10.6 w(θ) and the angular power spectrum

(a) A 10° by 10° random sky. The initial exercise is to take the random-sky data set provided for example 10.6 and to demonstrate that no signal is present in either the angular correlation function or the power spectrum. The uninformative Figures 1 and 2 do this.

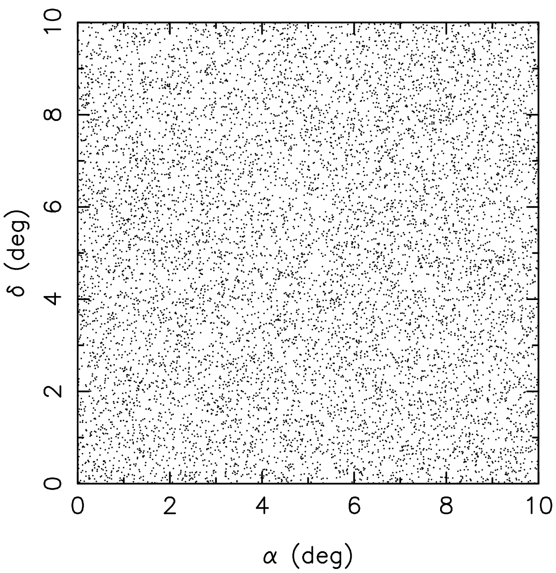

Figure 1: A toy sky, 10° by 10°, with 10000 points distributed at random; this is the data-set provided for exercise 10.6.

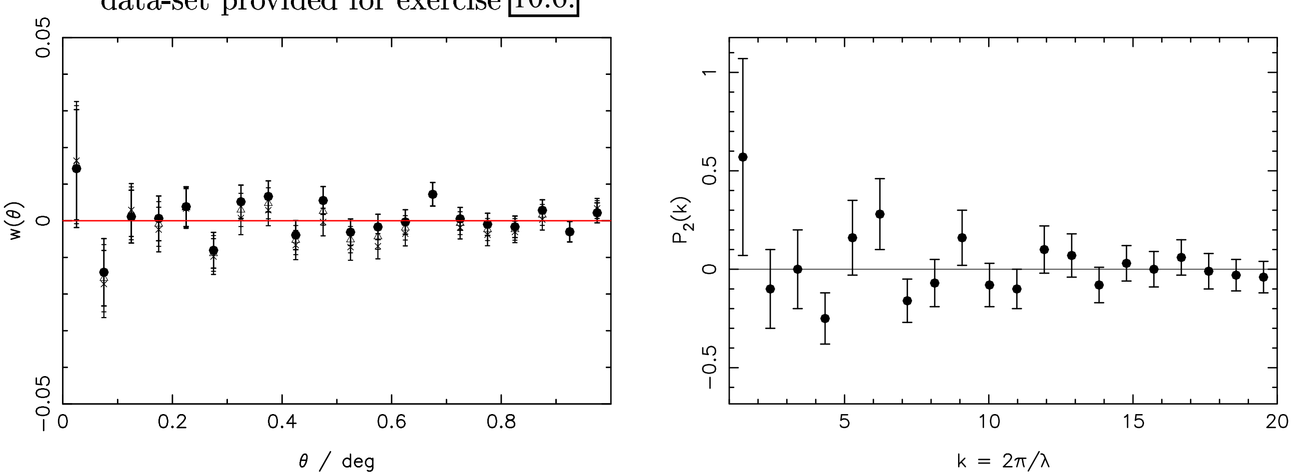

Figure 2: Left: $w(\theta)$, the two-point angular correlation function, computed for the random sky of Figure 1. The different symbols represent the estimators $w1$, $w2$ and $w3$, with results essentially indistinguishable. 10 random skies of 10000 points each were used in the calculation. Right: the power spectrum computed for the same sky. The results (see below) were obtained with a grid of 128 x 128, $k_{\rm min} = 1, k_{\rm max} = 20$, and a total of 20 bins in this interval.

The real object of the exercise was not to demonstrate a foregone conclusion – lack of signal in a random sky – but to ensure that formalism set up to do 2D power-spectrum analysis actually works. (Presumably the earlier exercises have done the same for angular correlation function and counts-in-cells measurements.) Making a 2D spectral analysis work is more difficult than either of these exercises, both conceptually and in practice.

You will want to use a 2D FFT routine provided by many numerical analysis packages; in Numerical Recipes for example, the use of the routine *fourn* is described in simple detail. The hard part is arranging the data input to conform precisely to specification.

The steps in the procedure are as follows:

```
Decide on FFT grid size NX x NY. We're not after much resolution here:  128 x 128 is ample

Specify the given survey dimensions X x Y, here 10 by 10 degrees

Work out the Nyquist frequencies (NX x pi/X) and (NY x pi/Y)

Work out the K-spacings of the grid in each direction:  dkx = 2 x pi/X, dky = 2 x pi/Y

Read in the data

Define the binning for the Pk estimation, start and end values of k, number of bins

Discretize the galaxy distribution on a grid, i.e.  'pixellate' the data onto a grid NX by NY.

Load the data as appropriate for the given 2D FFT routine, and run this

Work out the average value of the amplitudes**2 in radial bins

Subtract the shot noise and renormalize Pk coefficients

Work out the errors for all Pk

Output the results; plot and/or list
```

Loading of the data array for the 2D FFT routine can be done in the same loop as the pixellating.

The results shown here were obtained with a grid of 128 x 128, $k_{\min} = 1, k_{\max} = 20$, and a total of 20 bins in this interval.

(b) A hierarchy of galaxy clusters and clusters of clusters may be constructed by choosing

- a size or spread of sizes for individual galaxies
- a size or spread for clusters
- a size or spread for superclusters

For a patch of sky of say 10° by 10°, these dimensions can be related to reality using Hubble's law $D = cz/H_0$, with galaxies on scales of 10s of kiloparsec, clusters on 100s of kiloparsecs to megaparsecs, and superclusters on 10s of megaparsec scales. A mean redshift depth for the survey needs adopting. Randomly select several supercluster centres first. Then select many cluster centres with say a Gaussian spread in each supercluster; and finally populate the clusters with galaxies, scattered around the cluster centres according to a 2D Gaussians of the chosen cluster dimensions. With 10000 points at your disposal, you are at liberty to vary sizes of superclusters, clusters and the galaxies themselves - and of course to vary the number of galaxies per cluster, the number of clusters per supercluster, and the number of superclusters. Note that with this number of points to contribute to the surface structure, relatively monstrous signals will result in both angular correlations and power spectra, orders of magnitude above those actually observed. Such signal obviates the need to use more than one random sky in the correlation function calculation.

Rather than clusters and superclusters, the modern view sees a continuum of structure with nodes, filaments and voids in a sponge-like topology. The simple view will suffice for purposes of this example.

(c) Adjustment of hierarchy to produce a power-law of ~ -1 in the angular correlation function can be achieved in a number of ways. The sky of Figure 3 shows one of these. It is not difficult to realize the need to have sizes of structures in the hierarchy overlapping, and to have more small-scale entities ('galaxies') than large structures in order to tilt the spectrum in the right sense.

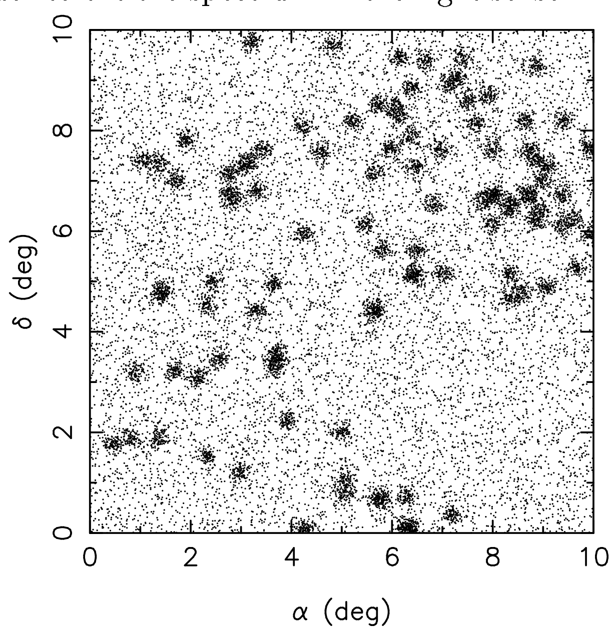

Figure 3: A toy sky, 10° by 10°, with 10000 points distributed at random; this is the data-set provided for exercise 10.6. In addition a further 10000 points have been distributed in 'galaxies' of 100 points in a Gaussian distribution with $\sigma = 0.1°$, 10 each of these in Gaussian 'clusters' of $\sigma = 1.0°$, themselves arranged in 10 Gaussian 'superclusters' distributed at random over the 10° by 10°.

The correlation-function slope shown in Figure 4 is -0.9, hanging in to about 2°, the rough extent of the clusters, at which point the amplitude dies. For a single hierarchy of clustering you can do better by further adjustments of the parameters. Moreover no attempt was made in this example to relate the structure sizes to redhifts as described above; you should also try to do this.

Of course the hierarchy should be continued. The dying of the amplitude after a scale of 2° can be rectified by clustering the superclusters within the 10° by 10° sky, as is necessary; with the current single generation of clustering it is assumed that the superclusters are 'clustered' on scales $> 10°$, whereas in reality there is continuity in

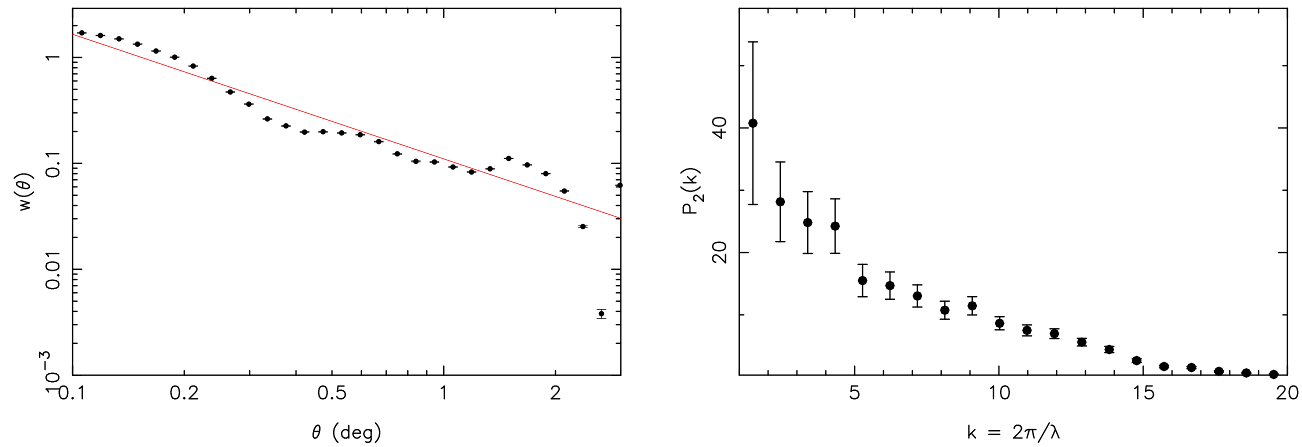

Figure 4: Clustering results for the 'hierarchical' sky described above and shown in Figure 3. Left: $w(\theta)$, the two point angular correlation function, computed using the Landy-Szalay ($w3$) estimator. A single random sky was used as the signal is strong enough that statistical errors are tiny. Right: the power spectrum computed for the same sky. The results, as before, were obtained with a grid of 128 x 128, $k_{\rm min} = 1, k_{\rm max} = 20$, and a total of 20 bins in this interval.

clustering to give the smooth power laws actually observed. See what needs to be done with the present setup and second-generation hierarchy to get a reasonable power law of slope -1 out to 4 or 5°. Explore the number of different ways in which a slope of -1 can be arranged.

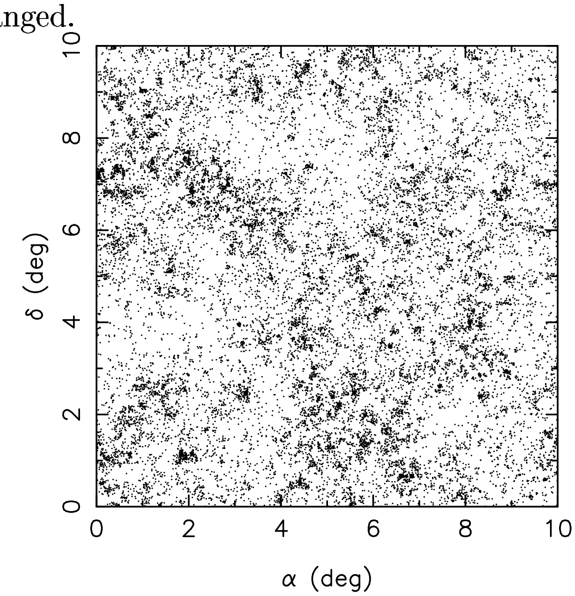

Figure 5: A model sky, 10° by 10°, with 20000 points distributed according to a power-law $w(\theta)$ of slope -1.0, amplitude 0.005.

The implication is that the two-point correlation function is far from a complete description of the real sky distribution, and as discussed in the book, the full picture requires the the three and four-point correlation functions and maybe higher. It is possible to demonstrate this lack of correspondence by approximating a model sky from a given correlation function, using a power-spectrum analysis. One example of this is shown in Figure 5; this is a sky with a two-point correlation function slope of ~ -1 and amplitude 0.005, put in as parameters to a power-law analysis program to provide a representation of such a sky. Figure 6 shows the two-point correlation function and the power spectrum that result.

This is a good demonstration of how incomplete the 2D angular correlation function is in terms of sky description. This model sky bears little resemblance to the blotchy countenance of Figure 3. With voids and filaments appearing, it looks much more like the real sky, with voids and filaments. And both skies have a slope of ~ -1.

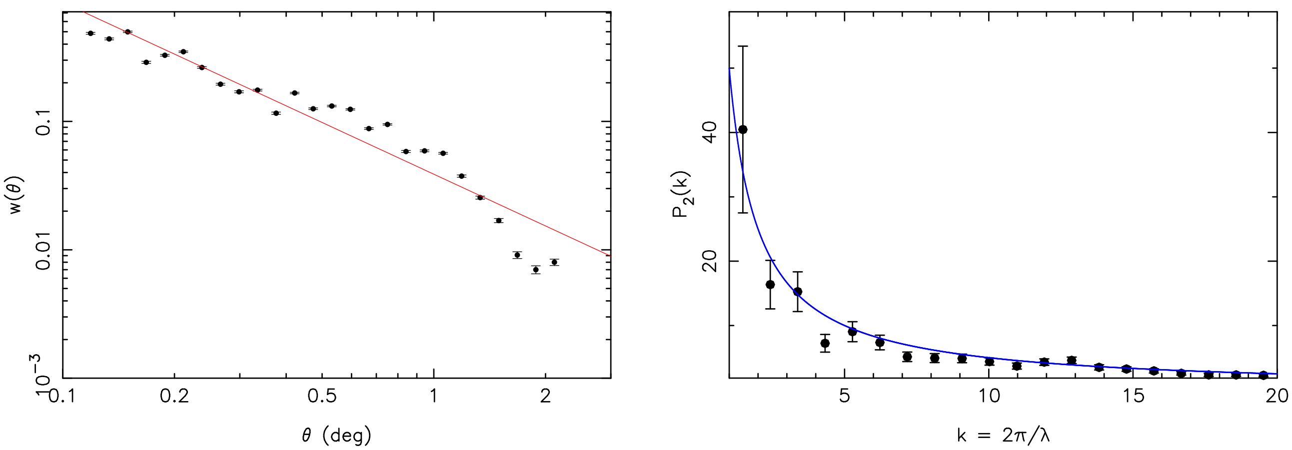

Figure 6: Clustering results for the 'model' sky of 20000 points distributed with a power-spectrum analysis roughly in accord with a power-law two-point correlation function of amplitude 0.005 and slope -1. Left: $w(\theta)$, the two point angular correlation function, computed using the Landy-Szalay ($w3$) estimator. A single random sky was used as the signal is so strong that statistical errors are tiny. Right: the power spectrum computed for the same sky. The results, as before, were obtained with a grid of 128 x 128, $k_{\rm min} = 1, k_{\rm max} = 20$, and a total of 20 bins in this interval. The model power spectrum corresponding to the designed $w(\theta)$ is shown as the solid line.

This sky simulation from a power spectrum is an advanced exercise, to be tried when you are confident that you have your original power-spectrum analysis working well. The basic steps in carrying it out are as follows:

```
As before, specify NX, NY, X, Y, and work out Nyquist frequencies.

Read in amp, slope of the angular correlation function

Generate a digitized version of the correlation function to twice Nyquist frequency

Make an input power spectrum grid for the realization.

Inverse FFT the digitized correlation function

FFT back to get the 2D power spectrum

Generate a lognormal realization of a power spectrum, and then

Populate the sky plane using a Poisson process.
```

There is much involved in getting all this right. Once you have done so, you have good tools to explore issues of signal-to-noise, and of the ranges in which each of the techniques, power-spectrum analysis vs correlation analysis, holds sway. Moreover you will have understood fully the errors, the relation and the transformation between the two sky descriptions.