Redshift Space Distortions

Overview

It is important to realize that our "third dimension" in cosmology is not radial distance but redshift. The two are related by the Hubble expansion but also affected by *peculiar* velocities. On small scales random motions within e.g. a cluster of galaxies will cause particles at the same distance to have slightly different redshifts. This elongates structures along the line of sight and leads to the so-called fingers-of-God effect: structures have a tendency to point toward the observer. On very large scales the opposite happens. Objects fall in towards overdense regions. This makes objects between us and the overdensity appear to be further away and objects on the other side of the overdensity appear closer. The net effect is to enhance the overdensity rather than smear it out. These effects are known as *redshift space distortions* – you can find a review in astro-ph/9708102 (specifically §§1, 2 5 & 6).

Contour plot of $\xi(r_\perp, r_{||})$ in r-space forms a butterly plot, and a sharpening of any feature along the line of sight. In k space the sense is reversed.

Galaxies are expected to be almost unbiased tracers of the cosmic velocity field, so a measurement of redshift space distortions allows a measure of the mass density field (through its relation to the velocity field) and hence of the growth of structure. **Note**: unlike lensing, which depends on both Φ and Ψ, the velocity field of non-relativistic tracers depends only on Ψ. The combination allows constraints on gravity theories.

The Kaiser factor

The former effect is fairly difficult to calculate, but the latter is a simple exercise in linear perturbation theory. You may wish to look at Nick Kaiser's original paper on this subject (MNRAS, 227, 1, 1987). Denote the redshift space coordinate as $\vec{s}$ and the real space coordinate as $\vec{r}$. If $\vec{v}(\vec{r})$ is the peculiar velocity in units of the Hubble constant then

$$\vec{s} = \vec{r}\left(1 + \frac{u(r)}{r}\right) \tag{1}$$

where $u = \hat{r} \cdot \vec{v}(\vec{r})$ and we have assumed $v(0) = 0$. Assuming that the object is so distant[1] that $kr \gg 1$ the Jacobian between $\vec{s}$ and $\vec{r}$ is

$$d^3s = \left(1 + \frac{du}{dr}\right) d^3r \tag{2}$$

where we have dropped the $(1 + u/r)^2$ factor. Number density conservation requires $\delta_s d^3s = \delta_r d^3r$ for plane-wave perturbation δ. Using the linear theory relation $\dot{\delta} = -ikv$ we have

$$\begin{aligned} \frac{du}{dr} &= \mu \frac{d}{dr} v && (3) \\ &= \mu(ik\mu)v && (4) \\ &= -\mu^2(-ikv) && (5) \\ &= -\mu^2 f(\Omega)\delta && (6) \end{aligned}$$

where $\mu = \hat{r} \cdot \hat{k}$ and $f(\Omega) \equiv d\log\delta/d\log a \approx \Omega^{0.6}$ so $\dot{\delta} = f(\Omega)\delta$ in units where $H = 1$. Putting it all together we have

$$\delta_s = \delta_r \left(1 + f\mu^2\right) \tag{7}$$

If we presume that the galaxy fluctuation is biased with respect to the mass by $\delta_{\rm gal} = b\delta_{\rm mass}$ then

$$\Delta^2_{\rm gal}(k,\mu) = b^2 \left(1 + \beta\mu^2\right)^2 \Delta^2_{\rm mass}(k) \tag{8}$$

[1]This is not a good approximation for wide-angle surveys, see Papai & Szapudi 0802.2940 or Castorina & White 1709.09730 for more details. In fact, once you go beyond plane-parallel you lose the translational symmetry of the problem that allows us to simply sum up independent k-modes here. This problem tends to be small in practice.

where $\beta \equiv f(\Omega)/b$. Note that β is not really a cosmological parameter like Ω as it is only defined in the context of the most simple linear biasing model (which is almost certainly wrong). However it is often referred to by people who study velocity fields because $f(\Omega)/b$ relates the velocity field to the density field in this linear bias model. (Non-linear models exist, Willick gives a review and people are revisiting these for density field reconstruction for BAO).

The μ dependence of $\Delta^2_{\rm gal}(k,\mu)$ can be expanded out in Legendre polynomials of order 0, 2, 4. The coefficient of P_0 is $1+\frac{2}{3}\beta+\frac{1}{5}\beta^2$ which is the amount by which clustering is enhanced on large-scales by redshift space distortions[2]. Notice that for $\Omega\sim b\sim 1$ this can by a significant effect! In Fourier space:

$$\frac{\Delta^2_{\rm red}(k,\mu)}{\Delta^2_{\rm real}(k)} = \left[\left(1+\frac{2}{3}\beta+\frac{1}{5}\beta^2\right)P_0(\mu)\right. \tag{9}$$

$$+ \left(\frac{4}{3}\beta+\frac{4}{7}\beta^2\right)P_2(\mu) \tag{10}$$

$$\left. + \frac{8}{35}\beta^2 P_4(\mu)\right] \tag{11}$$

For quite some time it was believed that the ratio of the quadrupole to monopole term would be a good way to measure β – see Berlind, Narayanan & Weinberg (2001; ApJ 549, 688) for a summary of past work and a discussion of issues with this approach.

Numerous numerical simulations have shown that the convergence to the Kaiser factor is very slow – we can understand some of this from PT as discussed below. Thus for precision cos-

[2] Note sign of elongation changes between k and r picture. In configuration space an overdensity generates a convergent flow, "sharpening" the structure so $\xi(r_{||})$ falls more quickly than without redshift distortions.

mology the Kaiser effect represents a limit which is never valid for real surveys.

The correlation function

It is straightforward to derive the correlation function(s) corresponding to the Kaiser power spectrum. Define

$$\Delta^2(k,\widehat{k}\cdot\widehat{z}) = \sum_\ell \Delta^2_\ell(k) P_\ell(\mu) \quad \text{and} \quad \xi(r,\widehat{r}\cdot\widehat{z}) = \sum_\ell \xi_\ell(r) P_\ell(\mu) \tag{12}$$

then the Rayleigh expansion of the plane wave gives

$$\xi_\ell(r) = i^\ell \int \frac{dk}{k}\ \Delta^2_\ell(k) j_\ell(kr) \tag{13}$$

If we use the recurrence relations between the j_ℓs we can write the ξ_ℓ in terms of integrals of ξ times powers of r, e.g.

$$\int \frac{dk}{k}\ \Delta^2_\ell(k) j_2(kr) = \frac{3}{s^3}\int_0^s s^2 ds\, \xi(s) - \xi(s) = \bar{\xi}(<s) - \xi(s) \tag{14}$$

Beyond plane-parallel

In deriving Eq. (8) we assumed the distant observer approximation. This is not necessary. In moving to wide fields it makes sense to make a radial-angular decomposition, since redshift space distortions involve only radial remapping. Thus one expands (Heavens & Taylor, 1995)

$$\rho_{\ell mn}(s) = c_{\ell n} \int d^3s\ \rho(\vec{s}) j_\ell(k_{\ell n} s) Y^*_{\ell m}(\theta,\phi) \tag{15}$$

where j_ℓ are the spherical Bessel functions and $Y_{\ell m}$ the spherical harmonics. If there is no boundary condition at finite r the

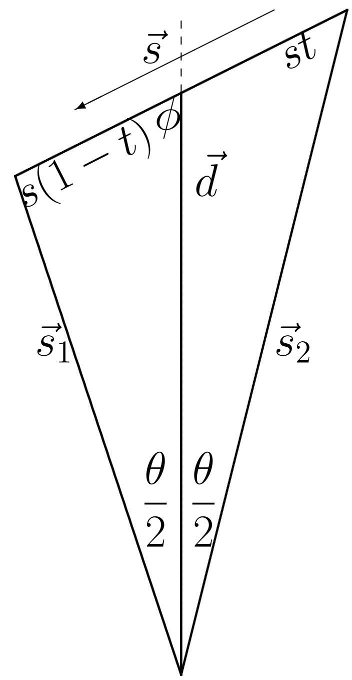

Figure 1: The assumed geometry and angles. The two galaxies lie at $\vec{s}_1$ and $\vec{s}_2$, with separation vector $\vec{s} = \vec{s}_1 - \vec{s}_2$ and enclosed angle θ. We take the line of sight to be parallel to the angle bisector, $\vec{d}$, which divides $\vec{s}$ into parts of lengths st and $s(1-t)$. The separation vector, $\vec{s}$, makes an angle ϕ with the line of sight direction, $\hat{d}$.

sum over n is an integral over k, otherwise the boundary condition must be applied. Redshift space distortions then become relations between the $\rho_{\ell mn}$. For example

$$j_\ell\left(k_{\ell n}s\right) \simeq j_\ell\left(k_{\ell n}r\right) + v(r)\frac{d}{dr}j_\ell\left(k_{\ell n}r\right) + \cdots \tag{16}$$

and the derivative can be written, using recurrence relations, in terms of the $j_{\ell+1}$ and $j_{\ell-1}$.

If one is going to drop the distant observer approximation, one also needs to worry about the fx_jv_j/x^2 term – see Papai & Szapudi (2008; MNRAS 389, 292).

Keeping to linear theory we begin with the redshift space density field

$$\delta(\mathbf{s}) = \int \frac{d^3k}{(2\pi)^3}\, e^{i\mathbf{k}\cdot\mathbf{x}} \left[1 + f(\mathbf{k}\cdot\mathbf{x})^2 - 2if\frac{\widehat{k}\cdot\widehat{x}}{kx}\right]\delta(\mathbf{k}) \tag{17}$$

so that

$$\begin{aligned}\langle\delta(\mathbf{x}_1)\delta^\star(\mathbf{x}_2)\rangle &= \int \frac{d^3k}{(2\pi)^3}\, P(k)e^{i\mathbf{k}\cdot(\mathbf{x}_1-\mathbf{x}_2)} \\ &\left[1+\frac{f}{3}+\frac{2f}{3}P_2(\widehat{k}\cdot\widehat{x}_1) - 2if\frac{P_1(\widehat{k}\cdot\widehat{x}_1)}{kx_1}\right] \\ &\left[1+\frac{f}{3}+\frac{2f}{3}P_2(\widehat{k}\cdot\widehat{x}_2) - 2if\frac{P_1(\widehat{k}\cdot\widehat{x}_2)}{kx_2}\right]\end{aligned} \tag{18}$$

We can expand ξ in terms of

$$S_{\ell_1\ell_2\ell}(\widehat{x}_1,\widehat{x}_2,\widehat{x}) = \sum_{m_1,m_2,m} \begin{pmatrix} \ell_1 & \ell_2 & \ell \\ m_1 & m_2 & m \end{pmatrix} C_{\ell_1 m_1}(\widehat{x}_1) C_{\ell_2 m_2}(\widehat{x}_2) C_{\ell m}(\widehat{x}) \tag{19}$$

where $C_{\ell m} = \sqrt{4\pi/(2\ell+1)}Y_{\ell m}$ and use the identities

$$P_\ell(\widehat{x}_1\cdot\widehat{x}_2) = \frac{4\pi}{2\ell+1}\sum_m Y_{\ell m}(\widehat{x}_1)Y^\star_{\ell m}(\widehat{x}_2) \tag{20}$$

$$e^{i\mathbf{k}\cdot\mathbf{x}} = 4\pi\sum_{\ell m} i^\ell j_\ell(kx) Y_{\ell m}(\widehat{k})Y^\star_{\ell m}(\widehat{x}) \tag{21}$$

and the Gaunt integral

$$\int d\Omega\, Y_{\ell_1 m_1}(\widehat{n})Y_{\ell_2 m_2}(\widehat{n})Y_{\ell_3 m_3}(\widehat{n}) \propto \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \tag{22}$$

to find expressions for the expansion coefficients $B^{\ell_1\ell_2\ell}$ in terms of β and ξ_ℓ. In the limit $\widehat{x}_1 \approx \widehat{x}_2$ we can use

$$S_{\ell_1\ell_2\ell}(\widehat{x}_1,\widehat{x}_1,\widehat{x}) = \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix} P_\ell(\widehat{x}_1\cdot\widehat{x}) \tag{23}$$

More general expressions can be found in Papai & Szapudi.
Derivation of plane-parallel limit: Expand $S_{\ell_1\ell_2\ell}(\widehat{x}_1,\widehat{x}_1,\widehat{x})$ in Legendre polynomials, $A_\ell P_\ell(\widehat{x}_1\cdot\widehat{x})$. Write $P_\ell(\widehat{x}_1\cdot\widehat{x})$ as a sum of $Y_{\ell m}(\widehat{x}_1)Y_{\ell m}(\widehat{x})$ and hence get an expression for $A_{\ell m}$. Using the definition of S the $A_{\ell m}Y_{\ell m}(\widehat{x}_1)$ is $3j$-symbols times a product $Y_{\ell_1 m_1}(\widehat{x}_1)Y_{\ell_2 m_2}(\widehat{x}_1)$. Integrate this against $Y_{\ell m}(\widehat{x}_1)$ to get $A_{\ell m}$ as the $3j$ symbol times the integral of a product of the three $Y_{\ell m}$s all of argument $\widehat{x}_1$. The integral is

$$\begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \tag{24}$$

Now use the identity

$$(2\ell+1)\sum_{m_1 m_2}\begin{pmatrix} \ell_1 & \ell_2 & \ell \\ m_1 & m_2 & m \end{pmatrix}\begin{pmatrix} \ell_1 & \ell_2 & \ell' \\ m_1 & m_2 & m' \end{pmatrix} = \delta_{\ell\ell'}\delta_{mm'} \tag{25}$$

to get the result above.

Explicit wide-angle expressions in linear theory

The linear theory correlation function for arbitrary triangles was derived by Szalay et al. (1998). We briefly recap how that derivation proceeds here. To keep the derivation as short as possible we only show some terms and in particular we omit the fx_jv_j/x^2 terms. The other terms follow a similar pattern and can be found in Szalay et al. (1998) if desired.

Recall the redshift-space density in linear theory is (Kaiser 1987)

$$\delta^{(s)}(\mathbf{s}) = \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{s}} \left(1 + \beta[\hat{k}\cdot\hat{s}]^2\right) \delta^{(r)}(\mathbf{k}) \tag{26}$$

If we define

$$\delta_\ell \equiv \int \frac{d^3k}{(2\pi)^3} \mathcal{L}_\ell(\hat{k}\cdot\hat{s}) e^{i\mathbf{k}\cdot\mathbf{s}} \delta(\mathbf{k}) \tag{27}$$

then using $\mu^2 = (2/3)\mathcal{L}_2(\mu) + (1/3)\mathcal{L}_0(\mu)$ we have

$$\delta^{(s)}(\mathbf{s}) = \left(1 + \frac{\beta}{3}\right) \delta_0 + \frac{2\beta}{3}\delta_2 \tag{28}$$

The correlation function is thus

$$\begin{aligned} \xi(\mathbf{s}_1, \mathbf{s}_2) = & \left(1 + \frac{\beta}{3}\right)^2 \langle\delta_0\delta_0\rangle + \frac{4}{9}\beta^2 \langle\delta_2\delta_2\rangle \\ & + \frac{2\beta}{3}\left(1 + \frac{\beta}{3}\right) \langle\delta_0\delta_2 + \delta_2\delta_0\rangle \end{aligned} \tag{29}$$

To evaluate the expectation values we expand $\mathcal{L}$ and $\exp[i\mathbf{k}\cdot\mathbf{s}]$ in spherical harmonics and integrate over $d\Omega_k$. For example

$$\left\langle\delta_0^2\right\rangle = \int \frac{k^2\,dk}{2\pi^2} P(k) j_0(ks) \tag{30}$$

as expected while

$$\langle\delta_0\delta_2 + \delta_2\delta_0\rangle = -\left[\mathcal{L}_2(\hat{s}\cdot\hat{s}_1) + \mathcal{L}_2(\hat{s}\cdot\hat{s}_2)\right] \int \frac{k^2\,dk}{2\pi^2} P(k) j_2(ks) \tag{31}$$

$$= -\left[2\mathcal{L}_2(\mu)\cos\theta + \frac{1}{2}(1 - \cos\theta)\right] \int \frac{k^2\,dk}{2\pi^2} P(k) j_2(ks) \tag{32}$$

and

$$\left\langle\delta_2^2\right\rangle = \int \frac{k^2\,dk}{2\pi^2}P(k)\sum_L i^L j_L(ks)\ \left(\frac{4\pi}{5}\right)^2 \times \sum_{M,m_1,m_2} \mathcal{G}_{L22}^{Mm_1m_2} Y^\star_{LM}(\hat{s})Y^\star_{2m_1}(\hat{s}_1)Y^\star_{2m_2}(\hat{s}_2) \qquad (33)$$

where $\mathcal{G}$ is the Gaunt integral. To evaluate the last line, set $\hat{d}$ to be the $\hat{z}$-axis and orient the triangle to lie in the $x-z$ plane (so all of the polar angles are zero or π). Note that $\hat{s}_1$ and $\hat{s}_2$ are both at angle $\theta/2$ to $\hat{d}$ while $\hat{s}$ is at angle $\pi-\phi$. Only $0\le L\le 4$ are non-zero and using the explicit forms of the $Y_{\ell m}$ then gives

$$\left\langle\delta_2^2\right\rangle \ni \frac{\mathcal{L}_2(\cos\theta)}{5}\int \frac{k^2\,dk}{2\pi^2}P(k)j_0(ks) \qquad (34)$$

for the $L=0$ contribution

$$\left\langle\delta_2^2\right\rangle \ni \frac{1}{28}\left[1-3\cos(2\theta) -3\cos(2\phi-\theta)-3\cos(2\phi+\theta)\right] \times\int \frac{k^2\,dk}{2\pi^2}P(k)j_2(ks) \qquad (35)$$

for the $L=2$ contribution and

$$\left\langle\delta_2^2\right\rangle \ni \frac{9}{1120}\left[6+35\cos(4\phi)+3\cos(2\theta)+ 10\cos(2\phi-\theta)+10\cos(2\phi+\theta)\right] \times\int \frac{k^2\,dk}{2\pi^2}P(k)j_4(ks) \qquad (36)$$

for $L=4$. Note that in the limit $\theta\to 0$

$$\left(1+\frac{\beta}{3}\right)^2+\frac{4\beta^2}{45}\mathcal{L}_2(\cos\theta)\to 1+\frac{2}{3}\beta+\frac{1}{5}\beta^2 \qquad (37)$$

and the $L = 4$ part of

$$\frac{4\beta^2}{9}\left\langle\delta_2^2\right\rangle \to \frac{8\beta^2}{35}\mathcal{L}_4(\mu)\int\frac{k^2\,dk}{2\pi^2}P(k)j_4(ks) \tag{38}$$

as desired. The other terms follow a similar pattern, and the results[3] can be found in Szalay et al. (1998; though beware that their θ is half ours). It is easy to show that the corrections to the plane-parallel limit start at $\mathcal{O}(\theta^2)$.

Beyond linear theory

Of course one can go beyond linear theory. Staying within the plane-parallel approximation (so that we keep translational invariance, so k-modes are the 'right' basis) and using number density conservation we have

$$d^3s \equiv J(x)d^3x \quad , \quad J(x) = |1 - f\nabla_z v_z(x)| \tag{39}$$

so

$$1 + \delta_s(s) = \frac{1+\delta(x)}{J(x)} \tag{40}$$

The correlation function in redshift space can therefore be written

$$1 + \xi_s\left(\pi, \sigma\right) = \int\frac{dr dk}{2\pi}e^{-ik(r-\pi)}\left\langle e^{ifk(u-u')}\left[1+\delta\right]\left[1+\delta'\right]\right\rangle \tag{41}$$

which we see involves all powers of the density and velocity cross correlation.

[3] Note that Eq. (15) of Szalay et al. (1998) contains a typographical error. The 4/15 should be 8/15.

If we expand the second exponential we have

$$\begin{aligned}\delta_s(k) &= \sum_{n=1}^{\infty}\int\frac{d^3k_1\cdots d^3k_n}{(2\pi)^{3n}}(2\pi)^3\delta_D(k-k_1-\cdots-k_n)\\ &\times \left[\delta(k_1)+f\mu_1^2\theta(k_1)\right]\frac{(f\mu k)^{n-1}}{(n-1)!}\frac{\mu_2}{k_2}\theta(k_2)\cdots\frac{\mu_n}{k_n}\theta(k_n)\end{aligned} \tag{42}$$

To first order we regain the Kaiser expression. Beyond linear order we have

$$\delta_s(k)=\sum_{n=1}^{\infty}\int d^3k_1\cdots d^3k_n\delta_D(k-k_1-\cdots-k_n)Z_n(k_1,\cdots,k_n)\delta_1(k_1)\cdots\delta_1(k_n) \tag{43}$$

where $Z_1=b_1+f\mu^2$ and

$$\begin{aligned}Z_2(k_1,k_2) &= b_1F_2(k_1,k_2)+f\mu^2G_2(k_1,k_2)+\frac{b_2}{2}\\ &+ \frac{f\mu k}{2}\left[\frac{\mu_1}{k_1}(b_1+f\mu_2^2)+\frac{\mu_2}{k_2}(b_1+f\mu_1^2)\right]\end{aligned} \tag{44}$$

and so on.

Another approach is to use Lagrangian perturbation theory. Here we shall work to 1st order – the Zel'dovich approximation. Particles initially at $\mathbf{q}$ are mapped to $\mathbf{x}=\mathbf{q}+\mathbf{\Psi}$ with $\mathbf{\Psi}=\sum\mathbf{\Psi}^{(n)}$ and $\mathbf{\Psi}^{(n)}\propto D^n$. Going to redshift space means replacing $\mathbf{\Psi}$ with

$$\mathbf{\Psi}\to\mathbf{\Psi}+\frac{\widehat{z}\cdot\dot{\mathbf{\Psi}}}{H}\widehat{z}=R\mathbf{\Psi} \tag{45}$$

and since $\mathbf{\Psi}^{(n)}\propto D^n$ we have $R_{ij}^{(n)}=\delta_{ij}+nf\widehat{z}_i\widehat{z}_j$. Taking the Fourier transform of the sum of δ-functions:

$$\delta(\mathbf{x}) = \int d^3q\ \delta^{(D)}(\mathbf{x}-\mathbf{q}-\mathbf{\Psi})-1 \tag{46}$$

$$\delta(\mathbf{k}) = \int d^3q\ e^{-i\mathbf{k}\cdot\mathbf{q}}\left(e^{-i\mathbf{k}\cdot\mathbf{\Psi}(\mathbf{q})}-1\right) \tag{47}$$

which gives

$$P(k) = \int d^3q \; e^{-i\mathbf{k}\cdot\mathbf{q}} \left(\left\langle e^{-ik_i \Delta\Psi_i(\mathbf{q})} \right\rangle - 1 \right) , \tag{48}$$

where $\mathbf{q} = \mathbf{q}_1 - \mathbf{q}_2$, and $\Delta\mathbf{\Psi} = \mathbf{\Psi}(\mathbf{q}_1) - \mathbf{\Psi}(\mathbf{q}_2)$. For a zero-mean Gaussian x we have $\langle e^x \rangle = \exp\left[\langle x^2 \rangle/2\right]$ so to 1$^{\rm st}$ order

$$P(k) = \int d^3q \; e^{-i\mathbf{k}\cdot\mathbf{q}} \exp\left[-\frac{1}{2} k_i k_j \left\{ \xi_{ij}(0) - \xi_{ij}(\mathbf{q}) \right\} \right] \tag{49}$$

where $\xi_{ij}(\mathbf{q}) \equiv \langle \Psi_i(\mathbf{q}_1) \Psi_j(\mathbf{q}_2) \rangle$. Leave the zero-lag piece exponentiated and expand the exponential of $\xi_{ij}(\mathbf{q})$. Using

$$k_i k_j R_{ia} R_{jb} \delta_{ab} = \left(k_a + f k \mu \widehat{z}_a\right) \left(k_a + f k \mu \widehat{z}_a\right) = k^2 \left[1 + f(f+2)\mu^2\right] \tag{50}$$

and that the Fourier transform of ξ is P_L we have

$$P(k) = \exp\left[- \left\{ 1 + f(f+2)\mu^2 \right\} k^2 \Sigma^2 / 2 \right] \left(1 + f\mu^2\right)^2 P_L(k) + \cdots \tag{51}$$

with

$$\Sigma^2 = \frac{1}{3\pi^2} \int dp \; P_L(p) \simeq \mathcal{O}\left[(10\,\mathrm{Mpc})^2 \right] \tag{52}$$

the displacement of particles in the Zel'dovich approximation. Note that this means any approach to the Kaiser limit is only accurate when $k\Sigma \ll 1$. Also note that including redshift space distortions increases the damping in the line-of-sight direction by $(1+f)$ compared to the transverse direction.

A discussion of the general effect of redshift space distortions for Gaussian fields, but not assuming linearity, is given in Shaw & Lewis (arxiv:0808.1724).

Fingers of God

To deal with small-scale velocities formally requires a model for the velocity field and its correlations with density. This is quite difficult. For the density we have the exact expression

$$\delta_D(k) + \delta_s(k) = \int d^3x \; e^{-ik\cdot x} e^{ifk_z u_z(x)} \left[1 + \delta(x)\right] \tag{53}$$

where δ_D is a Dirac δ-function. The power spectrum is thus

$$\delta_D(k) + P_s(k) = \int d^3r \; e^{-ik\cdot x} \left\langle e^{ifk_z \Delta u_z(x)} \left[1 + \delta(x)\right] \left[1 + \delta(x')\right] \right\rangle \tag{54}$$

where $r = x - x'$. Unfortunately this involves knowing all of the density and velocity correlations. The similar formula for the correlation function can be schemeatically written:

$$1 + \xi_s(s_{||}, s_\perp) = \int dr_{||} \;\; \left[1 + \xi(r)\right] P(r_{||} - s_{||}, r) \tag{55}$$

where P is the velocity distribution function which depends on r. In general it is almost impossible to make rigorous progress here. If we neglect this r dependence and assume isotropic velocity dispersion we get the "streaming model" first introduced by Peebles and used for many years: ξ_s is just a convolution of ξ_r and a velocity PDF. In Fourier space the convolution is a multiplication so Δ_s^2 becomes Δ_r^2 times the FT of the velocity model, usually a Gaussian or an exponential. This is an extremely simple, though not rigorously justified, model. For some further developments, see Reid & White (2011; MNRAS, 417, 1913) or Vlah et al. (2016; JCAP, 12, 007), Vlah & White (2019; JCAP, 03, 007)

Modeling the distortions

We saw above that frequently one uses a simplified model for the redshift space distortions. For example you could multiply the Kaiser factor by a small-scale factor which is either a Gaussian or an exponential. In Fourier space the latter leads to the "dispersion model" (see for example Park et al. 1994 or Peacock & Dodds, 1994, MNRAS 267, 1020)

$$\Delta^2_{\rm red}(k,\mu) = \Delta^2(k)\frac{\left(1+\beta\mu^2\right)^2}{1+k^2\mu^2\sigma^2} \tag{56}$$

where β and σ are parameters to be fit to the data. (Sometimes a model with $[1+k^2\mu^2\sigma^2/2]^2$ is used – note that they agree at low k.) Expressions for the moments can be found in Cole, Fisher & Weinberg (1995; MNRAS, 275, 512). A generalization to the halo model, and a description of why an exponential model is better than a Gaussian, can be found in (MNRAS, 321, 1 (2001); MNRAS 325, 1359 (2001)).

Beyond dispersion or streaming models one needs to turn to numerical simulations. Some work in this direction has been reported in e.g. Hatton & Cole (1999; MNRAS, 310, 113) who find a fit to the quadrupole-to-monopole ratio $Q = Q_{\rm lin}(1-x^{1.22})$ where $x = k/k_1$ and k_1 is a free parameter akin to σ above.

Projected statistics

In principle, if we project our density field along the line-of-sight direction we become immune to redshift space distortions. However in practice we never project to infinite distance and small effects can remain.

We follow Fisher et al. (1994) and define

$$1+\delta_2(\widehat{n}) = \int d\chi\ \phi(s)\left[1+\delta_3(\chi,\chi\widehat{n})\right] \tag{57}$$

where ϕ is a weighting function which integrates to unity and whose argument is the redshift space position, not the real space position. If peculiar velocities are small compared to the width of ϕ we can write

$$\phi(s) \simeq \phi(\chi) + \frac{d\phi}{d\chi} u \tag{58}$$

Fourier transforming and using the Rayleigh expansion of a plane wave we have two terms. The density piece looks like

$$\delta_2(\widehat{n}) = \int d\chi\, \phi(\chi) \int \frac{d^3k}{(2\pi)^3} \delta_3(\chi, \mathbf{k}) \sum_{\ell=0}^{\infty} i^\ell (2\ell+1) j_\ell(k\chi) P_\ell(\widehat{k} \cdot \widehat{n}) \tag{59}$$

with Legendre moment

$$\delta_\ell = i^\ell \int \frac{d^3k}{(2\pi)^3} \delta_3(\mathbf{k}) W_\ell(k) \Rightarrow C_\ell \equiv \langle \delta_\ell \delta_\ell^\star \rangle = 4\pi \int \frac{dk}{k}\, \Delta^2(k) W_\ell^2(k) \tag{60}$$

The velocity term is slightly more complicated

$$\delta_2(\widehat{n}) = \int d\chi\, \frac{d\phi}{d\chi} \int \frac{d^3k}{(2\pi)^3} (-i\beta\mu/k) \delta_3(\chi, \mathbf{k}) \sum_{\ell=0}^{\infty} i^\ell (2\ell+1) j_\ell(k\chi) P_\ell(\widehat{k} \cdot \widehat{n}) \tag{61}$$

Integrating by parts and putting the two terms together using $\ell j_{\ell-1} - (\ell+1) j_{\ell+1} = (2\ell+1) j_\ell'$ one gets

$$\begin{aligned} W_\ell(k) &= \int d\chi\, \phi(\chi) j_\ell(k\chi) + \beta \int d\chi\, \phi(\chi) \left[\frac{2\ell^2 + 2\ell - 1}{(2\ell+3)(2\ell-1)} j_\ell \right. \\ &- \left. \frac{\ell(\ell-1)}{(2\ell+1)(2\ell-1)} j_{\ell-2} - \frac{(\ell+2)(\ell+1)}{(2\ell+1)(2\ell+3)} j_{\ell+2} \right] \end{aligned} \tag{62}$$

The second term, proportional to β, is much smaller than the former when $\ell \gg 1$, but can be significant for narrow redshift shells and small ℓ. In general redshift space distortions enter these expressions down by $k\,\Delta\chi$ compared to the real-space term where $\Delta\chi$ is the width of the "shell".

Light-cone effects

Some people include light-cone effects, the fact that we observe not a constant time slice but along the past light-cone, in redshift space distortions. A nice discussion of the full light-cone formalism in linear perturbation theory is given in Matsubara (2000; ApJ, 535, 1).

The halo model

The halo model lends itself naturally to a RSD treatment, since it conceptually separates the virial motions (1-halo) from the supercluster infall (2-halo). The 2-halo term gets increases by the usual Kaiser factor, and there is a suppression in the bias integral:

$$P^{2-\mathrm{halo}} \to \left(1 + \frac{2}{3}f + \frac{1}{5}f^2\right) B^2(k)P(k) \tag{63}$$

where

$$B = \int f(\nu)d\nu\ b(\nu)\mathcal{R}(k\sigma/2)y(k) \tag{64}$$

while for the 1-halo term

$$P^{1-\mathrm{halo}} \to \frac{1}{(2\pi)^3}\int f(\nu)d\nu\ b(\nu)\frac{M(\nu)}{\bar{\rho}}\mathcal{R}(k\sigma)y^2(k;M) \tag{65}$$

Here $\mathcal{R} \propto \mathrm{erf}$ is the integral of the assumed-Gaussian velocity dispersion. We find the 1-halo term is suppressed relative to the 2-halo term, making the power spectrum closer to linear in redshift as opposed to real space. Note also that the suppression factor at high k is the integral of Gaussians over the mass function, leading to a close-to-exponential profile (as observed).

Cosmological constraints

Under the assumption that the density field has Gaussian statistics and uncorrelated Fourier modes, the Fisher matrix for a set of parameters $\{p_i\}$ is

$$F_{ij} = \frac{1}{2} \int \frac{d^3k}{(2\pi)^3} \left(\frac{\partial \ln P}{\partial p_i}\right) \left(\frac{\partial \ln P}{\partial p_j}\right) V_{\rm eff}\left(\vec{k}\right) \tag{66}$$

where P is the power spectrum and the mode counting is determined by the effective volume

$$V_{\rm eff}\left(\vec{k}\right) \equiv V_0 \left(\frac{\bar{n}P}{1+\bar{n}P}\right)^2 \tag{67}$$

which depends on the geometric volume of the survey, V_0, and the number density, $\bar{n}$, of the tracer. If $\bar{n}$ is high enough then $V_{\rm eff} \simeq V_0$. The constraints are dominated by regions where $\bar{n}P \geq 1$, so it is safe to neglect the higher order (in $\bar{n}^{-1}$) terms which arise assuming that galaxies are a Poisson sample of the underlying density fluctuations.

The simplest model for the observed galaxy distribution is a linear, deterministic, and scale-independent galaxy bias, with redshift space distortions due to super-cluster infall and no observational non-idealities. In this case $P_{\rm obs} \propto \left(b + f\mu^2\right)^2 P_{\rm lin}(k)$ where $P_{\rm lin}$ is the linear theory mass power spectrum in real space, b is the bias and μ the angle to the line-of-sight. The quantity of most interest here is $f \equiv d\ln D/d\ln a$, the logarithmic derivative of the linear growth rate, $D(z)$, with respect to the scale factor $a = (1+z)^{-1}$. In general relativity $f \approx \Omega_{\rm mat}(z)^{0.6}$ while in modified gravity models it can be smaller by tens of percent. Redshift space distortions allow us to constrain f times the normalization of the power spectrum (e.g. $f(z)\sigma_8(z)$), or $dD/d\ln a$.

The derivatives in Eq. (66) are particularly simple

$$\frac{\partial \ln P}{\partial b} = \frac{2}{b + f\mu^2} \quad , \quad \frac{\partial \ln P}{\partial f} = \frac{2\mu^2}{b + f\mu^2} \quad , \quad \frac{\partial \ln P}{\partial \sigma_z^2} = -k^2\mu^2 \tag{68}$$

independent of the shape of the linear theory power spectrum, and hence of the spectral index and transfer function. Since we hold the normalization of the power spectrum fixed for these derivatives, the fractional error on $f(z)\sigma_8(z)$ is equal to that on f in our formalism. The bias and f turn out to be anti-correlated, with a correlation coefficient of $70 - 75\%$, depending on the precise sample.

Using multiple populations of objects of very different biases, and their cross correlations, improves the constraints on b_i and f if you don't have to pay any shot-noise penalty. Usually it doesn't help. Note, what is most of interest is the constraint on f, not the constraint on β. Even perfect knowledge of β doesn't help unless you know b, and you can only know b to precision $\sqrt{2/N}$.