9.7 Filtering and mean values

A discrete Fourier transform (of which a FFT is an example) doesn't know anything about the Fourier variable; it just returns a string of numbers, which are indexed in a way which we interpret. See Numerical Recipes for details of how to make sense of the output of the FFT. My favourite method for ascribing a frequency / wavenumber scale for the output of a FFT is just to figure out where the Nyquist frequency lies in the array, since I know the Nyquist frequency is 1/ 2 $\times$ sampling interval.

"Integrating" the output of a FFT is thus not obvious. With one definition of the normalization (see §9.2) the level of the of the power spectrum is (on average) the variance on the input. If we integrate by just summing all the numbers in our N-long array that was derived from the FFT, the answer will be $N\sigma^2$. In practice we just *choose* the normalization to give the right answer, which we know from the true (continuous, not discrete) transform. The simplest proof of this is from the Wiener-Khintchine theorem. The integral of the power spectrum is the autocorrelation function at zero lag, which is the variance by definition of the autocorrelation.

Now for the variance on the mean. From the above argument, the output of the FFT must be divided by N to give the correct variance on integration. This means the zero-frequency value of the power spectrum is, on average, σ^2/N – precisely what we expect for the variance on the mean of uncorrelated data. Filtering with a normalized filter does not change this, because a normalized filter must have a Fourier transform that is unity at zero frequency. The zero-frequency of the transform of a filter $f(t)$ is just

$$\int f(t)\, dt = 1$$

because f is normalized. The width of the filter is irrelevant; if we make the filter wider it will have to get less high, to maintain normalization.

A more intuitive way of looking at this is that filtering reduces the noise level *and* increases the correlation between nearby data points. The two effects just cancel out so that the variance on the mean remains the same. The fluctuations are smaller but there are fewer independent points as well.