G_\phi(\mathbf{r},\mathbf{r}') = G_\phi(\mathbf{r}',\mathbf{r})$, since D^{-1} is self-adjoint, whereas $\mathbf{G_v}$ and G_δ lack this symmetry.

4 The distant survey approximation

An interesting limit occurs for a distant-survey (Kaiser 1987, Zaroubi & Hoffman 1994) in which $\mathbf{r} \approx \mathbf{r}'$ or $|\mathbf{r}'-\mathbf{r}| \ll r$. For definiteness, let us choose $\mathbf{r}' = \hat{\mathbf{z}}$, the unit vector in the z- direction. Then $\partial_r \approx \partial_z$, the radial part of the Laplacian $\approx \partial_z^2$, and the Kaiser equation reduces to

$$\left[\partial_x^2 + \partial_y^2 + (1+\beta)\partial_z^2\right]\phi = A\delta_s. \tag{34}$$

A simple rescaling of the z-axis transforms the above expression into the Poisson equation. This fact and the application of the gradient and Laplacian operators to the Green function for the potential immediately give us the three Green functions in the distant-survey limit: defining

$$\mathbf{\Delta r} \equiv \mathbf{r} - \mathbf{r}' \equiv (\sin\vartheta\cos\phi, \sin\vartheta\sin\phi, \cos\vartheta)\Delta r \tag{35}$$

and $\mu_l \equiv \cos(\vartheta)$, with ϑ being the angle between $\mathbf{\Delta r}$ and the line-of-sight (the z-axis, parallel to both $\mathbf{r}$ and $\mathbf{r}'$ in this limit),

$$G_\phi(\mathbf{r},\mathbf{r}') = -\frac{1}{4\pi\Delta r}\left(1+\beta-\beta\mu_l^2\right)^{-1/2}; \tag{36}$$

$$\mathbf{G_v}(\mathbf{r},\mathbf{r}') = \frac{1}{4\pi\Delta r^2}\left(1+\beta-\beta\mu_l^2\right)^{-1/2}\left(\widehat{e}_r + \frac{\beta\mu_l\sqrt{1-\mu_l^2}}{1+\beta-\beta\mu_l^2}\widehat{e}_\vartheta\right); \tag{37}$$

$$G_\delta(\mathbf{r},\mathbf{r}') = \frac{(3+\beta)\mu_l^2-1-\beta}{4\pi\Delta r^3\left(1+\beta-\beta\mu_l^2\right)^{5/2}} + \frac{3+2\beta}{3+3\beta}\delta^3(\mathbf{r}'-\mathbf{r}). \tag{38}$$

Thus G_δ vanishes on the cone given by

$$|z'-z| = \left(\frac{1+\beta}{2}\right)^{1/2}\sqrt{(x'-x)^2+(y'-y)^2}, \tag{39}$$

takes positive values inside this cone and takes negative values outside of it.

5 The real world: nonlinearity, shot noise and finite volume

The above treatment applied to an idealized world were the Kaiser equation was strictly valid, and where we could measure the redshift space density field δ_S accurately throughout all space. Alas, this is not the world in which we live. Below, we discuss the following three sources of error:

1. The Kaiser equation, which is exact to first order, is only accurate when $\delta \ll 1$ and thus breaks down on small scales where nonlinear evolution has become important.
2. Our galaxy surveys do not measure the continuous field δ_R but the discrete galaxy distribution, so our estimates of δ_S are contaminated by Poissonian "shot noise".
3. Our galaxy surveys sample only a finite volume, which forces us to truncate our volume integrals and accept truncation errors in our reconstructed fields.

There are of course many more potential sources of error, such as nonlinear biasing and flux errors to mention two, but we will restrict our discussion to the three listed problems, which are *always* present. In what follows, we also limit ourselves to reconstruction of the density field, the analogy for the velocity field and the peculiar gravitational potential being obvious.

5.1 Nonlinear clustering

On small scales, the gravitational evolution has entered into the deeply non-linear regime, where the "fingers-of-God effect" causes virialized clusters to appear elongated along the line of sight. Although some authors have attempted to extend the Kaiser equation into the quasi-linear regime using the Zel'dovich approximation, an exact reconstruction of the real-space density is of course impossible in the deeply non-linear regime. Essentially, this is because the gravitational n-body problem exhibits chaotic behavior, rendering the inversion problem numerically ill-posed. We thus follow the same approach as all previous authors, and smooth all our fields (δ_S, δ_R, *etc.*):

$$\delta_S(\mathbf{r}) \mapsto (W \star \delta_S)(\mathbf{r}) = \int W(\mathbf{r}' - \mathbf{r})\delta_S(\mathbf{r}')d^3r', \tag{40}$$

the star denoting convolution. For definiteness, let us choose a Gaussian smoothing kernel

$$W(\mathbf{r}) \equiv \frac{1}{(2\pi)^{3/2}r_s^3}e^{-\frac{1}{2}(r/r_s)^2}, \tag{41}$$

and choose the *smoothing scale* r_s to be large enough that the resulting fields can be expected to obey linear theory[2].

5.2 Shot noise

What we measure is of course not the smooth field δ_S but the discrete galaxy density $n(\mathbf{r}) = \sum \delta(\mathbf{r} - \mathbf{r}_i)$, with a delta function at the location $\mathbf{r}_i$ of each galaxy. We estimate the unsmoothed density by simply $\delta n/\bar{n} \equiv (n - \bar{n})/\bar{n}$, where $\bar{n}$, the *selection function*, specifies the survey geometry by giving the expected density of observed galaxies at each position in space. Our estimate of of the redshift space density δ_S is thus the convolution of $n/\bar{n} - 1$ with the smoothing kernel W. Hence our real-world generalization of equation (30) becomes

$$\delta_R'(\mathbf{r}) \equiv \int_V G_\delta(\mathbf{r}, \mathbf{r}') \left(W \star \left[\frac{n}{\bar{n}} - 1\right]\right)(\mathbf{r}'), \tag{42}$$

where the prime on the left hand side has been added to distinguish the estimate δ_R' from the true value δ_R. We will return to the issue of how to choose V, the volume to integrate over, in the following subsection. Since

[2] Since smoothing suppresses high-frequency Fourier components, its success of course hinges on the standard assumption that the gravitationally induced mode coupling does not cause a significant "blue leak" of power from short to long wavelengths.

convolution is associative (simply corresponding to multiplication in Fourier space), equation (42) is equivalent to

$$\delta'_R(\mathbf{r}) = \int_V (W \star G_\delta)(\mathbf{r}, \mathbf{r}') \left[\frac{n(\mathbf{r}')}{\bar{n}(\mathbf{r}')} - 1\right], \tag{43}$$

i.e., to smoothing the Green function (with respect to $\mathbf{r}'$) instead of the galaxy distribution.

Let us, as is customary, model n as a Poisson process whose intensity (average point density) is $\lambda(\mathbf{r}) \equiv \bar{n}(\mathbf{r})[1 + \delta_S(\mathbf{r})]$. We model (the smoothed) δ_S as a random field with $\langle \delta_S(\mathbf{r}) \rangle = 0$ and define its correlation function as

$$\xi_s(\mathbf{r}, \mathbf{r}') \equiv \langle \delta_S(\mathbf{r}) \delta_S(\mathbf{r}') \rangle, \tag{44}$$

but make no assumptions of it being homogeneous, isotropic or Gaussian. Defining the *reconstruction error* as

$$\Delta(\mathbf{r}) \equiv \delta'_R(\mathbf{r}) - \delta_R(\mathbf{r}), \tag{45}$$

we now proceed to compute the statistical properties of this error. Since a Poisson process satisfies

$$\langle n(\mathbf{r}) \rangle = \lambda(\mathbf{r}), \tag{46}$$

$$\langle n(\mathbf{r}) n(\mathbf{r}') \rangle = \lambda(\mathbf{r})\lambda(\mathbf{r}') + \delta(\mathbf{r} - \mathbf{r}')\lambda(\mathbf{r}), \tag{47}$$

a straightforward calculation shows that

$$\langle \Delta(\mathbf{r}) \rangle = 0, \tag{48}$$

$$\langle \Delta(\mathbf{r})^2 \rangle = \sigma_n(\mathbf{r})^2 + \sigma_t(\mathbf{r})^2, \tag{49}$$

where

$$\sigma_n(\mathbf{r})^2 \equiv \int_V \frac{[(W \star G_\delta)(\mathbf{r}, \mathbf{r}')]^2}{\bar{n}(\mathbf{r}')} d^3r', \tag{50}$$

$$\sigma_t(\mathbf{r})^2 \equiv = \int_{\not V} \int_{\not V} G_\delta(\mathbf{r}, \mathbf{r}') G_\delta(\mathbf{r}, \mathbf{r}'') \xi_S(\mathbf{r}', \mathbf{r}'') d^3r' d^3r''. \tag{51}$$

When computing σ_n^2, the *shot noise error*, the integral is to be taken over the volume V (see below). When computing σ_t^2, the *truncation error*, the integrals are to be taken over all space *except* the volume V. Equation (48) simply tells us that our reconstruction is *unbiased*, in the sense that the expectation value of the error is zero. The remainder of this subsection is devoted to the shot noise term in Equation (49).

Although the shot noise error σ_n is easily computed numerically using equation (50), it is instructive to make a crude estimate of it. First, let us assume that $\bar{n}(\mathbf{r})$ is constant, so that we can factor it out of the integral in equation (50). Second, let us make the distant survey approximation for G_δ (this is a good approximation if $\mathbf{r}$ is more than a few smoothing lengths away from us, as it is easy to show that the smoothed Green function $W \star G$ falls off rapidly, at least as fast as r^{-3}, for $|\mathbf{r}-\mathbf{r}'| \gg r_s$, so that the main contribution to the integral comes from within a few smoothing lengths of $\mathbf{r}$). Third, let us overestimate σ_n^2 slightly by extending the integration to all space (this makes only a minor difference, as we will find it optimal to integrate considerably beyond the smoothing scale anyway). Now we can apply Parseval's theorem and the convolution theorem to the remaining integral, and obtain

$$\sigma_n^2 \approx \frac{1}{(2\pi)^3 \bar{n}} \int |\widehat{G_\delta}(\mathbf{k})|^2 |\widehat{W}(\mathbf{k})|^2 d^3k, \tag{52}$$

where hats denote Fourier transforms. From equation (34), we immediately recover Kaiser's familiar result

$$\widehat{G_\delta}(\mathbf{k}) \approx \left[1+\beta\left(\frac{k_z}{k}\right)^2\right]^{-1}, \tag{53}$$

so substituting this into equation (52), we obtain the estimate

$$\sigma_n^2 \approx \frac{1}{(2\pi)^3 \bar{n}} \int \frac{e^{-(r_s k)^2} k^2 \sin\theta}{[1+\beta\cos^2\theta]^2}\, dk\, d\theta\, d\phi = \left(\frac{\arctan\sqrt{\beta}}{8\pi^{3/2}\sqrt{\beta}}\right)\left(\frac{1}{\bar{n} r_s^3}\right). \tag{54}$$

Defining the effective *smoothing volume*

$$V_s \equiv \left(\frac{8\pi^{3/2}\sqrt{\beta}}{\arctan\sqrt{\beta}}\right) r_s^3, \tag{55}$$

we can conclude this section with the familiar-looking shot noise formula

$$\sigma_n^2 \approx \frac{1}{\bar{n} V_s}. \tag{56}$$

In other words, the crucial quantity is the expected number of galaxies in a smoothing volume.

5.3 Finite volume

Although the truncation error σ_t is straightforward to compute numerically given any prescribed volume V, either by evaluating the integral in equation (51) or by making reconstructions from Monte Carlo galaxy catalogs, we now make a very crude estimate σ_t to obtain a qualitative understanding of how to best choose V and the smoothing scale. Although the unsmoothed fields δ_R and δ_S have ample structure on very small scales, the corresponding smoothed fields are virtually featureless on scales $r \ll r_s$. Hence the correlation ξ_s is almost perfect for $|\mathbf{r}'-\mathbf{r}| \ll r_s$, and typically falls off like some power law for $|\mathbf{r}'-\mathbf{r}| \gg r_s$. Thus let us make the crude approximation

$$\xi_s(\mathbf{r},\mathbf{r}') \approx \begin{cases} \sigma^2 & \text{for } |\mathbf{r}'-\mathbf{r}| \ll r_s, \\ 0 & \text{for } |\mathbf{r}'-\mathbf{r}| \gg r_s, \end{cases} \tag{57}$$

for some constant σ which is the r.m.s. fluctuation in a smoothing volume. The only non-negligible contributions to the integrals in equation (51) thus arise when $|\mathbf{r}'-\mathbf{r}| \lesssim r_s$. This can be interpreted as there being a large number of independent volumes in which the field is coherent, and that the total variance σ_t^2 is simply a sum of the variance arising from each coherence volume that we do not look at, *i.e.*, that falls outside of V. Assuming that V is large relative to the smoothing scale, replacing $G_\delta(\mathbf{r},\mathbf{r}'')$ by $G_\delta(\mathbf{r},\mathbf{r}')$ will thus be a good approximation, and equation (51) reduces to simply

$$\sigma_t(\mathbf{r})^2 \approx a\,\sigma^2 V_s \int_{\bar{V}} G_\delta(\mathbf{r},\mathbf{r}')^2 d^3r', \tag{58}$$

where a is some constant of order unity. In the distant survey approximation of equation (38), we see that $G_\delta(\mathbf{r},\mathbf{r}') \propto |\mathbf{r}'-\mathbf{r}|^{-3}$ far away, so if V is chosen to be a sphere of radius r_c around $\mathbf{r}$, then

$$\sigma_t^2 \propto \left(\frac{r_s}{r_c}\right)^3 \sigma^2 \propto \frac{V_s}{V}\sigma^2. \tag{59}$$

5.4 How to choose V and the smoothing scale

In this section, we give some useful rules of thumb for how to choose the volume V and the smoothing scale r_s.

Since the integrand in equation (50) is always positive, the shot noise error σ_n always *increases* if we increase the volume V, whereas equation (59) shows that the truncation error σ_t *decreases* as we increase V. Since all we care about is the total error $\sigma_n^2+\sigma_t^2$, we find ourselves facing a trade-off. Similarly, equation (56) shows that the shot noise error decreases if we increase the smoothing volume V_s, whereas equation (59) shows that the relative truncation error σ_t/σ increases if we increase the smoothing volume V_s. Minimizing the total error by differentiating $\sigma_n^2+\sigma_t^2$ with respect to V and V_s, we of course end up with parameter choices such that the two sources of error are of comparable importance, *i.e.*, such that σ_n is of the same order of magnitude as σ_t. Another way of phrasing this is that given some finite level of truncation error σ_t, there is no point in trying to reduce the shot noise error σ_n way below this value (and vice versa), as this will cause almost no reduction in the total error. We thus obtain one of our rules of thumb by simply equating σ_n and σ_t, which after substituting equations (56) and (59) gives

$$\sigma V_s \approx \sqrt{\frac{V}{\bar{n}}}\ , \tag{60}$$

so that the optimal smoothing volume V_s then goes roughly as the geometric mean of the total volume V and the average volume per galaxy, $1/\bar{n}$.

In practice, the survey specifics often place a firm lower limit on either σ_n or σ_t, which makes the best choices of V and V_s quite simple. We now turn to a couple of specific examples:

5.4.1 Slices and pencil beams

If the survey geometry is that of a thin slice, such as for instance the CfA or Las Campanas redshift surveys, or that of a so-called pencil beam, then the dominant source of error at a point $\mathbf{r}$ will usually be the truncation error σ_t caused by the most nearby edge of the slice/pencil. This means that we should choose the smoothing scale as small as we feel comfortable with given the non-linearities on small scales, perhaps around $r_s \approx 10h^{-1}$Mpc. Equation (58) shows that since ${G_\delta}^2$ falls off so rapidly (about as $|\mathbf{r}'-\mathbf{r}|^{-6}$), the truncation error will be completely dominated by the smallest distance from $\mathbf{r}$ to the edge of the volume V. In other words, given any volume V, we can get almost as small truncation errors by replacing it by the largest sphere around $\mathbf{r}$ that will fit inside it, and this defines a characteristic σ_t that we refer to below. In practice, the reconstruction calculations are so fast that there is little need for such attempts to save computer time, especially considering the amount of effort that has gone into the preceding data collection. Rather, the main concern is that we want to avoid extending

V out to such large radii that the shot noise error becomes as large as σ_t. Since $\bar{n}(\mathbf{r})$ is usually a decreasing function of r, Equation (56) shows that this would happen if r where allowed to be so large that $\bar{n}(r) \lesssim 1/\sigma_t V_s$. In conclusion, when reconstructing the real-space fields at a point $\mathbf{r}$, we will do fine if we simply choose V to be the entire slice volume, but truncated at some cutoff radius r_c such that

- $r_c - r$ exceeds the sideways distance from $\mathbf{r}$ to the edge of the slice/pencil beam
- $\bar{n}(r_c) \gtrsim 1/\sigma_t V_s$.

5.4.2 "Non-skinny" surveys

For surveys where all three dimensions are large relative to the smoothing length (such as the IRAS 1.2Jy all-sky survey), the situation is quite different. When reconstructing the real-space field at points r that are many smoothing lengths away from the survey boundaries, truncation errors will not a priori dominate as above. Hence an efficient analysis strategy is to simply truncate the data at some radius r_c where $\bar{n}$ begins to fall off rapidly, take V to be the the entire remaining survey volume (since σ_n saturates a few smoothing lengths away from $\mathbf{r}$ anyway), and then select the smoothing scale according to equation (60).

6 Examples

We now evaluate the Green function method in two simple numerical examples. First, for a qualitative assessment, we realize a single 3D field in real space, remap the density into redshift space with peculiar flows derived from linear theory, and then perform the reconstruction of the real density field with the Green function method. The results of this procedure are given in Figure 2 for an $\Omega = \beta = 1$ Universe, where 2D slices of the density fields are shown. In addition to the reconstructed field (lower-left plot) we provide a reconstruction using an incorrect β value of 0.25 (lower-right plot). This illustrates the sensitivity of the method to β, a feature which may be useful in constraining Ω and the degree of linear bias.

Details of the procedure that led to the data in Figure 2 are as follows: We start with a triply periodic 250 h^{-1} Mpc cube and realizes a Gaussian random field on grid of 64^3 points. The power spectrum of density fluctuations has the form of standard cold dark matter (e.g., Efstathiou, Bond

& White 1992), smoothed with a Gaussian filter on scales of 50 h^{-1} Mpc. The normalization is chosen near the COBE value to approximate observed large-scale structure; before smoothing, the r.m.s. density fluctuations in a sphere of radius 8 h^{-1} Mpc are at a level of 1.1 relative to the mean density, *i.e.*, $\sigma_8 = 1.1$. We then estimate the velocity field at the grid points and displace the the points by an amount appropriate for an observer at some chosen origin to mimic the distribution in redshift space. These points are interpolated onto a lower resolution grid (32^3) to reduce discreteness effects. This lower-resolution grid represents the volume-limited red-shift space field that is used in the reconstruction. For the 2D plots in Figure 2, we perform the reconstruction on a plane of grid points; at each point a 3D integral over the redshift-space grid is evaluated with the kernel of equation (32) under the assumption that the density field is a set of weighted Dirac-delta functions on the grid. For speed enhancement, the sum in equation (32) was evaluated at a grid of points (x, μ) once and for all in a lookup table. When $\mathbf{r} \approx \mathbf{r}'$, the order zero analytic reconstruction kernel (Ozark) of equation (38) was used.

Evidently, the reconstruction works well for regions away from the edges of the mock survey in Figure 2. The reconstruction has recovered both the amplitude of the original field and the topology of the isodensity contours. However, a qualitative assessment of the errors is in order. Our focus in this section is on truncation errors as well as errors incurred during interpolation and integration. We estimate the errors with a Monte Carlo method; we realize 72 random density fields, perform reconstruction on each as described above, and determine the r.m.s. point-wise difference between the real density field and the reconstruction. Figure 3 shows the r.m.s. error in reconstruction for points on one of the principle axes of the cubical survey volume (the points lie on a path which bisects the plane shown in Figure 2).

Truncation effects are clear in Figure 3 from the increase in error near the boundary of the survey volume. The truncation error worsens as the correlation length of the density field increases, as can be seen from a comparison of the error curves with correlation lengths of 50 and 100 h^{-1} Mpc (induced by smoothing on those scales); large correlation lengths correspond to low frequency modes which are more susceptible to the ill effects of a finite survey volume.

For the case of the realizations with a smoothing scale of 50 h^{-1} Mpc, we estimate the theoretical truncation error by directly integrating equation (51). The relatively high dimensionality of the integral and the rapid variation of the integrand near the boundaries of the survey volume suggest that this problem is best handled by parallel computation (we are grateful for time on a Cray T-3D at Jet Propulsion Laboratory). For reference, an approximate theoretical error curve is shown in Figure 3. The remaining error comes from grid interpolation, the numerical integrator, and to a lesser extent the breakdown in the Kaiser equation, which is of course only exact for $\delta \ll 1$. This latter source of error scales roughly as δ^2, and does become more important at smaller smoothing scales. While in this example we are content to let these errors lie between 5% and 10% near the center of the survey, they may be further reduced with a finer grid mesh and a higher-order integrator.

7 Discussion

We have presented the Green functions for computing the gravitational potential, the peculiar velocity field and the real-space density fluctuations from the distribution of matter in redshift space. Our results are based on Kaiser's (1987) analysis of redshift-space distortions in the linear regime, and are applicable to surveys of arbitrary geometry.

One virtue of the method presented here is that the entire calculation is performed in real space. This means that the effects of complications such as survey boundaries and the presence of galaxy clusters, which are manifestly in physical space rather than Fourier space, are easy to understand and quantify. For example, the magnitude of the Green function in the parts of space that we are ignoring immediately gives an indication of how much information we are loosing, *i.e.*, on the magnitude of the error bars on the reconstructed field. As we have discussed, this allows the problems caused by shot noise, survey boundaries and nonlinear clustering to be minimized by appropriate choices of the smoothing scale and the integration volume V.

We wish to emphasize several possible uses of the Green function method. The first is toward ferreting out estimates of the linear growth parameter $\beta = \Omega^{0.6}/b$. While β has been estimated in elegant manners from the correlation anisotropy (e.g., Hamilton 1992), it may also be determined with the Green function G_δ: The best estimate for β is the value which minimizes the anisotropy in the real-space two-point correlation function of the reconstructed field. Quantitative conditions for minimizing the anisotropy in ξ_R may be derived with an orthogonal decomposition in spherical harmonics (e.g., Hamilton 1992). For example, a least-squares criterion may be imposed on the amplitudes of terms with spherical harmonics Y_{lm} for which $l > 0$, with greatest weights assigned to quadrupole terms, as they couple strongly to the redshift-space distortions.

The second application of the Green function method is to extract information on the nature of biasing. Specifically, the reconstructed real-space density field contains the implicit assumption of a linear bias between luminous and dark matter. Thus it is worth comparing the field with density reconstructions from other techniques such as the POTENT algorithm (Bertschinger & Dekel 1989). While such a comparison will not alone break the degeneracy between b and Ω, it can test the validity of the linear biasing hypothesis.

Ultimately, the method's most important use is the estimation of peculiar velocities. Even linear peculiar flows can provide stringent constraints on cosmological models (e.g., Lauer & Postman 1994; Tegmark, Bunn & Hu 1994). Since the Green function method generates peculiar velocities directly from redshifts without the difficult measurement of real-space distances, it may have fruitful application in the large-scale digital sky surveys that are currently coming on line.

Our purpose here has been to provide the mathematical underpinnings of the Green function approach to cosmic field reconstruction from redshift data. We are now embarking on the next step, the analysis of redshift surveys.

We thank James Anglin, Marc Davis, Andrew Hamilton, Saleem Zaroubi and our referee for useful comments and suggestions. BCB acknowledges support from the NASA HPCC program and help from Isabel Dulfano.

References

Bertschinger, E., & Dekel, A. 1989, ApJ, 336, L5

Bromley, B. C. 1994, ApJ, 423, L81

Cole, S., Fisher, K. B., & Weinberg, D. H. 1994, MNRAS, 267, 785

Efstathiou, G., Bond, J. R., & White, S. D. M. 1992, MNRAS, 258, 1P

Fisher, K. B., Scharf, C. A., & Lahav, O. 1994, MNRAS, 266, 219

Giovanelli, R., & Haynes, M. P. 1991, Ann. Rev. Astr. Ap., 29, 499

Gramann, M., Cen, R. Y. & Gott III, J. R. 1994, ApJ, 425, 382

Hamilton, A. J. S. 1992, ApJ, 385, L5

Hamilton, A. J. S. 1993, ApJ, 406, L47

Kaiser, N. 1987, MNRAS, 227, 1

Lauer, T. & Postman, M. 1994, ApJ, 425, 418

Nusser, A. & Davis, M. 1994, ApJ, 421, L1

Peebles, P. J. E., 1993, Principles of Physical Cosmology (Princeton: Princeton University Press)

Smoot, G. F. et.al. 1992, ApJ, 396, L1

Taylor, A. & Rowan-Robinson, M. 1993; MNRAS, 265, 809

Tegmark, M., Bunn, E., & Hu, W. 1994, ApJ, 434, 1

Yahil, A., Strauss, M. A., Davis, M., & Huchra, J. P. 1991, ApJ, 372, 380

Zaroubi, S. & Hoffman, Y. 1994, preprint

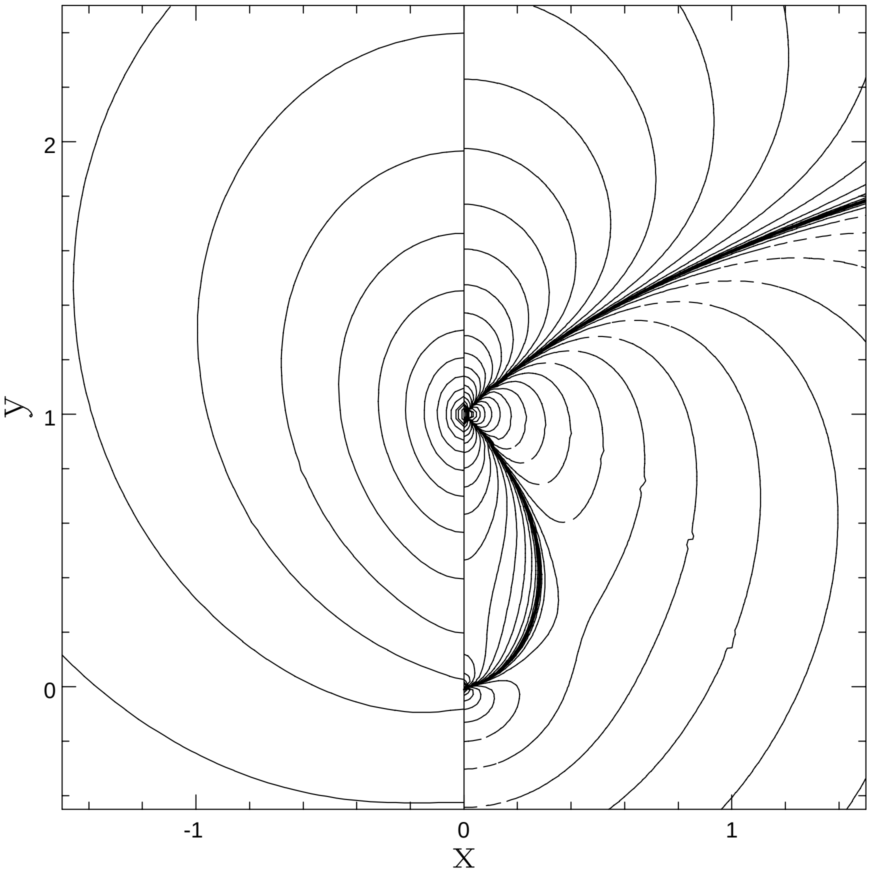

Figure 1: The Green functions G_ϕ and G_δ, for $\beta = 1$.
To the left of the y-axis are contours of $G_\phi(\mathbf{r}, \mathbf{r}')$, while the contours to the right correspond to $G_\delta(\mathbf{r}, \mathbf{r}')$ for $\mathbf{r} = (x, y, 0)$ and $\mathbf{r}' = (0, 1, 0)$. The light solid lines and dashed contours indicate positive and negative values, respectively, with logarithmic spacing. The heavy solid line is the $G_\delta(\mathbf{r}) = 0$ contour. The trend in both functions is toward high absolute values near the center of the figure. Values of these functions for more general choices of $\mathbf{r}$ and $\mathbf{r}'$ follow directly from symmetry and scaling properties. The functions essentially give the contribution of a point mass at $\mathbf{r}$ to the reconstructed field at $\mathbf{r}'$ (point at center of the figure).

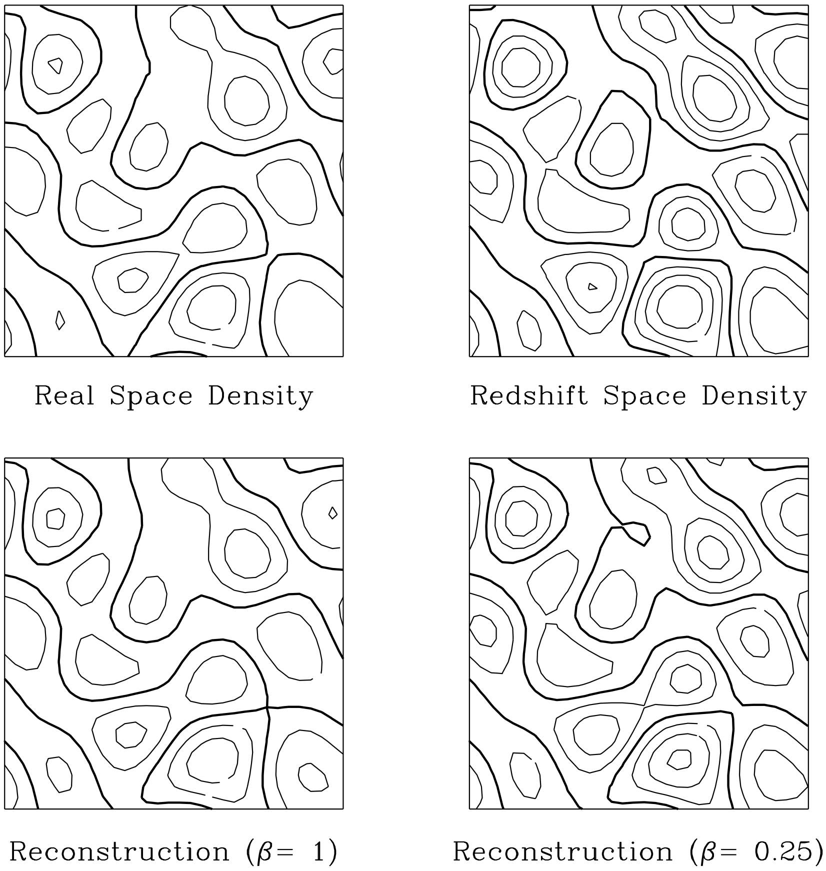

Figure 2: An example of the reconstruction method.
As discussed in the text, the reconstruction method is applied to a linearly evolved Gaussian random field in a 250 h^{-1} Mpc cube. The linearly spaced contours, denoting isodensity surfaces at a mid-plane of the cube, show overdensities (solid lines), underdensities (broken lines), and the mean density (heavy solid lines). The observer is located at a corner of the cube. The upper-left plot is the density in real space, generated from a CDM power spectrum with 50 h^{-1} Mpc smoothing. The corresponding redshift-space density (upper-right) is used to reconstruct the real-space fields; the lower-left plot was derived with the correct β value while an incorrect value yielded the plot on the lower right.

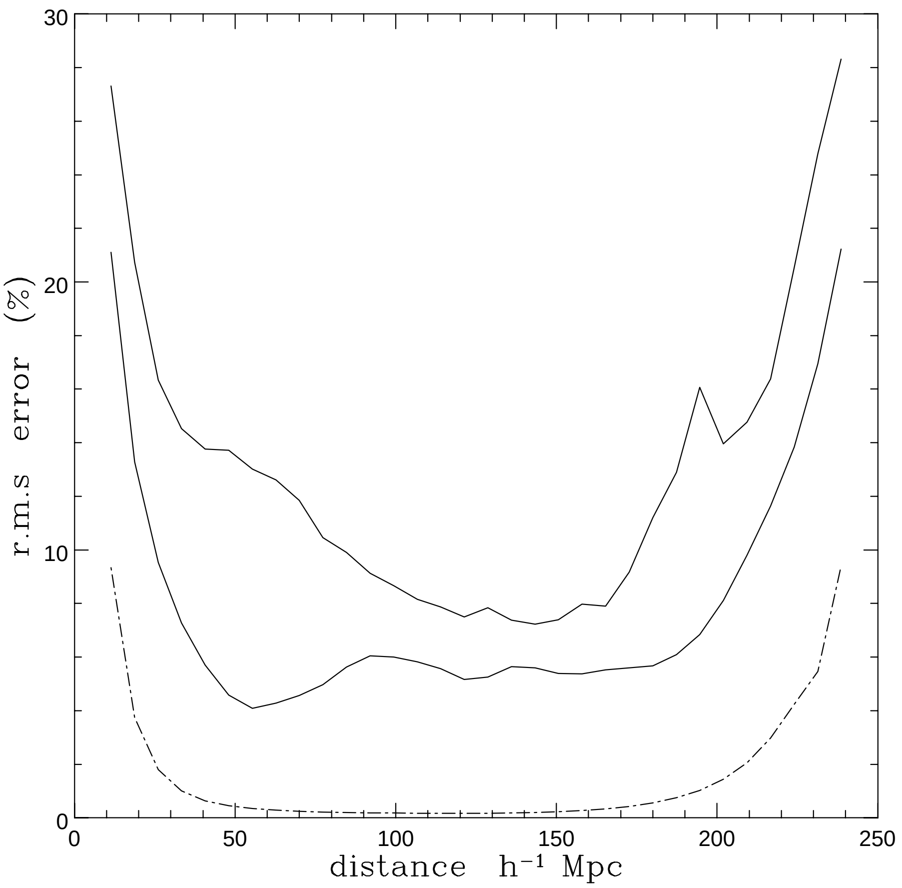

Figure 3: Density reconstruction errors.

The r.m.s. reconstruction errors, $\sigma(\mathbf{r})/\sqrt{\xi_r(0)}$, are estimated from 72 realizations of a linearly evolved CDM field, with 100 h^{-1} Mpc smoothing (top curve) and 50 h^{-1} Mpc smoothing (middle curve), in the survey region used for Figure 2. The horizontal axis is distance along a principle axis of the cube. The lower (dashed) curve is an estimate of the expected truncation noise, $\sigma_t(\mathbf{r})$ for the case of 50 h^{-1} Mpc smoothing.