The Stochastic Geometry of non-Gaussian Fields

Beuman, T.H.

Author: Beuman, Thomas Hubertus
Title: The stochastic geometry of non-gaussian fields
Issue Date: 2015-12-08

5

NON-GAUSSIANITY AND COARSE-GRAINING

When a measurement is made, the resulting picture is usually not pixel perfect. A point source for instance may look like a blob of finite size. This is because, when slightly away from the point source, its presence still leaves its mark. In effect, a measured signal is typically a weighted average of the signal in a small neighborhood of the location in question. This process is known as coarse-graining.

One of the features of our geometric approach is that it does not require accurate measurements of the field in question, since the removal or addition of a singularity or topological defect typically requires a nontrivial operation. Having said that, it is natural to study the effect coarse-graining has on the statistics.

Coarse-graining is also interesting from another point of view: If the scale over which the field is coarse-grained is much larger than its correlation length, the coarse-graining effectively entails taking an average of a large number of independent regions. Therefore, the process of coarse-graining itself may, by virtue of the central limit theorem, give rise to near-Gaussian statistics.

An example in which coarse-graining plays a different but intriguing role is gravitational lensing tomography. Images from distant galaxies are sheared due to mass present between that galaxy and us. This phenomenon is called weak gravitational lensing. Measuring the shear offers a window on the distribution of mass in the universe [25]. The singularities of the shear field for instance correspond to the umbilics of the projected gravitational potential [45]. It is a two-dimensional window though, providing only information about the projected gravitational potential. If one however incorporates the redshift—which can be translated to distance—a three-dimensional picture can be constructed [26]. The entire range of redshifts is divided into bins that can be processed individually. Therefore each bin yields an average of sources from a range of distances, rather than one specific distance. This can be interpreted as an effective source of coarse-graining.

The outline of this chapter is as follows. Section 5.1 sets up the mathematical framework for coarse-graining and discusses its effects. In section 5.2 we derive what the imbalance between maxima and minima looks like in the limit that one coarse-grains over a large scale. This is demonstrated in section 5.3, using a sample case that allows the imbalance to be determined analytically for arbitrary coarse-grain length scales. In section 5.4, the analytical result is compared to numerical simulations. Finally, section 5.5 summarizes our findings.

5.1 COARSE-GRAINING

In general, mathematically, coarse-graining a field $h(\boldsymbol{r})$ can be expressed as

$$\tilde{h}(\boldsymbol{r}) = \int \mathrm{d}^2\boldsymbol{u}\, K(\boldsymbol{u}) h(\boldsymbol{r}+\boldsymbol{u}), \tag{5.1}$$

with $\int \mathrm{d}^2\boldsymbol{u}\, K(\boldsymbol{u}) = 1$. It may be interpreted as a kind of averaging, with the function $K(\boldsymbol{u})$ corresponding to the weight. Coarse-graining may represent blurring as the result of an imprecise measurement of h, for example. Typically then, $K(\boldsymbol{u})$ is largest at $\boldsymbol{u} = \boldsymbol{0}$ and decreases for increasing $|\boldsymbol{u}|$. Ad extremum, $K(\boldsymbol{u}) = \delta(\boldsymbol{u})$ corresponds to no coarse-graining at all, or a perfectly accurate measurement.

When coarse-graining is applied to a Gaussian field $H(\boldsymbol{r})$, the result is (in complex notation):

$$\begin{aligned} \tilde{H}(\boldsymbol{r}) &= \int \mathrm{d}^2\boldsymbol{u}\, K(\boldsymbol{u}) \sum_{\boldsymbol{k}} A(k) e^{i(\boldsymbol{k}\cdot(\boldsymbol{r}+\boldsymbol{u})+\phi_{\boldsymbol{k}})} \\ &= \sum_{\boldsymbol{k}} A(k) \left(\int \mathrm{d}^2\boldsymbol{u}\, K(\boldsymbol{u}) e^{i\boldsymbol{k}\cdot\boldsymbol{u}} \right) e^{i(\boldsymbol{k}\cdot\boldsymbol{r}+\phi_{\boldsymbol{k}})} \end{aligned} \tag{5.2}$$

It is thus easily seen that coarse-graining only affects the amplitude spectrum of H, but not its Gaussianity.

Consider now what happens when applied to a non-Gaussian field h. Let us set $K(\boldsymbol{r}) = f(r/l)/l^2$, where l is a length scale that controls the size of the coarse-graining, while $f(x)$ is a dimensionless function that goes to zero for $x \gg 1$. Let ξ be (a measure of) the correlation length of h, so that $\langle h(\boldsymbol{r})h(\boldsymbol{r}+\boldsymbol{R})\rangle$ vanishes when $|\boldsymbol{R}| \gg \xi$. If $l \gg \xi$, the coarse-graining effectively entails taking the average over a large number (of the order of $(l/\xi)^2$) of independent regions, causing $\tilde{h}$ to acquire Gaussian characteristics on account of the central limit theorem.

There is a link between coarse-graining and the deterministic Kardar-Parisi-Zhang (KPZ) equation that was investigated in chapter 4. This equation is a diffusion equation that reads

$$\frac{\partial h}{\partial t} = \nu \nabla^2 h + \frac{\lambda}{2}(\nabla h)^2. \tag{5.3}$$

Following the substitution $u = \exp((\lambda/2\nu)h)$, this transforms into $\frac{\partial u}{\partial t} = \nu \nabla^2 u$, with the solution

$$u(\boldsymbol{r}, t) = \int \mathrm{d}^2\tilde{\boldsymbol{r}}\, \frac{1}{4\pi\nu t} e^{-\frac{(\boldsymbol{r}-\tilde{\boldsymbol{r}})^2}{4\nu t}} u(\tilde{\boldsymbol{r}}, 0). \tag{5.4}$$

This relation has the same structure as that of the formula for coarse-graining: $u(\boldsymbol{r}, 0)$ can be identified as the original field and $l = \sqrt{\nu t}$ as the coarse-graining scale. The dimensionless coarse-graining function is therefore

$$f(\rho) = \frac{1}{4\pi} e^{-\rho^2/4}. \tag{5.5}$$

In summary, coarse-graining a non-Gaussian field over a large scale causes it to obtain Gaussian characteristics. A Gaussian field remains Gaussian, regardless of the scale over which it is coarse-grained.

5.2 LARGE SCALE LIMIT

In this section, the consequences of coarse-graining a non-Gaussian field h are investigated, in the limit that the coarse-graining scale becomes very large, as compared to the typical correlation length of the field. In particular, focus is put on the consequences for the densities of maxima and minima.

For a Gaussian field, the densities of maxima and minima are the same due to symmetry. For a non-Gaussian field, this is in general not the case. In chapter 4, a general expression was derived for the relative difference between the two densities for a perturbed Gaussian field, up to first order in the perturbation (see eq. (4.15)),

$$\Delta n \equiv \frac{n_{max} - n_{min}}{n_{max} + n_{min}} = \sqrt{\frac{6}{\pi\alpha}} \left(\frac{4}{3}\frac{\beta}{\sigma} + \frac{4}{9}\frac{\gamma}{\alpha} - \frac{10}{27}\frac{\delta}{\alpha} \right). \tag{5.6}$$

The parameters are second- and third-order correlations (see eq. (4.6)),

$$\sigma = \langle h_z h_{z^*} \rangle = C_2(h_z, h_{z^*}), \tag{5.7a}$$
$$\alpha = \langle h_{zz^*}^2 \rangle = C_2(h_{zz^*}, h_{zz^*}), \tag{5.7b}$$
$$\beta = \langle h_z h_{z^*} h_{zz^*} \rangle = C_3(h_z, h_{z^*}, h_{zz^*}), \tag{5.7c}$$
$$\gamma = \langle h_{zz^*}^3 \rangle = C_3(h_{zz^*}, h_{zz^*}, h_{zz^*}), \tag{5.7d}$$
$$\delta = \langle h_{zz} h_{z^*z^*} h_{zz^*} \rangle = C_3(h_{zz}, h_{z^*z^*}, h_{zz^*}). \tag{5.7e}$$

Here the subscripts denote partial differentiation, with $\partial_z = \frac{1}{2}(\partial_x - i\partial_y)$ and $\partial_{z^*} = \frac{1}{2}(\partial_x + i\partial_y)$, as before (eq. (3.9)). In each cumulant, the variables are all taken at the same point $\boldsymbol{r}$, e.g. $\sigma = C_2(h_z(\boldsymbol{r})h_{z^*}(\boldsymbol{r}))$. Due to homogeneity, the choice of $\boldsymbol{r}$ is irrelevant.

Recall that, in general, correlations and cumulants are not the same; in an identity (see eqs. (3.4) and (3.5)), there would also be terms with correlations / cumulants of a single variable, but those are zero in this case (compare eq. (4.7)).

For the coarse-grained field $\tilde{h}$, the cumulants can be calculated in the following way (using β as an example):

$$\begin{aligned}\beta &= C_3(\tilde{h}_z, \tilde{h}_{z^*}, \tilde{h}_{zz^*}) \\ &= \partial_{z_1}\partial_{z_2^*}\partial_{z_3}\partial_{z_3^*} C_3(\tilde{h}(\boldsymbol{r_1})\tilde{h}(\boldsymbol{r_2})\tilde{h}(\boldsymbol{r_3}))\Big|_{\boldsymbol{r_1}=\boldsymbol{r_2}=\boldsymbol{r_3}}.\end{aligned} \tag{5.8}$$

Therefore, we see that the main ingredients for calculating the cumulants are $C_2(\tilde{h}(\boldsymbol{r_1}), \tilde{h}(\boldsymbol{r_2}))$ (for σ and α) and its third-order equivalent (for β, γ and δ).

Let ξ be a measure of the correlation length of h, in the sense that $h(\boldsymbol{r_1})$ and $h(\boldsymbol{r_2})$ can be said to be roughly uncorrelated when $|\boldsymbol{r_1} - \boldsymbol{r_2}| > \xi$. If $l \gg \xi$, then $\tilde{h}$ is nearly Gaussian (on account of the central limit theorem), so one can find the imbalance of maxima and minima using the approximation.

The second-order cumulant of $\tilde{h}$ can be expressed as

$$\begin{aligned}C_2(\tilde{h}(\boldsymbol{r_1}), \tilde{h}(\boldsymbol{r_2})) = \frac{1}{l^4}\iint \mathrm{d}^2\boldsymbol{R_1}\mathrm{d}^2\boldsymbol{R_2}\, C_2(h(\boldsymbol{R_1}), h(\boldsymbol{R_2})) \\ \times f\Big(\frac{\boldsymbol{R_1}-\boldsymbol{r_1}}{l}\Big) f\Big(\frac{\boldsymbol{R_2}-\boldsymbol{r_2}}{l}\Big).\end{aligned} \tag{5.9}$$

Since $C_2(h(\boldsymbol{R_1}), h(\boldsymbol{R_2}))$ is only appreciable when $|\boldsymbol{R_1} - \boldsymbol{R_2}| < \xi \ll l$, the approximation $f((\boldsymbol{R_2} - \boldsymbol{r_2})/l) \approx f((\boldsymbol{R_1} - \boldsymbol{r_2})/l)$ can be applied. This gives

$$C_2(\tilde{h}(\boldsymbol{r_1}), \tilde{h}(\boldsymbol{r_2})) = \frac{1}{l^4} \iint \mathrm{d}^2\boldsymbol{R_1} \mathrm{d}^2\boldsymbol{a}\, C_2(h(\boldsymbol{0}), h(\boldsymbol{a})) \\ \times f\Big(\frac{\boldsymbol{R_1} - \boldsymbol{r_1}}{l}\Big) f\Big(\frac{\boldsymbol{R_1} - \boldsymbol{r_2}}{l}\Big), \tag{5.10}$$

where $\boldsymbol{a} = \boldsymbol{R_2} - \boldsymbol{R_1}$ and use was made of the homogeneity of h. This integration can now be split into two parts:

$$\begin{aligned} &C_2(\tilde{h}(\boldsymbol{r_1}), \tilde{h}(\boldsymbol{r_2})) \\ &\quad = \frac{1}{l^4} \int \mathrm{d}^2\boldsymbol{a}\, C_2(h(\boldsymbol{0}), h(\boldsymbol{a})) \int \mathrm{d}^2\boldsymbol{R_1} f\Big(\frac{\boldsymbol{R_1} - \boldsymbol{r_1}}{l}\Big) f\Big(\frac{\boldsymbol{R_1} - \boldsymbol{r_2}}{l}\Big) \\ &\quad = \frac{1}{l^2} I_2 K_2\big(\frac{\boldsymbol{r_1}}{l}, \frac{\boldsymbol{r_2}}{l}\big), \end{aligned} \tag{5.11}$$

where

$$I_2 \equiv \int \mathrm{d}^2\boldsymbol{a}\, C_2(h(\boldsymbol{0}), h(\boldsymbol{a})), \tag{5.12}$$

and

$$K_2(\boldsymbol{\rho_1}, \boldsymbol{\rho_2}) \equiv \int \mathrm{d}^2\boldsymbol{v}\, f(\boldsymbol{v} - \boldsymbol{\rho_1}) f(\boldsymbol{v} - \boldsymbol{\rho_2}). \tag{5.13}$$

For the third-order correlation an analogous derivation can be made. The result is

$$C_3(\tilde{h}(\boldsymbol{r_1}), \tilde{h}(\boldsymbol{r_2}), \tilde{h}(\boldsymbol{r_3})) = \frac{1}{l^4} I_3 K_3\big(\frac{\boldsymbol{r_1}}{l}, \frac{\boldsymbol{r_2}}{l}, \frac{\boldsymbol{r_3}}{l}\big), \tag{5.14}$$

with

$$I_3 \equiv \iint \mathrm{d}^2\boldsymbol{a} \mathrm{d}^2\boldsymbol{b}\, C_3(h(\boldsymbol{0}), h(\boldsymbol{a}), h(\boldsymbol{b})), \tag{5.15}$$

and

$$K_3(\boldsymbol{\rho_1}, \boldsymbol{\rho_2}, \boldsymbol{\rho_3}) \equiv \int \mathrm{d}^2\boldsymbol{v}\, f(\boldsymbol{v} - \boldsymbol{\rho_1}) f(\boldsymbol{v} - \boldsymbol{\rho_2}) f(\boldsymbol{v} - \boldsymbol{\rho_3}). \tag{5.16}$$

Note in particular that I_2 and I_3 depend on h only, whereas K_2 and K_3 depend only on f. Also note that none of these terms depends on l.

For the correlations, the z- and z^*-derivatives, as used in eq. (5.8), act only on $K_2(\frac{\boldsymbol{r_1}}{l}, \frac{\boldsymbol{r_2}}{l})$ and $K_3(\frac{\boldsymbol{r_1}}{l}, \frac{\boldsymbol{r_2}}{l}, \frac{\boldsymbol{r_3}}{l})$. Each derivative introduces a factor

$1/l$ as a result of the chain rule. It can therefore already be deduced how the relevant correlations scale with l:

$$\sigma \sim l^{-4}, \qquad \alpha \sim l^{-6}, \qquad \beta \sim l^{-8}, \qquad \gamma, \delta \sim l^{-10}, \tag{5.17}$$

and therefore $\Delta n \sim 1/l$. More precisely, we have the following:

$$\lim_{l/\xi \to \infty} \Delta n = \frac{c_h c_f}{l}, \tag{5.18}$$

where $c_h = I_3/I_2^{3/2}$ is a parameter that depends on the statistics of h only, and c_f is a parameter that depends on the coarse-graining function f only.

5.3 EXAMPLE

5.3.1 *Large scale limit*

As an example, let us consider the coarse-grain function from eq. (5.5). This gives

$$K_2(\boldsymbol{\rho_1}, \boldsymbol{\rho_2}) = \frac{1}{8\pi} e^{-\frac{1}{8}(\boldsymbol{\rho_1} - \boldsymbol{\rho_2})^2}, \tag{5.19a}$$

$$K_3(\boldsymbol{\rho_1}, \boldsymbol{\rho_2}, \boldsymbol{\rho_3}) = \frac{1}{48\pi^2} e^{-\frac{1}{4}\left(\boldsymbol{\rho_1}^2 + \boldsymbol{\rho_2}^2 + \boldsymbol{\rho_3}^2 - \frac{1}{3}(\boldsymbol{\rho_1} + \boldsymbol{\rho_2} + \boldsymbol{\rho_3})^2\right)}. \tag{5.19b}$$

Applying the method as exemplified in eq. (5.8), we obtain

$$\beta = -\frac{I_3}{6912\pi l^8}. \tag{5.20}$$

Calculating all relevant cumulants ultimately leads to

$$\Delta n = \frac{2^{9/2}}{3^{11/2}\pi} \frac{I_3}{I_2^{3/2}} \frac{1}{l}, \tag{5.21}$$

for the KPZ-inspired Gaussian coarse-grain function eq. (5.5).

To get the parameter c_h, we use a non-Gaussian field of the form $h = H + \varepsilon H^2$, where H is a Gaussian field with a given two-point correlation function $\langle H(\boldsymbol{r_1}) H(\boldsymbol{r_2}) \rangle$, and ε is a small constant. The two-point correlation function of h differs from that of H only in second order of ε; this difference will be ignored.

The third-order cumulant of h is zero in leading order, since H is Gaussian, but in first order we find

$$\begin{aligned} C_3(h(\boldsymbol{r_1}), h(\boldsymbol{r_2}), h(\boldsymbol{r_3})) = C_3(\varepsilon H(\boldsymbol{r_1})^2, H(\boldsymbol{r_2}), H(\boldsymbol{r_3})) \\ + C_3(H(\boldsymbol{r_1}), \varepsilon H(\boldsymbol{r_2})^2, H(\boldsymbol{r_3})) \\ + C_3(H(\boldsymbol{r_1}), H(\boldsymbol{r_2}), \varepsilon H(\boldsymbol{r_3})^2). \end{aligned} \tag{5.22}$$

This can be expanded with the help of Wick's theorem, e.g.

$$\begin{aligned} &C_3(H(\boldsymbol{r_1})^2, H(\boldsymbol{r_2}), H(\boldsymbol{r_3})) \\ &\quad = \langle H(\boldsymbol{r_1})^2 H(\boldsymbol{r_2}) H(\boldsymbol{r_3})\rangle - \langle H(\boldsymbol{r_1})^2\rangle\langle H(\boldsymbol{r_2}) H(\boldsymbol{r_3})\rangle \\ &\quad = 2\langle H(\boldsymbol{r_1}) H(\boldsymbol{r_2})\rangle\langle H(\boldsymbol{r_1}) H(\boldsymbol{r_3})\rangle, \end{aligned} \tag{5.23}$$

Therefore, the result is

$$\begin{aligned} &C_3(h(\boldsymbol{r_1}), h(\boldsymbol{r_2}), h(\boldsymbol{r_3})) \\ &\quad = 2\varepsilon\Big(\langle H(\boldsymbol{r_1}) H(\boldsymbol{r_2})\rangle\langle H(\boldsymbol{r_1}) H(\boldsymbol{r_3})\rangle \\ &\qquad + \langle H(\boldsymbol{r_2}) H(\boldsymbol{r_3})\rangle\langle H(\boldsymbol{r_2}) H(\boldsymbol{r_1})\rangle \\ &\qquad + \langle H(\boldsymbol{r_3}) H(\boldsymbol{r_1})\rangle\langle H(\boldsymbol{r_3}) H(\boldsymbol{r_2})\rangle\Big) \end{aligned} \tag{5.24}$$

If we consider the Gaussian two-point correlation function

$$\langle H(\boldsymbol{r_1}) H(\boldsymbol{r_2})\rangle = e^{-\frac{1}{2}k_0^2(\boldsymbol{r_1}-\boldsymbol{r_2})^2}, \tag{5.25}$$

which corresponds to the spectrum $A(k)^2 \sim \exp(-k^2/(2k_0^2))$, we get

$$I_2 = \int \mathrm{d}^2\boldsymbol{a}\, C_2(0, \boldsymbol{a}) = \int \mathrm{d}^2\boldsymbol{a}\, \langle H(0) H(\boldsymbol{a})\rangle = \frac{2\pi}{k_0^2}, \tag{5.26}$$

and

$$\begin{aligned} I_3 &= \iint \mathrm{d}^2\boldsymbol{a}\mathrm{d}^2\boldsymbol{b}\, C_3(h(\boldsymbol{0}), h(\boldsymbol{a}), h(\boldsymbol{b})) \\ &= \iint \mathrm{d}^2\boldsymbol{a}\mathrm{d}^2\boldsymbol{b}\, 2\varepsilon\Big(e^{-\frac{1}{2}k_0^2(\boldsymbol{a}^2+\boldsymbol{b}^2)} + e^{-\frac{1}{2}k_0^2(\boldsymbol{a}^2+(\boldsymbol{a}-\boldsymbol{b})^2)} \\ &\qquad + e^{-\frac{1}{2}k_0^2(\boldsymbol{b}^2+(\boldsymbol{a}-\boldsymbol{b})^2)}\Big) \\ &= \frac{24\pi^2\varepsilon}{k_0^4}. \end{aligned} \tag{5.27}$$

This gives $c_h = I_3/I_2^{3/2} = 6\sqrt{2\pi}\varepsilon/k_0$.

Combined with the coarse-grain function as given above, we thus get

$$\Delta n \to \frac{c_f c_h}{l} = \frac{64}{81\sqrt{3\pi}k_0}\frac{\varepsilon}{l}. \tag{5.28}$$

5.3.2 *Analytic result*

The separation of the dependence on f and h, as displayed in eq. (5.18), is only valid in the limit of $l \gg \xi$. In general however, f and h can no longer be treated separately. Due to the complexity of the integrals involved, it is only in very specific cases possible to calculate Δn analytically, for arbitrary l. Not coincidentally, the f and h chosen in the previous section allow precisely this.

For example, the exact expression for β is

$$\begin{aligned}\beta &= \langle \tilde{h}_z \tilde{h}_{z^*} \tilde{h}_{zz^*} \rangle \\ &= \partial_{z_1} \partial_{z_2^*} \partial_{z_3} \partial_{z_3^*} \iiint \mathrm{d}^2\boldsymbol{u_1} \mathrm{d}^2\boldsymbol{u_2} \mathrm{d}^2\boldsymbol{u_3} \Big(K(\boldsymbol{u_1}) K(\boldsymbol{u_2}) K(\boldsymbol{u_3}) \\ &\qquad \times \langle h(\boldsymbol{r_1}+\boldsymbol{u_1}) h(\boldsymbol{r_2}+\boldsymbol{u_2}) h(\boldsymbol{r_3}+\boldsymbol{u_3}) \rangle \Big) \Big|_{\boldsymbol{r_1}=\boldsymbol{r_2}=\boldsymbol{r_3}}. \end{aligned} \tag{5.29}$$

The three-point correlation can be expanded in the same way as before (see eqs. (5.22) and (5.23)). The final result is

$$\beta = -\frac{k_0^4 \varepsilon}{2(1+2k_0^2 l^2)^2 (1+6k_0^2 l^2)^2}, \tag{5.30}$$

Determining and combining all the correlations gives the following result, which is exact with respect to l but still perturbative with respect to ε:

$$\Delta n = \frac{64 a^3 (1+4a)^{7/2} \varepsilon}{\sqrt{3\pi}(1+2a)^3 (1+6a)^4}, \tag{5.31}$$

where $a \equiv k_0^2 l^2$. In the limit of large l ($a \gg 1$) we find that it matches the perturbative result eq. (5.28).

One may also note that eq. (5.31) matches the result from the deterministic KPZ equation for the Gaussian spectrum (eq. (4.32)), following the substitution $\nu t \to l^2$.

5.4 NUMERICAL TESTS

5.4.1 *Setup*

The validity of eq. (5.31) was checked using computer simulations. The setup of the simulations and the identification of the extrema is similar to the process used in previous chapters and outlined in appendix A: For

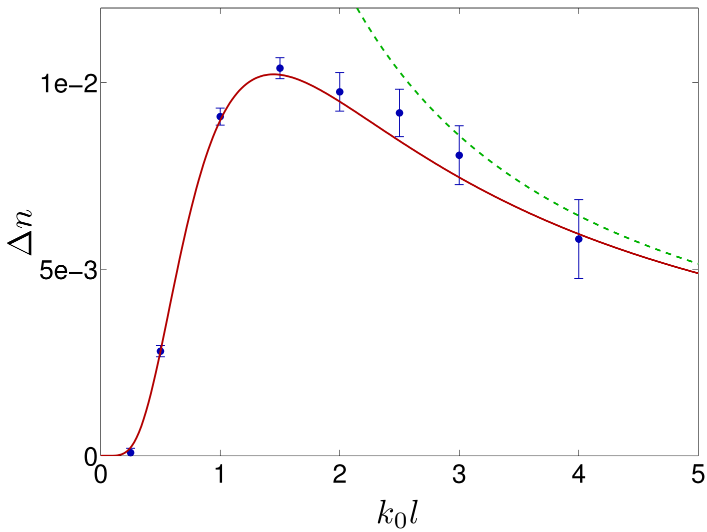

Figure 5.1: The imbalance between maxima and minima Δn for a field $h = H + \varepsilon H^2$, where H is a Gaussian field with a Gaussian power spectrum (with typical wavelength k_0^{-1}), coarse-grained with a Gaussian function $f(r) = \exp(r^2/(4l^2))$. The solid line is the exact result eq. (5.31) (perturbative in ε but not in l), while the data points stem from simulations. The dashed line is the theoretical result for large l, eq. (5.28).

each data point, corresponding to a particular value of l, thousands of Gaussian fields $H(\boldsymbol{r})$ were generated, following eq. (1.9). To each field, the perturbation εH^2 was added to it, with $\varepsilon = 0.1$. This new field h was then coarse-grained, after which the extrema were identified.

5.4.2 *Results*

Fig. 5.1 shows the theoretical result, as well as results from simulations at various values of l. As can be seen, there is an excellent agreement between the two. Also shown is the prediction of eq. (5.28), illustrating the large l limit, which matches well for $k_0 l \gg 1$.

An interesting point is that, for no coarse-graining at all, the imbalance is very close to zero. This general feature of local perturbations of the type $h = H + f(H)$ was already established in detail in chapter 2. Measuring the imbalance between maxima and minima thus does not reveal the non-Gaussianity of h. However, it is clear from fig. 5.1 that coarse-graining may significantly increase the imbalance to measurable values, thereby not only granting the possibility of detecting non-Gaussianity, but also potentially identifying the size and type of the perturbation.

5.4.3 *Large scale coarse-graining*

It is difficult to test the formula for a large coarse-graining (eq. (5.18)) accurately in a numerical setting, since the size of the system is limited.

As explained in appendix A, the periodic boundary conditions are enforced by only using waves with wave vectors $\boldsymbol{k}$ for which the components are multiples of $\frac{2\pi}{L}$, where L is the system size. As long as L is large, this quantization has a high enough resolution to be of no significant source of error.

Now consider what happens when it is coarse-grained. We already saw in eq. (5.2) that, effectively, its amplitude spectrum changes:

$$\tilde{A}(k) = A(k)\left(\int \mathrm{d}^2\boldsymbol{u}\, K(\boldsymbol{u})e^{i\boldsymbol{k}\cdot\boldsymbol{u}}\right) = A(k)\left(\int \mathrm{d}^2\boldsymbol{\rho}\, f(\rho)e^{il\boldsymbol{k}\cdot\boldsymbol{\rho}}\right). \quad (5.32)$$

In the limit that $lk \gg 1$, the phase factor causes the integral to vanish. Hence, for large l, only the waves with small wave vector k (in the order of $1/l$ or less) prevail. This is the technical justification of the statement that coarse-graining causes a field to become smoother, and thus dominated by long waves.

However, in combination with the periodic boundary conditions, this means that, in the case that l becomes comparable to L, there are only a few wave vectors left that are of importance from the coarse-graining point of view. The accuracy with which the simulated coarse-grained field represents an actual field thus becomes compromised. Therefore, L should be larger than l. Increasing L however naturally increases computation time. As a result, probing large coarse-grain scales indirectly requires a lot of computation time, making it difficult to properly explore the regime in which the imbalance Δn decays as $1/l$.

5.5 CONCLUSIONS

Coarse-graining a non-Gaussian field has the effect of giving it Gaussian characteristics, as the coarse-graining scale goes to infinity. More precisely, when this scale l is significantly larger than the correlation length of the field, the imbalance between maxima and minima (which is zero for Gaussian fields) scales as $1/l$. The corresponding constant factor can be written as the product of two independent scalars: One depends on the field only, whereas the other depends on the coarse-graining function only.

Coarse-graining a signal on purpose can also be useful, because the imbalance between maxima and minima depends on the length scale of the coarse-graining for a non-Gaussian field. For example, locally perturbed fields, such as $h = H + \varepsilon H^2$, where H is Gaussian, do not show a significant imbalance between maxima and minima (if the resolution is perfect). However, coarse-graining—which would not produce an effect for Gaussian fields—creates an imbalance allowing ε to be measured. In general, coarse-graining by various amounts can give a multitude of data that can be used to shed light on some unknown parameters of the perturbation.