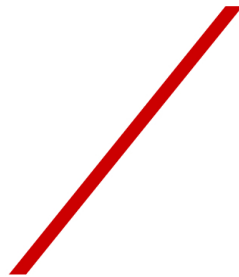


UNIVERSITY OF GRONINGEN

BACHELOR THESIS

Phase correlations in cosmic structure formation



university of
groningen

Author:
Roi Kugel

Supervisor:
Prof. dr. Rien van de Weijgaert
Second examiner:
Dr. Pratika Dayal

Abstract

This thesis looks at using the Fourier phases as a method of describing cosmic structure formation. In previous efforts it was pointed out that structure appears in the Fourier phases as expansion factor increases. We identify the cosmological structure responsible for the majority of the phase structure and provide a quantity that quantifies the structure found in the Fourier phases. This quantity, and thus also the phase structure, relates to the very extremes of the density field. For cosmological simulations these are mostly overdensities. We also look at using phase shifts and provide a way of comparing phase shift results for different models. Both models were tested for invariance and are independent of initial coordinates. We also provide a way of making any phase field invariant under translations by use of the Taylor invariant.

Contents

1	Acknowledgements	3
2	Introduction	4
3	Fourier Transforms and Random Fields	6
3.1	Fourier Transforms	6
3.1.1	Fourier filters	6
3.1.2	Power spectra	9
3.1.3	Fourier phases	10
3.1.4	Visualisation of phases	11
3.2	Gaussian random fields	12
3.2.1	Definition	12
3.2.2	Gaussian fields and the power spectrum	13
4	Cosmological Background	14
4.1	Characterising Densities and times	14
4.1.1	Densities	14
4.1.2	Time evolution	14
4.2	Cosmic structure formation	15
4.2.1	Initial conditions and Gaussian random fields	15
4.2.2	The phase of linear growth	15
4.2.3	The phase of non-linear growth	17
4.2.4	The final state of the universe	17
5	Numerical Methods	18
5.1	Numerical Fourier transforms	18
5.1.1	Fourier phases and amplitudes	18
5.1.2	Numerical power spectrum	18
5.1.3	Numerical Phase spectrum	19
5.2	Generation of Gaussian fields	19
5.3	Particle locations to density fields	20
5.4	The Phase gradient	21
5.5	σ_{∇}	21

6	Models and Simulations	22
6.1	Images	22
6.2	Gaussian Fields	22
6.3	Voronoi models	23
6.3.1	Specific Voronoi models used	23
6.3.2	N-Body Simulations	27
7	Phase Invariants	30
7.1	The Taylor invariant	31
7.1.1	The center of mass	31
7.1.2	A center of mass in a periodic system	31
7.2	The Taylor invariant of a static image	32
8	Data analysis	34
8.1	Fourier phases of the cosmic web	34
8.1.1	Origin of phase structure	34
8.2	Quantifying the phase structure	36
8.3	Quantifying σ_{∇}	37
8.3.1	σ_{∇} for N-body simulations	39
8.4	Phase shifts to track cosmic structure formation	40
8.5	The cosmic Taylor invariant	42
8.5.1	The invariance of σ_{∇}	42
8.5.2	The invariance of $ \Delta\phi $	43
9	Conclusion	45
A	Appendix	48
A.1	Numerical Mathematical tools	48
A.1.1	Romberg integration	48
A.1.2	Numerical Gradient	48
A.2	Residual due to round of errors	49
A.3	Voronoi 3D Boxes	50
A.3.1	Voronoi kinematic model experiment	50
A.3.2	Voronoi elements experiment	51
A.3.3	Increasing cells experiment	52
A.3.4	Explosion model experiment	53
A.4	Voronoi Slices	54
A.4.1	Voronoi kinematic model experiment	54
A.4.2	Voronoi elements experiment	55
A.4.3	Increasing cells experiment	56
A.4.4	Explosion model experiment	57

1 Acknowledgements

I would first like to thank my supervisor Rien van de Weygaert for his continuous enthusiasm, support and for allowing me to work on such a interesting topic which such a large amount of freedom. I would also like the thank Folkert Nobels and Ewout Wempe for greatly reducing the amount of computational time and for helping me create very insightful videos. Finally I would like to thank Jasper Steringa for keeping me company.

2 Introduction

Structure is formed after the universe leaves its initially uniform state. Gravitational collapse allows for structure formation to cascade upward. Stars are formed first, followed by galaxies and eventually structure is formed at larger and larger scales. This multiscale structure found in cosmic density fields is conventionally described by using Fourier analysis and more specifically the Fourier power spectrum. The power spectrum is taken from the Fourier amplitudes while the phases are discarded. The power spectrum provides the image with intensities as a function of scale but it is the phases that provide the image with structure. An example of this can be seen in Figure 1:

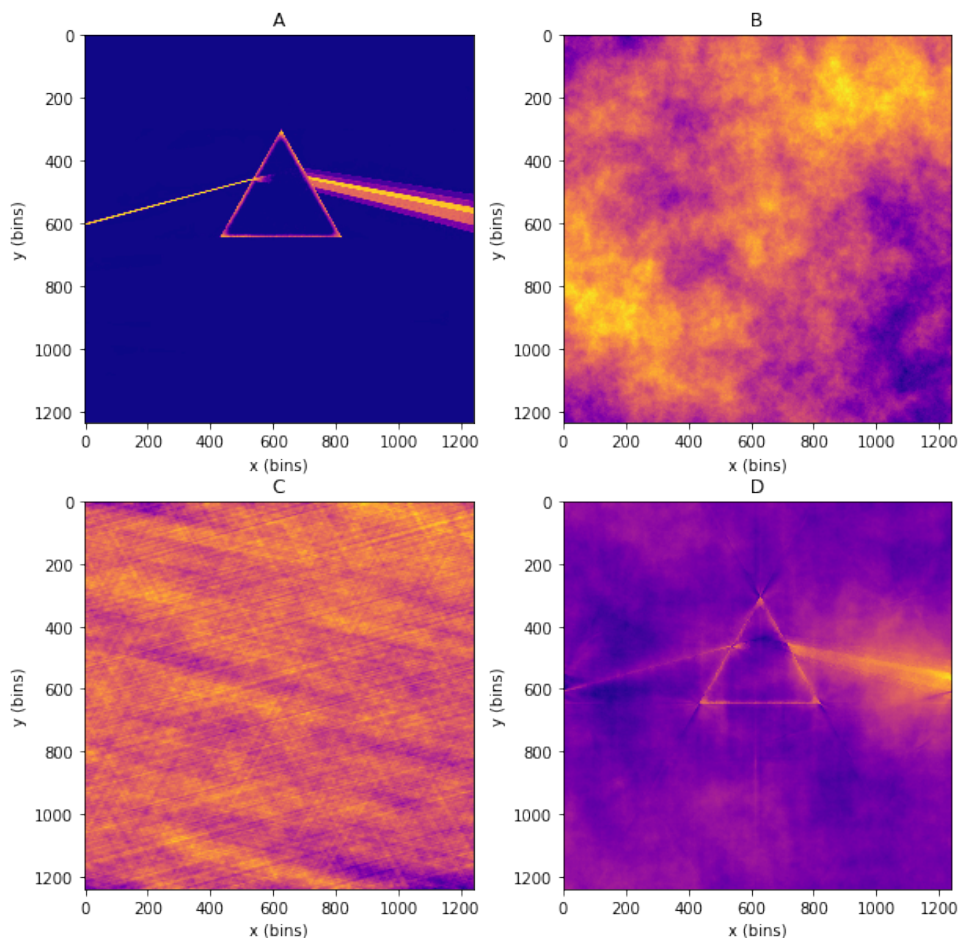


Figure 1: (A) Has the image of a prism of the famous album The dark side of the moon (Pink Floyd) by Hardie, Thorgerson, and Powell (1973), (B) is a Gaussian random field with a power law power spectrum with $n=-2$. (C) is the result of combining the Fourier phases of (B) with the amplitudes of (A), (D) is the result of combining the amplitudes of (B) with the phases of (A)

When the phases are changed to the random phases specific of a Gaussian random field, all structure is lost. In the opposite case, where the phases are conserved and the amplitudes are that of a Gaussian random field, the structure is mostly conserved and only the relative intensities are changed due to the difference in the power spectrum of both images. The phases are strictly non-random and provide the resulting field with structure. The phases hold the structural information and analysing them should relate back to structure formation. The main problem with analysing the Fourier phases is that they are not invariant under translations and the value of the phases changes during translations. In a cosmological setting the coordinate system should not matter. There have been important contributions that look into how the phases can be used as a measure for cosmic structure formation. Coles and Chiang (2000) give a

look at a possible way to analyse phases using a quantity called phase entropy which is further explained by Chiang and Coles (2000). This uses mainly the gradient of the Fourier phases as a guide for structure. B. S. Ryden and Gramann (1991) instead look at how the phase field moves away from the initial conditions at different scales. In this thesis a Fourier invariant is discussed and tested in a physical system. A new way of looking at phase structure is also explained and a new quantity is defined that can quantify the amount of phase structure. Finally a repeat of the analysis done by B. S. Ryden and Gramann (1991) is done and an effort is made to quantify the results in a way that can be compared to other simulations.

3 Fourier Transforms and Random Fields

In this section the mathematical formalism of Fourier transforms and their uses is explained. A way of visualising phase fields is explained and the definition of Gaussian random fields is discussed.

3.1 Fourier Transforms

The analysis used is based around and relies solely on a mathematical transformation from normal space, often referred to as real space, to k space or wave space. The Fourier transform is defined as:

$$\hat{f}(\vec{k}) = \int_{-\infty}^{\infty} d\vec{x} f(\vec{x}) e^{i\vec{x} \cdot \vec{k}}. \quad (1)$$

this definition is for a continuous function $f(\vec{x})$ and maps it into k space. To retrieve the original function from the Fourier transformed function the Fourier transform can simply be applied to the function a second time:

$$f(\vec{x}) = \int_{-\infty}^{\infty} \frac{d\vec{k}}{(2\pi)^3} \hat{f}(\vec{k}) e^{-i\vec{x} \cdot \vec{k}}, \quad (2)$$

A Fourier transformation takes apart the original function and shows the amplitude per wavenumber needed to reconstruct the original functions from simple harmonic waves. Here in also lies the strength of the Fourier transform as it can be used to find for example the key frequencies produced by a musical instrument or find the scale of structure in a 2d image.

3.1.1 Fourier filters

One of the applications of the Fourier transform is to use it for filtering. Filtering in normal space can be quite difficult, for a top hat filter the average over a certain scale needs to be taken at every location and this is quite cumbersome. A much easier and more beautiful approach is applying the filter in Fourier space. This can be done by first taking the Fourier transform of the field and then convolving it with the filter in Fourier space. The reverse Fourier transform will then return the filtered field:

$$f(\vec{x}) = \int_{-\infty}^{\infty} \frac{d\vec{k}}{(2\pi)^3} \hat{f}(\vec{k}) \hat{W}^2(kR_f) e^{-i\vec{x} \cdot \vec{k}}, \quad (3)$$

where $f(\vec{x})$ is the filtered function, $\hat{f}(\vec{k})$ is the Fourier transform of the original function, \hat{W} is the Fourier filter and R_f is the filtering scale. The Fourier filters are the result of Fourier transforming the real space filter and these filters can also provide some simple examples of what a Fourier transform does to a function. This thesis makes use of three distinct filter functions.

1. The Fourier top hat

The Fourier top hat is one of the simplest filters available in Fourier space. The Fourier top hat filter is defined as:

$$\hat{W}(k, R_f) = \begin{cases} 1 & kR_f \leq 1 \\ 0 & kR_f > 1 \end{cases}. \quad (4)$$

A plot of the filter can be seen in Figure 2:

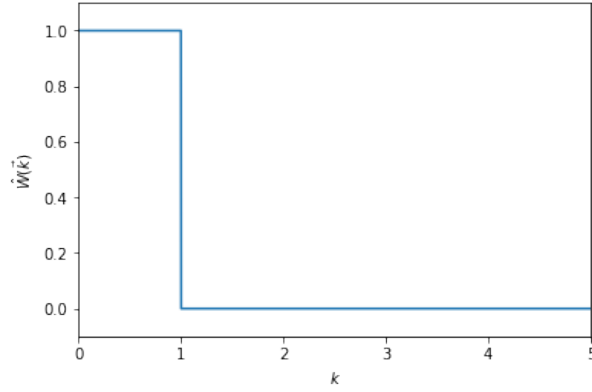


Figure 2: A Fourier top hat filter as defined by Equation 4 with $R_f = 1$.

The filter returns zero for all wavenumbers larger then the filtering scale and 1 for all wavenumbers smaller then the filtering scale. This process filters out the wavenumbers that lead to smaller scales. This filter has a sharp cutoff where the functions jumps from 0 to 1. It requires a very precise combination of harmonic waves to be able to recreate such a sharp cutoff and this filters creates sidelobes in real space. This behaviour is best explained by taking the Fourier transform of this filter.

2. Real space top hat

We now take Equation 4 as our real space function. The Fourier transform of a now real space top hat filter results in the real space top hat filter:

$$\hat{W}(k, R_f) = \frac{3 [\sin(kR_f) - (kR_f) \cos(kR_f)]}{(kR_f)^3}. \quad (5)$$

A plot of the filter can be seen in Figure 3:

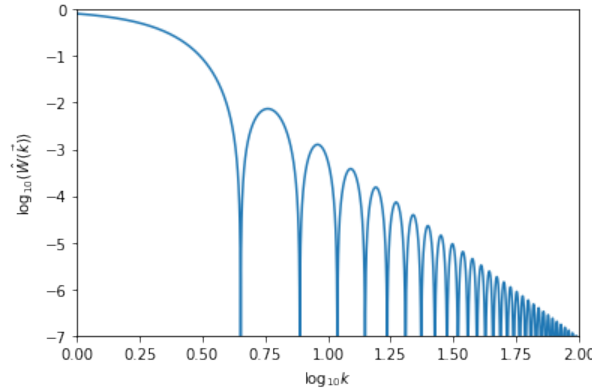


Figure 3: A Real space top hat filter in Fourier space given by Equation 5 with $R_f = 1$.

The sharp nature of the real space filter causes a ringing harmonic pattern in Fourier space. The filter can no longer be defined by a simple central distribution but instead has side structure. This side structure is often referred to as side lobes and is a structure that is always found when taking the Fourier transform of functions or images with very sharp edges.

3. The Gaussian filter

The final filter used in this thesis is the Gaussian filter. The Gaussian filter is defined as:

$$\hat{W}(k, R_f) = e^{-\frac{k^2 R_f^2}{2}} \quad (6)$$

A plot of the filter can be seen in Figure 4:

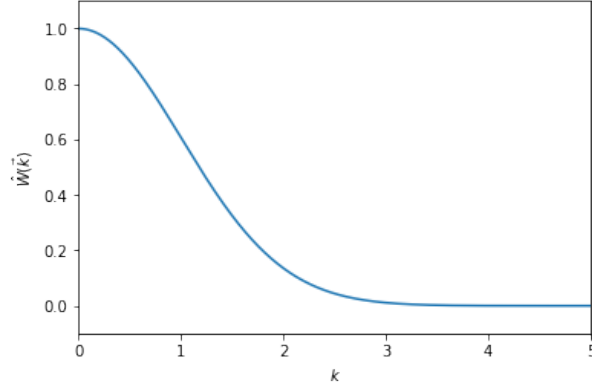


Figure 4: A Gaussian filter given by Equation 6 with $R_f = 1$.

The Gaussian filter is often used because it has the same shape in both real and Fourier space. The Fourier transform of a Gaussian is again a Gaussian. The only thing that will change is that a very broad Gaussian is transformed into a narrow Gaussian in Fourier space but the Gaussian shape is conserved. The Gaussian shape is "perfectly" smooth and has no side lobe structure in Fourier or real space.

2D Filtering

The effect of the different filtering functions on an image can be seen in Figure 5.

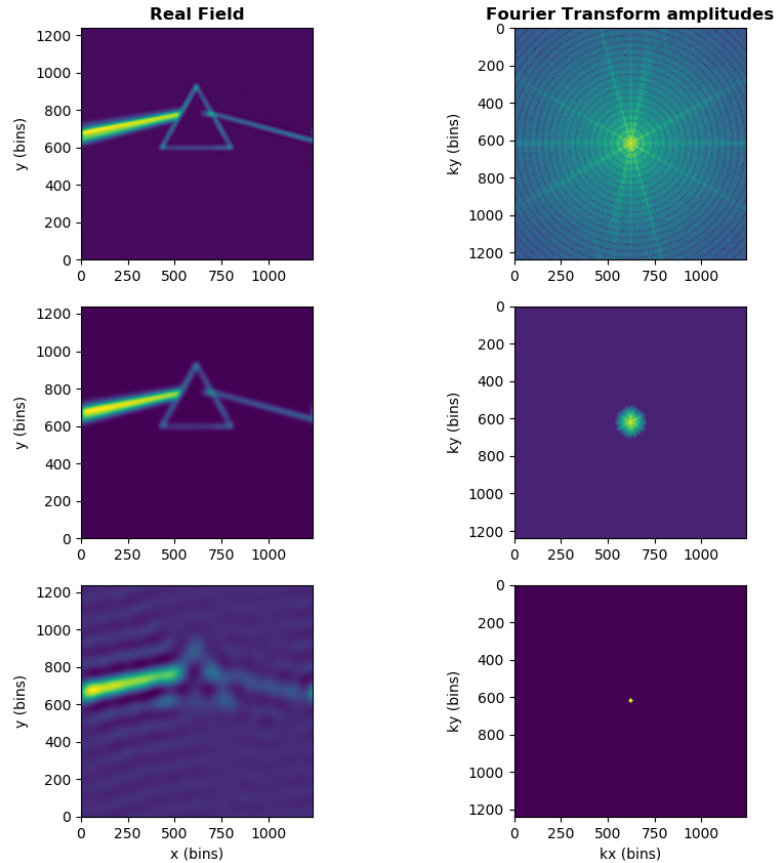


Figure 5: The real and Fourier field images using from top to bottom, A real top hat filter, A Gaussian filter and a Fourier top hat filter. Image by Hardie, Thorgerson, and Powell (1973).

Each function has a distinct effect on the original image. The side lobe structure of both the real and Fourier top hat filters is clearly visible. The real top hat has the side lobe structure visible in the Fourier image and they appear as circular waves in the Fourier image. The Fourier tophat instead induces side lobes in the filtered image. The filter has filtered out the waves with a higher wave number. The low wave number structure that remains no longer gets cancelled out causing huge side lobe effects on the sharp edges of the original image. Both the real top hat and the Gaussian filter create a very smooth filtered image. Both filters extend to infinity. In a discrete set, like for an image, the cutoff of the function can lead to aliasing effects but this is not displayed in this example.

3.1.2 Power spectra

Power spectra are usually the main statistic used when using Fourier analysis. The power spectrum of a Fourier field relates to the average amplitude of the complex numbers found in the Fourier field. Power spectra are defined by:

$$\langle \hat{f}(\vec{k}) \hat{f}^*(\vec{k}') \rangle = (2\pi)^3 P(k) \delta_D(\vec{k} - \vec{k}'), \quad (7)$$

where $\hat{f}(\vec{k})$ is the Fourier field and $P(\vec{k})$ is the power spectrum. The power spectrum shows the relative power per wavenumber and can be used to find out at which scales the data is most periodic/structured. Taking the Fourier transform of the sound coming from the A string of a guitar would result in a power spectrum with it's main peak at 110 Hz and smaller peaks at harmonics of 110 Hz (220Hz, 330Hz, ...). The power spectrum of a data set can also help find periodic signals hidden in that data set. This method can for example be used for getting an estimate of the period of pulsars. This method is described by for example Lorimer and Kramer (2004). In cosmological simulations power spectra are used to set the initial conditions in the form of Gaussian fields as the properties of a Gaussian field are completely described by its power spectrum (which is explained in more detail in section 4.2.1). One cosmological application of the power spectrum is found by taking the Fourier transform of the power spectrum:

$$\xi(\vec{r}) = \int_{-\infty}^{\infty} \frac{d\vec{k}}{(2\pi)^3} P(\vec{k}) e^{-i\vec{k} \cdot \vec{r}}. \quad (8)$$

The Fourier transform of the power spectrum is the two-point covariance function. The two-points covariance function, first introduced by Totsuji and Kihara (1969), describes the probability for two density perturbations to be at a certain distance of each other. In cosmology there are certain models for the power spectrum that are often used. The simplest form of a power spectrum is the power law:

$$P(k) \propto k^n \quad (9)$$

For negative n the most dominant structure will be at small wavenumbers which translate to large real structure, for positive n large k dominates which translates to small scale structure. One of the properties of the power law power spectrum is that it is not normalizable since it always tends to non-zero at one of it's boundaries. The power law is also scale independent so any choices for the size of the box are completely arbitrary.

The standard model power spectrum for cosmology is $P_{\Lambda CDM}$ [B. Ryden (2003)]. This spectrum was obtained from fitting data and is described by Bardeen et al. (1986) and Eisenstein (1997):

$$P_{\Lambda CDM}(k) \propto \frac{k^n}{[1 + 3.89q + (16.1q)^2 + (5.46q)^3 + (6.71q)^4]^{1/2}} \times \frac{[\ln(1 + 2.34q)]^2}{(2.34q)^2} \quad (10)$$

$$q = k/\Gamma \quad (11)$$

$$\Gamma = \Omega_{m,o} h \exp \left[-\Omega_b - \frac{\Omega_b}{\Omega_{m,o}} \right] \quad (12)$$

k has unit of $h^{-1}\text{Mpc}^{-1}$. To describe the slope of the powerspectrum we use:

$$n_k = \frac{\partial \log P(\vec{k})}{\partial \log \vec{k}}. \quad (13)$$

$P_{\Lambda\text{CDM}}$ starts of with $n_k = 1$ for small \vec{k} . This is known as the Harrison-Zeldovich spectrum and stems from analytic calculations. The result was found independently by Harrison (1970) and Y. B. Zeldovich (1972). The power spectrum has its maximum at around $0.1h^{-1}$. For large wavenumbers n_k tends to -3. This is needed for hierarchical clustering, if n_k is smaller than -3 there is no dominant small scale structure and structure will not be able to form at small scales (van de Weygaert (2016)). $P_{\Lambda\text{CDM}}$ is also normaliseable, since it goes to zero in for the limit to 0 and to ∞ . To normalise using real data we need to convolve the power spectrum with the real top hat filter found in Equation 5. To normalise we usually calculate the average dispersion (σ) of the field on a scale of $8h^{-1}\text{Mpc}$. This is based on the historical observations by Davis and Peebles (1983) where σ_8 was originally found to have a value close to one. and is the conventional quantity for normalisation. Normalisation can be done with the following integral:

$$\sigma(R_f)^2 = \int_0^\infty \frac{k^2 d\vec{k}}{2\pi^2} P_{\Lambda\text{CDM}}(\vec{k}) \hat{W}^2(\vec{k}, R_f), \quad (14)$$

where \hat{W} is the real top hat filter (Equation 5). If we then take $R_f = 8\text{Mpc}$ we can equate the integral to the observed value of σ_8 . A plot of $P_{\Lambda\text{CDM}}$ is shown in Figure 6.

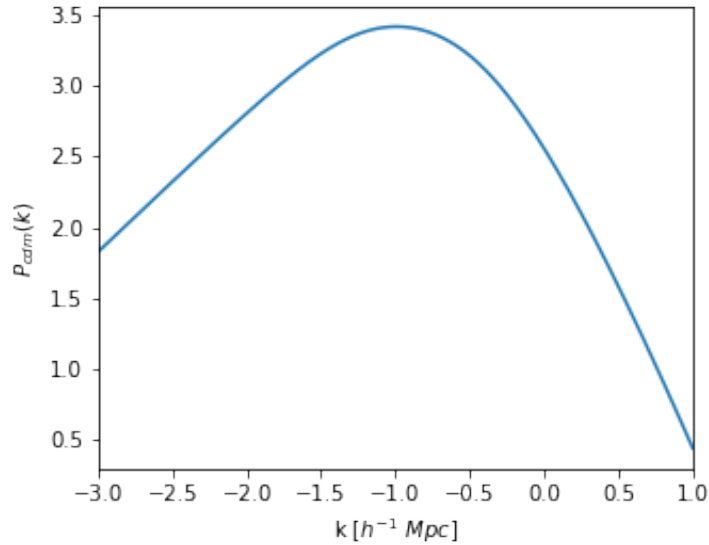


Figure 6: The P_{cdm} power spectrum in between 10^{-3} and 10 in a $\log_{10} - \log_{10}$ scale.

3.1.3 Fourier phases

The Fourier phases come from the second part of the complex field: The complex phase/angle. The Fourier phase field is defined by:

$$\phi(\vec{k}) = \arg(\hat{f}(\vec{k})), \quad (15)$$

where $\phi(\vec{k})$ is what is defined as the phase spectrum. The phases give information about the phase of the waves that construct the eventual image. In the 1 dimensional picture this translates to a shift in the location of the individual waves. This translation can be different for each wave and this governs the destructive and constructive interference of waves. For images this is very

important and the phases cause a large majority of the structure as can be seen in Figure 1. Contributions to analytic descriptions of phases for cosmology have been made by for example Coles (2005). The phases are very sensitive to the field and translations which makes them very hard to analyse.

Translations

The phases govern the phase shifts of the individual waves and make sure the individual shifts lead to the formation of structure. Changing the phase of one of the Fourier components results in a translation of that wave. Generalising this leads to the conclusion that phases can govern translations or equivalently that the phases change under translations. In a cosmological setting the phases should be able to be studied in an invariant basis and the initial coordinate system should not matter. To show how the phases respond to a translation of the original function we define (\vec{u}) which is the function $f(\vec{x})$ translated by the vector \vec{u} :

$$g(\vec{u}) = f(\vec{x} + \vec{u}). \quad (16)$$

And we define the Fourier transform of $f(\vec{x})$ as:

$$\hat{f}_0(\vec{k}) = \int_{-\infty}^{\infty} d\vec{x} f(\vec{x}) e^{-i(\vec{x} \cdot \vec{k})}. \quad (17)$$

We can then inspect the Fourier transform of a translated function by Fourier transforming the function $g(\vec{u})$:

$$\hat{g}(\vec{u}) = \int_{-\infty}^{\infty} d\vec{x} g(\vec{u}) e^{-i(\vec{x} \cdot \vec{k})} = \int_{-\infty}^{\infty} d\vec{x} f(\vec{x} + \vec{u}) e^{-i(\vec{x} \cdot \vec{k})}. \quad (18)$$

By performing a change of basis we can rewrite the integral to:

$$\hat{g}(\vec{u}) = \int_{-\infty}^{\infty} d\vec{l} f(\vec{l}) e^{-i[(\vec{l} - \vec{u}) \cdot \vec{k}]}, \quad (19)$$

where $\vec{l} = \vec{x} + \vec{u}$. This can then be rewritten to:

$$\hat{g}(\vec{u}) = e^{i(\vec{u} \cdot \vec{k})} \int_{-\infty}^{\infty} d\vec{l} f(\vec{l}) e^{-i(\vec{l} \cdot \vec{k})} = \hat{f}_0(\vec{k}) e^{i(\vec{u} \cdot \vec{k})}, \quad (20)$$

where \hat{f}_0 is the Fourier transform of the untranslated function. The Fourier amplitudes are complete unchanged by a translation but the phases get an additional term that scales linearly with \vec{k} and the phase spectrum becomes:

$$\phi'(k) = \phi(k) + \vec{k} \cdot \vec{u}, \quad (21)$$

where $\phi'(k)$ is the phase spectrum of $\hat{g}(\vec{u})$ and $\phi(k)$ is the phase spectrum of $\hat{f}_0(\vec{k})$. It is important to keep in mind that there is no physical difference between the 2 fields and only the coordinates have changed. Not only is the phase spectrum different but the difference is dependent on both \vec{k} and the translation \vec{u} which makes it very difficult to study a phase spectrum in systems where the coordinates should not matter. In Section 7 we look at a solution to this problem.

3.1.4 Visualisation of phases

The Fourier phases are periodic as there is no difference between a harmonic waves with phases related by $\phi = \phi_0 + k2\pi$ for any integer k . This causes the Fourier phases to be cyclical, linearly increasing the phases results in a rotation along the unit circle. Using a simple linear intensity scale is not a valid way of representing phases as the jump from one branch edge to the next is unnatural

and a nonphysical eye guide. Coles and Chiang (2000) discuss a nice way of visualising phases in a visually friendly manner by mapping angles to a colourmap that loops on itself. This colourmap also gives a clear representation of phase rotations as this causes the colours to cycle through the map. This method of displaying the phases will be used throughout this thesis.



Figure 7: The visualisation of phases used in this thesis. The phase can be read of as the angle of an imaginary line starting horizontally at the red colour (Red = 0 , Yellow/Green (top) = 0.5π , Light blue (left) = π and purple (bottom) = 1.5π). Image by Coles and Chiang (2000)

3.2 Gaussian random fields

The second formalism that is needed for the understanding of the Fourier phases is the formalism of Gaussian random fields. Gaussian fields are essential to the field of cosmology and their cosmological background will be discussed in more depth in Section 4.2.1.

3.2.1 Definition

Gaussian fields are defined by the N point distribution function:

$$\mathcal{P}_N = \frac{\exp \left[-\frac{1}{2} \sum_{i,j=1}^M f_i (M_{ij}^{-1}) f_j \right]}{[(2\pi)^N (\det M)]^{1/2}}, \quad (22)$$

this is a multivariate Gaussian where f_i is the function describing the field and M_{ij} is the covariance matrix:

$$M_{ij} = \langle f(\vec{x}_i) f(\vec{x}_j) \rangle = \xi(\vec{x}_i - \vec{x}_j) = \xi(|\vec{x}_i - \vec{x}_j|). \quad (23)$$

The fact that the N point distribution function is a multivariate Gaussian is the reason these random fields are named Gaussian random fields. The covariance matrix is directly related to the 2 point correlation function and thus also to the power spectrum. The covariance matrix is the only adjustable variable in the distribution and thus Gaussian fields are completely described by their 2 point correlation function/power spectrum. Gaussian fields also translate nicely into Fourier space. To define what makes a field Gaussian in Fourier space we separate the Fourier nodes into its individual components:

$$\hat{f}(\vec{k}) = \hat{f}_r(\vec{k}) + i \hat{f}_i(\vec{k}) = |\hat{f}(\vec{k})| e^{i\phi_k}, \quad (24)$$

where $\hat{f}(\vec{k})$ represents one of the Fourier modes, $\hat{f}_r(\vec{k})$ is the real component of the Fourier mode, $\hat{f}_i(\vec{k})$ is the imaginary component of the Fourier mode, $|\hat{f}(\vec{k})|$ is the amplitude of the Fourier mode and ϕ_k is the phase of the Fourier mode. For a field to be Gaussian these components need to satisfy three conditions [van de Weygaert 2016]:

- The phases are uniformly distributed over the 0 to 2π range: $\phi(\vec{k}) \in \mathcal{U}[0, 2\pi]$
- The amplitudes are Rayleigh distributed:

$$P(|\hat{f}(\vec{k})|) dP(|\hat{f}(\vec{k})|) = \exp \left(-\frac{|\hat{f}(\vec{k})|^2}{2P(\vec{k})} \right) \frac{|\hat{f}(\vec{k})| d|\hat{f}(\vec{k})|}{2P(\vec{k})}, \quad (25)$$

where $P(|\hat{f}(\vec{k})|)$ is the probability density of a certain value of $|\hat{f}(\vec{k})|$ and $P(\vec{k})$ is the power spectrum of the Gaussian random field.

- $\hat{f}_r(\vec{k})$ and $\hat{f}_i(\vec{k})$ are independently drawn from the Gaussian distribution and are thus mutually independent.

A couple of conclusions can be made from these conditions. The phases are uniformly distributed and the value for each individual node is drawn randomly from this distribution, this causes the phases to hold no information. The conditions actually allow for only one free parameter, the power spectrum $P(\vec{k})$ found in the Rayleigh distribution or equivalently the 2 point correlation function ξ in the . This implies that the characteristics of Gaussian fields are completely specified by its power spectrum.

3.2.2 Gaussian fields and the power spectrum

Gaussian fields are completely described by their power spectra and this makes them a nice way of visualising power spectra. An example of four Gaussian fields with simple power law power spectra ($P(\vec{k}) \propto k^n$) can be seen in Figure 8. The difference between initial power spectra is easily visible. The fields with negative n have dominant large scale structure and the fields with positive n have dominant small scale structure.

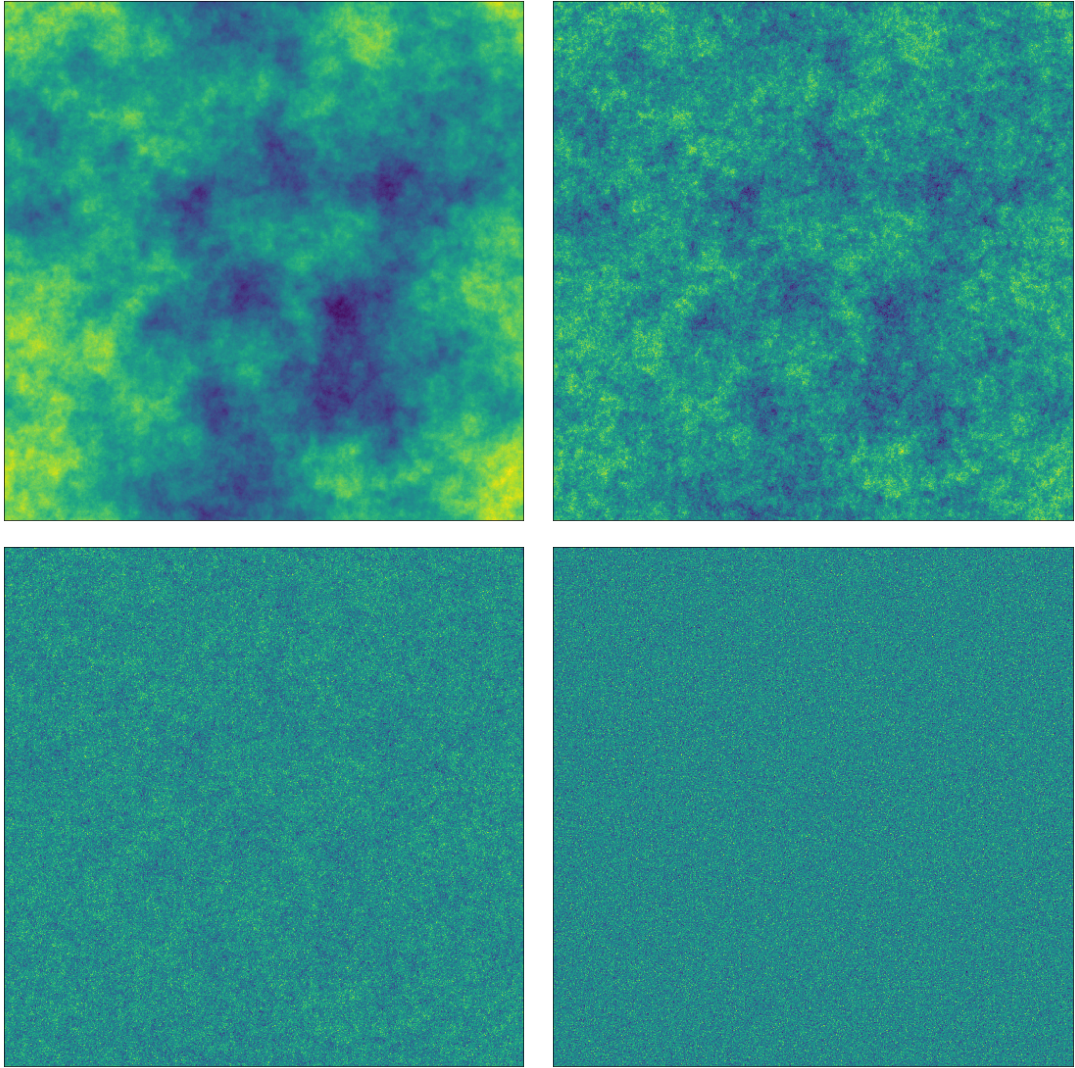


Figure 8: Four Gaussian fields with simple power law power spectra (Equation 9). Top left has $n = -3$, top right $n = -2$, bottom left has $n = -1$ and bottom right has $n = 1$.

4 Cosmological Background

In this section the basics of cosmic structure formation will be explained. This is needed to appreciate the context of Fourier phases. Cosmic time and density will be characterised and the most important concepts of cosmic structure formation are explained.

4.1 Characterising Densities and times

Cosmological simulations and theory span a huge range of scales and cosmic time and cosmic densities have to be characterised in order to be able to compare results.

4.1.1 Densities

In cosmology densities are usually normalised with the mean in to be able to work in convenient dimensionless quantities. Densities are characterised by:

$$\rho = \bar{\rho}(1 + \delta), \quad (26)$$

where ρ is the density at a certain point, $\bar{\rho}$ is the average density and δ is the density parameter. The density parameter has the underdensities below zero and the overdensities above zero and has the mean density at zero. δ is often used to analyse a density field instead of ρ as it is dimensionless. The cosmological densities can range from -1, completely empty, to values $\gg 1$. This often asks for a logarithmic scaling. In this thesis logarithmic density field intensities are characterised by $\log(1 + \delta)$.

One of the parameters that describes cosmic density fields is σ_8 . σ_8 specifies the amplitudes of the density perturbations, quantifies the spread in the densities and is a tracer of structure formation. The simplest definition of σ_8 is:

$$\sigma_8^2 = \langle (\rho - \bar{\rho})^2 \rangle_{8\text{Mpc}}. \quad (27)$$

σ_8 is the variance in the density field at a scale of $8 \text{ h}^{-1} \text{ Mpc}$. σ_8 can also be calculated from Fourier space:

$$\sigma_8^2 = \int_0^\infty \frac{k^2 dk}{2\pi^2} P(k) \hat{W}^2(k, 8\text{Mpc}), \quad (28)$$

here $\hat{W}(k, 8\text{Mpc})$ represents a real space top hat filter (Equation 5) on a scale of $8 \text{ h}^{-1} \text{ Mpc}$ and $P(k)$ is the Fourier power spectrum of the density field. The 8 Mpc scale used for σ_8 is a convention stemming from Davis and Peebles (1983). They found the value of $\sigma(R_f)$ to be close to 1 at the scale of 8 Mpc and it has since become a convention. Recent observations by for example Planck Collaboration, Ade, Aghanim, Arnaud, et al. (2016) have found from the CMB that the value is closer to 0.8.

4.1.2 Time evolution

When analysing cosmological simulations density fields are usually given at different slices of cosmic time usually represented by either redshift or expansion factor. The expansion factor is found as the scale factor in the Robertson–Walker metric and describes how much the universe has expanded. The expansion factor starts of $\ll 1$ and increases with cosmic time and is defined as 1 at current time. The expansion of the universe is observed as a redshift. Light that is emitted travels through an expanding space and this causes the light waves to expand as well. The expansion of the light causes the wavelength to be stretched creating a redshift. Redshift is directly related to the expansion of the universe and thus also the expansion factor:

$$a = \frac{1}{1 + z}, \quad (29)$$

where a is the expansion factor and z is the redshift. This allows the two quantities to be used interchangeably. Different models can provide different levels of structure formation and evolve differently as a function of redshift/expansion factor. In this case σ_8 can be used to compare structure formation and function as a parameter for time evolution as $\sigma_8(t) = D(t)\sigma_8(t_0)$ and thus directly relates to the growth factor $D(t)$.

4.2 Cosmic structure formation

This section will cover the main phases of structure formation mainly focusing on the initial condition and the linear phase as they are best understood and are analytically solvable.

4.2.1 Initial conditions and Gaussian random fields

Gravitational cosmic structure formation is largely determined by the initial conditions of the universe. One way of studying the initial conditions of our universe is by looking at the cosmic microwave background (CMB). The CMB was first discovered by Penzias and Wilson (1965) and its discovery was backed by theoretical work by Dicke et al. (1965). The CMB is a remnant of the very early universe. After the epoch of recombination photons were no longer trapped by the free electrons and were allowed to stream free throughout the universe. This snapshot of the early universe is a treasure trove and provides us with a better understanding of the initial conditions of cosmic structure formation. Observations of the CMB, by for example Planck Collaboration, Ade, Aghanim, Armitage-Caplan, et al. (2014) and Komatsu et al. (2003), show that the CMB has a very close resemblance to a Gaussian field. The assumption that the initial conditions are Gaussian are further backed by theories of inflation, with great contributions from Guth (1981) and Linde (1982) and the central limit theorem, discussed by Bardeen et al. (1986), which states that all distributions with a well defined mean and variance become Gaussian for large N . The theory of Gaussian fields and their applications for cosmology has been extensively researched with big contributions from Bardeen et al. (1986) and Adler (1981).

Gaussian phases

Gaussian fields are multiscale and therefore best described in Fourier space. A Gaussian field is completely described by its Fourier power spectrum as the Fourier phases are randomly and uniformly distributed over the zero to 2π range ($\phi(\vec{k}) \in U[0, 2\pi]$). The random nature of the initial Fourier phases puts constraints on studying how the phases evolve. The phases evolve from a random distribution into a strictly non-random distribution (Figure 1) and this is the behaviour that will be studied in this thesis.

4.2.2 The phase of linear growth

After the initial conditions are set the universe enters the first phase of structure formation. In this phase the small initial perturbations in the density field start to evolve. To understand how perturbations evolve during this phase a new set of quantities need to be defined.

The density perturbation:

$$\delta(\vec{x}, t) = \frac{\rho(\vec{x}, t) - \bar{\rho}}{\bar{\rho}}, \quad (30)$$

where $\rho(\vec{x}, t)$ is the density at a certain point in space and time and $\bar{\rho}$ is the mean density.

The co-moving coordinate:

$$\vec{x}(t) = \frac{\vec{r}(t)}{a(t)}, \quad (31)$$

where $\vec{r}(t)$ is the regular space coordinate and $a(t)$ is the expansion factor. The equations needed to solve for the evolution of these quantities are the fluid equations, detailed by Peebles (1980).

The fact that most structure formation happens during the matter-dominated phase allows for the use of the simplified fluid equations. The continuity equation, which governs conservation of mass:

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot (\rho \vec{u}) = 0, \quad (32)$$

the Euler equation, which is the equation of motion and governs conservation of momentum:

$$\frac{\partial \vec{u}}{\partial t} + (\vec{u} \cdot \vec{\nabla}) \vec{u} = -\vec{\nabla} \Phi, \quad (33)$$

and the Poisson equation, which couples the Euler equation with gravity:

$$\vec{\nabla}^2 \Phi = 4\pi G \rho, \quad (34)$$

in these expressions ρ is the density, \vec{u} is the velocity and G is the gravitational constant. These equations can then be transformed to co-moving coordinates and can then be solved for the quantity δ . This results in the evolution equation[van de Weygaert (2016)]:

$$\frac{dD}{dt} + 2\frac{\dot{a}}{a} \frac{dD}{dt} = \frac{3}{2} \Omega_0 H_0^2 \frac{1}{a^3} D, \quad (35)$$

where $D(t)$ is known as the universal growth factor and is related to the density perturbation by:

$$\delta(\vec{x}, t) \propto D(t) \delta(\vec{x}, t_0) \text{ for } \delta \ll 1. \quad (36)$$

The density perturbation evolves from the initial conditions linearly with $D(t)$. Solutions of the evolution equation are dependent on H_0 , Ω_0 and on the time evolution of the expansion factor $a(t)$ and are thus linked to the type of universe used.

Solutions for $D(t)$

One of the solutions of this problem is for an Einstein-de Sitter universe with $\Omega_0 = 1$ and a constant H_0 . For such an universe:

$$D(t) \propto a(t) \propto t^{2/3}, \quad (37)$$

the densities evolve linearly with expansion factor in the linear regime. We can also find $D(t)$ for an empty universe. We can get to an empty universe by starting with an Einstein-de Sitter universe ($\Omega_0 = 1$) and then asymptotically go to $\Omega_0 \ll 1$ as $a(t) \gg 1$, this results in:

$$\begin{aligned} D(t) &\propto a(t) \propto t^{2/3} \text{ for } a(t) \ll 1, \\ D(t) &\propto \text{cst. for } a(t) \gg 1. \end{aligned}$$

$D(t)$ starts off proportional to $a(t)$, but as $a(t)$ approaches values much bigger than one there is a halt on structure formation and $D(t)$ becomes constant.

The power spectrum

This linear phase is also seen in the evolution of the power spectrum:

$$P(\vec{k}) \propto D^2 P(\vec{k}, t_0), \quad (38)$$

here $P(\vec{k})$ is the power spectrum and $P(\vec{k}, t_0)$ is the initial power spectrum. In the linear phase the power spectrum rises at the same rate for all scales i.e. it rises independently of scale. This behaviour for a Λ CDM power spectrum can be seen in Figure 9:

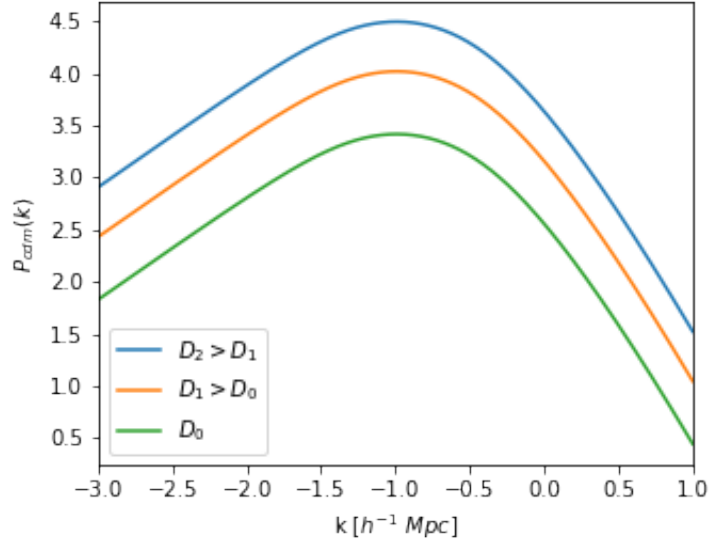


Figure 9: An extreme representation of the linear increase of a P_{cdm} power spectrum with increasing $D(t)$. The bottom line is the initial power spectrum.

The power spectrum increases uniformly over the entire range and is independent on \vec{k} . This approximation is only valid when $\delta \ll 1$ and when this approximation breaks the universe enters the non-linear phase.

4.2.3 The phase of non-linear growth

When the density perturbations become too large they can no longer be approximated as increasing linearly with $D(t)$. At this point an analytical approach becomes impossible and the system must be solved numerically using for example N-body simulations. Hierarchical structure formation starts at small scales. This causes the power spectrum to rise more for large \vec{k} . Peacock and Dodds (1996) and Smith et al. (2003) have provided analytic expressions that describe this behaviour. These expressions are usually simple fitting functions represented by broken power laws and allow for simple comparison between simulations. Similar power law fits were used by B. S. Ryden and Gramann (1991) when dealing with phases which could imply that these analytic models could be of use for the modelling of phases.

4.2.4 The final state of the universe

In a Λ -dominated or an open matter-dominated universe structure formation eventually comes to a halt. An ever expanding universe does not just become larger but it also becomes emptier and the contribution of matter in the universe eventually becomes negligible. Using Equation 35 we can then find an analytic solution for $D(t)$. In an empty universe $D(t)$ becomes constant for large expansion factors. If the universe became empty due to expansion it is safe to assume a is large and all structure formation will come to a halt.

5 Numerical Methods

In this Section the numerical methods used in this thesis will be described with the more basic functions described in Section A.1 of the appendix. All calculations were made using `python3` and all modules used are available for `python3`.

5.1 Numerical Fourier transforms

For this thesis all numerical Fourier transforms were carried out using the various functions from the `scipy` module `fftpack`. This module has very simple to use functions for Fourier transforms in one, two or more dimensions and uses very efficient algorithms. For two dimensional the function `fft2` was used and for higher dimensional Fourier transform `fftn` was used.

5.1.1 Fourier phases and amplitudes

After taking a Fourier transform the Fourier amplitudes and Fourier phases need to be separated. In order to extract the amplitudes the `numpy` function `numpy.abs` was used which returns the amplitude of each complex number in the Fourier grid:

$$r = |Pe^{i\phi}| = P \quad (39)$$

Where r is the Fourier amplitude, P is the power at a certain grid point and ϕ is the complex angle. For the Fourier phases the `numpy` function `numpy.angle` was used. The function returns the argument of the complex value:

$$\theta = \arg(Pe^{i\phi}) = \phi \quad (40)$$

Where θ is the Fourier phase, P is the power at a certain grid point and ϕ is the complex angle. In order to make sure the branch of the argument is known the modulus of θ is taken with 2π in order to map the phases in the range from 0 to 2π . A function was also made that can shuffle the Fourier phases. It works by first drawing a random number for a uniform distribution in between 0 and 2π and then adding this number to the Fourier phase:

$$\hat{f}(k_x, k_y)_{new} = \hat{f}(k_x, k_y)_{old} e^{i\phi_{random}} \quad (41)$$

Here $\hat{f}(k_x, \phi_x, k_y, \phi_y)_{old}$ is the original value of the Fourier gridpoint, $\hat{f}(k_x, \phi_x, k_y, \phi_y)_{new}$ is the new value and ϕ_{random} is the random number. This needs to be done for each point on the Fourier grid separately. This process causes all phase information to be lost. This method does break the symmetry in Fourier space that ensures that the resulting image is real, namely:

$$\hat{f}(\vec{k}) = \hat{f}^*(-\vec{k}). \quad (42)$$

This problem is circumvented by just taking the real part of the resulting Fourier transform. One could instead choose to iterate over half the grid and fill in the conjugate at the same time. This process takes more computational time and we choose to just use the real part of the resulting field.

5.1.2 Numerical power spectrum

After performing the Fourier transform it is also useful to extract the power spectrum. This is done by first creating the grid of wavenumbers that underlies the Fourier grid. This can be done for both the x wavenumber and the y wavenumber. The main interest is the dependence of the power spectrum on the length of the wavenumber so the x and y grid are combined with:

$$|\vec{k}| = \sqrt{k_x^2 + k_y^2} \quad (43)$$

Combining this grid with the amplitudes of the Fourier transform gives the discrete values to have $P(|\vec{k}|)$. This now needs to be made into one dimensional data which can be done by flattening both grids, this keeps the x and y values linked while converting it to one dimension. These array will now not be sorted probably and we can sort them both by sorting both the wavenumber and the power values according to the wavenumbers. This then gives the sorted x values along with the accompanying y values. These can now simply be re-binned linearly or logarithmically to lower the noise. This is then a discrete representation of the power spectrum as a function of the length of the wavenumber.

5.1.3 Numerical Phase spectrum

The phase spectrum, defined by Equation 15, can be extracted in a comparable way to the power spectrum. Again a length of wavenumber grid is constructed along side the Fourier phase grid. This can then again be flattened and sorted. The difference comes from the fact that the domain of the phases is known, namely from 0 to 2π . This can then again be plotted in the same way as the power spectrum. Instead of displaying it as a function of wavenumber it is nicer to instead display it as a histogram, this also removes the need for the wavenumber grid reducing computational time.

5.2 Generation of Gaussian fields

Gaussian fields are easiest to generate in Fourier space, all that is needed is the power spectrum which will completely determine how the final field will look like. We start by creating a grid that represents the wavenumber at that point in the grid. This grid is needed since we need to be able to calculate the value of the power spectrum at each point. To ensure the real field will actually be real we need:

$$\hat{f}(\vec{x}) = \hat{f}^*(-\vec{x}) \quad (44)$$

Where $*$ indicates the conjugate of the function. Therefore only half of the grid needs to be iterated over and each point corresponds to filling in two points in the complete grid. The individual gridpoints are created by drawing the R_1 and R_2 values of a complex number independently from a random Gaussian distribution and are then convolved with the value of the chosen powerspectrum at that point:

$$\hat{f}(\vec{k}) = (R_1 + iR_2)\sqrt{P(|\vec{k}|)} \quad (45)$$

Where $\hat{f}(\vec{k})$ is the complex value at a certain grid point, R_1 and R_2 are random Gaussian values and $P(|\vec{k}|)$ is the powerspectrum wanted for the Gaussian field. When the grid is completely filled up with values all that is left is Fourier transforming it back into real space. Due to numerical errors only the real part of the transform is taken, even after ensuring the symmetry of the Fourier grid there will still be negligible imaginary components that are discarded when taking only the real part. An example of a Gaussian can be seen in Figure 10.

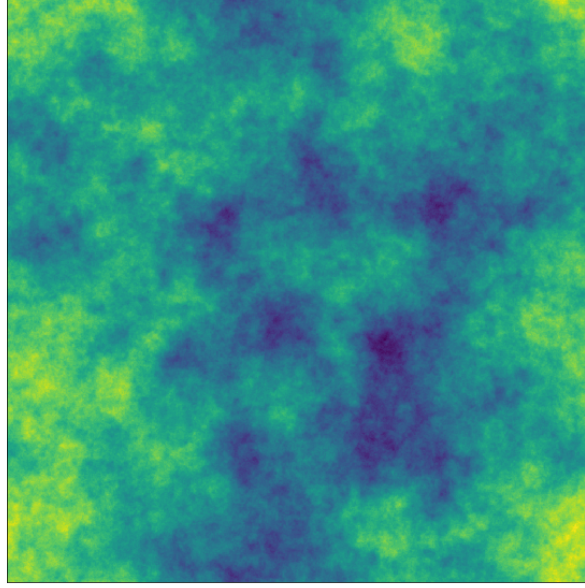


Figure 10: Example of a Gaussian field with a $n=-2$ power law power spectrum ($P(k) \propto k^n$)

5.3 Particle locations to density fields

N-body simulations do not result in simple density fields and an algorithm is needed to convert particle locations into a density field. For this thesis a simple Nearest-Grid-Cell method is used. There are more sophisticated methods, the simplest alternative being a Cloud-In-Cell method which interpolates the particles in the nearest cells, but for the purposes of this thesis this was not required. The reasoning for this is explained in more detail in Section 6. To calculate the density fields we start by constructing a density grid on which the particles can be counted. This grid can then be set to have a zero value at every gridpoint. The next step is to iterate over every particle and for each particle calculate the gridpoint it belongs to. This can be done with:

$$x_{grid} = \frac{x_i}{L} * N \quad (46)$$

Where x_{grid} is the location on the new grid, x_i is the particle position, L is the length of the original box and N is the number of bins in a sidelength in the new density grid. To get the index of the gridpoint x_{grid} needs to be rounded to an integer and if it is equal to N it needs to be set to 0 (a set with length N goes from 0 to $N-1$). For each particle one is added to the gridpoint it corresponds to resulting in a 3d histogram. To get cosmological densities all that is left to do is to apply Equation 26 and the resulting grid will have the correct densities.

5.4 The Phase gradient

In order to fully study the Fourier phases we define two quantities, the phase gradient and σ_{∇} . The phase gradient is the normal gradient applied to the phase spectrum:

$$\vec{\nabla}_{\phi}(\hat{f}(\vec{k})) = \vec{\nabla}_k(\arg(\hat{f}(\vec{k}))) = \vec{\nabla}\phi(\vec{k}), \quad (47)$$

where $\vec{\nabla}_{\phi}$ is the phase gradient, $\hat{f}(\vec{k})$ is a Fourier field, $\phi(\vec{k})$ is the phase spectrum of the Fourier field and $\vec{\nabla}$ is the \vec{k} gradient. Numerically this is achieved by first applying Equation 40 and then either Equation 73 or 74.

5.5 σ_{∇}

Below is the mathematical definition of σ_{∇} for a discrete set of data. σ_{∇} describes the spread in the average value of the phase gradient (Eq. 47) in each slice. Consider a 3 dimensional box with discrete z coordinate i:

$$\phi_i(n, m) = \arg \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx dy f(x, y, i) e^{i(nx+my)} \right), \quad (48)$$

where $f(x, y, i)$ is a slice at z-coordinate i in the original field. We can then calculate the phase gradient of this slice:

$$\nabla_i(n, m) = \vec{\nabla}\phi_i(n, m) \quad (49)$$

We can then calculate the mean of the gradient at each slice of the box need to be calculated:

$$\bar{\nabla}_i = \frac{1}{N_b M_b} \sum_{n=1}^{N_b} \sum_{m=1}^{M_b} \nabla_i(n, m), \quad (50)$$

here $\nabla_{n,m,i}$ is the value of the 2d phase gradient of slice at z coordinate i at the point (n, m), N_b and M_b are the sidelights of the 2d slice at $z = i$ and N is the number of 2d slices in the z direction. $\bar{\nabla}_i$ is then the average value of the gradient in that slice. σ_{∇} is the standard deviation for this subset of values. It can then be calculated by:

$$\sigma_{\nabla}^2 = \frac{1}{N} \sum_{i=1}^N (\bar{\nabla}_i - \bar{\bar{\nabla}})^2, \quad (51)$$

where $\bar{\bar{\nabla}}$ is the average of the values $\bar{\nabla}_i$:

$$\bar{\bar{\nabla}} = \frac{1}{N} \sum_{i=1}^N \bar{\nabla}_i. \quad (52)$$

Slices can be taken in any direction and not just the z direction. The z direction is what is used in this thesis. A full description of why this quantity is used is given in Section 8.2.

6 Models and Simulations

In this section the various models and Simulations used will be detailed. 4 types of models were used: simple images, Gaussian fields, Voronoi models and an N-Body simulation.

6.1 Images

As a base example we used the image of the prism found on The dark side of the moon album by Pink Floyd. The design of the album cover was created by Hardie, Thorgerson, and Powell (1973). This image has sharp edges which ensure there is structure in both the Fourier amplitudes and the phases. The distinct structure makes it perfect for use in tests of invariance as the differences between translated images are easier to spot for a distinct image. The image used can be seen in Figure 11:



Figure 11: The image used as a guide for phases and power spectra. Image by Hardie, Thorgerson, and Powell (1973)

6.2 Gaussian Fields

Gaussian fields were used as the initial conditions of both the power spectrum and the phases as the universe initially started of in a state close to Gaussian. The power spectrum is a free parameter and can be set to whatever is needed but it was mainly set to either a power law or $P_{\Lambda CDM}$. The Gaussian fields were used as a way of visualising the power spectra and to look at the random initial distribution of the phases. An example of a Gaussian field an its phases can be seen in Figure 12

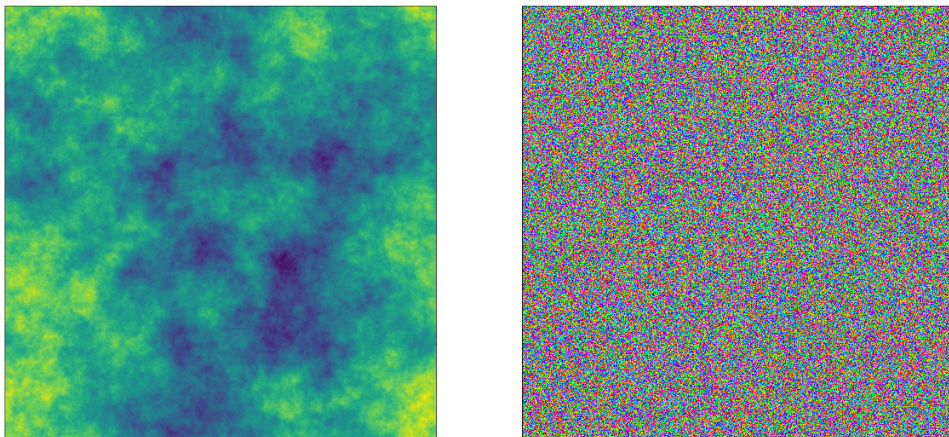


Figure 12: Example of two Gaussian fields with power law $P(\vec{k}) \propto k^n$. Left image has $n = -3$ and the right image has $n = -2$

6.3 Voronoi models

The cosmic web is the nonlinear structure that emerges as the universe enters the early non-linear phase of gravitational cosmic structure formation. The cosmic density field starts forming the first types of structure consisting of flattened walls, elongated filaments, high density cluster nodes and large underdense voids. This morphology has been a topic of many studies with contributions from Zel'dovich (1970), Shandarin and Ya. B. Zeldovich (1989), Bond, Kofman, and Pogosyan (1996), van de Weygaert and Bond (2008), Aragón-Calvo et al. (2010) and Cautun et al. (2014). The morphology of the cosmic web lends itself to a Voronoi tessellation based description and Voronoi clustering models have been described by van de Weygaert and Icke (1989), Icke and van de Weygaert (1991), van de Weygaert (2003) and van de Weygaert (2003). The Voronoi grid was first defined by Voronoi (1908). Voronoi tessellations are based on a Voronoi cell structure. A Voronoi cell is defined as the region of space in which all points are closest to the Voronoi cell center of that cell. In 2D this results in the structure seen in Figure 13. The white regions are the Voronoi cells defined by the blue cell centers. In 3D this becomes a cell like structure. The cell-like Voronoi tessellation structure can then be identified with the structures found in the cosmic web. The planes/faces of the tessellation become the walls, the tessellation edges with filaments, the nodes with clusters and the cell interiors with voids.

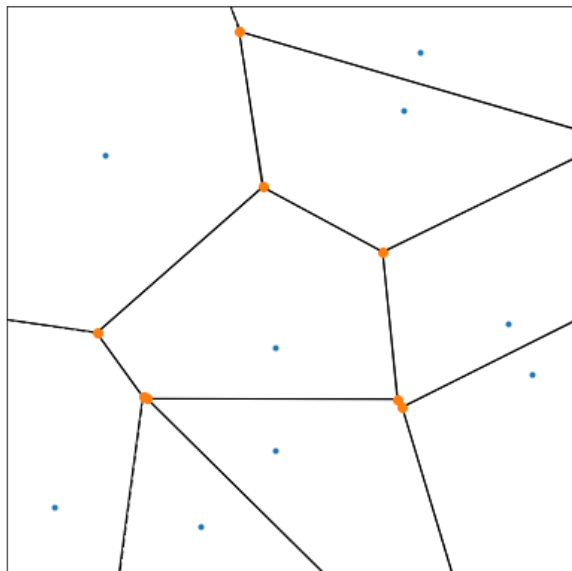


Figure 13: A Voronoi grid, the blue dots represent the cell centers, the black lines show the Voronoi grid and the orange dots represent the nodes at which multiple grid lines intersect.

6.3.1 Specific Voronoi models used

The Voronoi models can simulate cosmic structure formation in a simple way and were described by van de Weygaert and Icke (1989) and Icke and van de Weygaert (1991). The models were used to test what kind of structure the quantity σ_{∇} responds to (this quantity will be defined in Section 8.2). The only prior knowledge about σ_{∇} was that it responds to extreme overdensities in the density field. The Voronoi models have a large number of adjustable values that allow for tests on a wide variety of structures. The Voronoi models are also largely deterministic and the final conditions could be set allowing for a very complete view of what happens in the box. This in turn make understanding what our new quantity responds to a much simpler process. All models used one million particles in a 200 Mpc box. And were converted to a 128^3 bins density field. This meant that there was a larger number of cells that what not filled. For the purpose of this thesis this was not a problem as it just increased contrast between under and overdense

regions while retaining a higher resolution in the overdense regions. The code provided several models to use and compare. The full results of the models can be found in Section A.3.

Voronoi kinematic model

The Voronoi kinematic model is the most accurate model that was available and its functionality is the backbone of all the other models. First the particle locations are generated and a Voronoi tessellation with a set number of cells is generated. The particles then start moving according to the path seen in Figure 14.

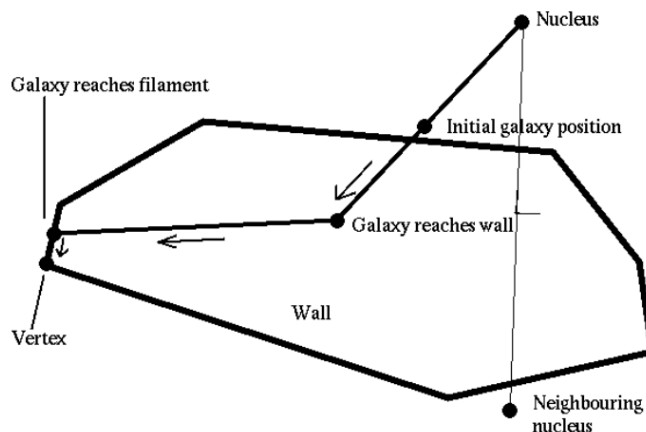


Figure 14: The path a particle follows in the Voronoi models. Image from van de Weygaert (2002), Courtesy: Jacco Dankers.

The particles start moving towards the planes defined by the tessellation on the line drawn from the Voronoi center through the particles to the nearest wall. When the particles reach the wall they lose their velocity perpendicular to the plane and start moving towards the filaments. When the particles reach the filaments they again lose the velocity perpendicular to the filaments and the velocity that the particles have left causes them to move to the clusters where they eventually start to pile up. This model gives results that are closest to a regular N-body simulation. The main differences between this model and a regular N-Body simulations are caused by the low amount of particles. The overdensities forming at the nodes are way better behaved. Due to the lower number of particles the most dense regions are still significantly less dense than the most overdense regions in a regular N-Body simulations. The voids are also way emptier than for a regular N-Body simulation. This is a good thing as this causes the extreme regions to be less dominated by overdensities allowing different effects to leave their mark on our quantity. Results of this model can be seen in Figure 15.

Voronoi elements model

The second model makes direct use of the Voronoi morphology and puts the particles in the 4 distinct structures defined by the Voronoi tessellation. The path described in the Kinematic model is used to determine the final position of the particles but the particles now stop when they reach the assigned structure. The first phase is just the random distribution of particles, The second phase has all the particles in the walls, the thirds has all the particles in filaments and the final phase puts all the particles in the nodes. This model allows for a quick look at very structured density fields and provides a nice baseline for the other models. The results of the individual phases can be seen in Figure 16.

Increasing Number of Voronoi cells model

In this model we use the fact that structure is defined by Voronoi tessellations to set up a unique type of experiment. We start with a baseline evolved universe according to the Voronoi kinematic model, in this universe most of the structure is found in the walls and filaments which should provide a good baseline for an evolved universe. We then start increasing the number of Voronoi cells while keeping the level of evolution in the universe the same. This results in an increase in the number of nodes and very overdense regions but due to the low ammount of particles the structures also start becoming less dense and overdense regions become less peaked as structure becomes to smaller and smaller, eventually becoming to small to be resolved. The resulting effect is almost like zooming out as structure decreases in size. The results can be seen in Figure 17.

Bubble explosion model

Finally we made use of a more extreme model, a bubble explosion model. The movement of the particles is now governed by spherical shockwaves that emanate from the Voronoi cell centers. The shockwaves push out all particles within a certain radius of the cell center and move them along the path described for the Voronoi kinematic model. This process results in huge pileups of particles along the edges of the shockwaves and leaves behind truly empty areas. This process can be seen in Figure 18. Particles only move with the shockwaves and particles not yet hit remain in their random state. This create a very large contrast between the shockwaves, the empty cells and the random parts of the field. These contrasts are what make it an interesting model to use even though it is not very physical.

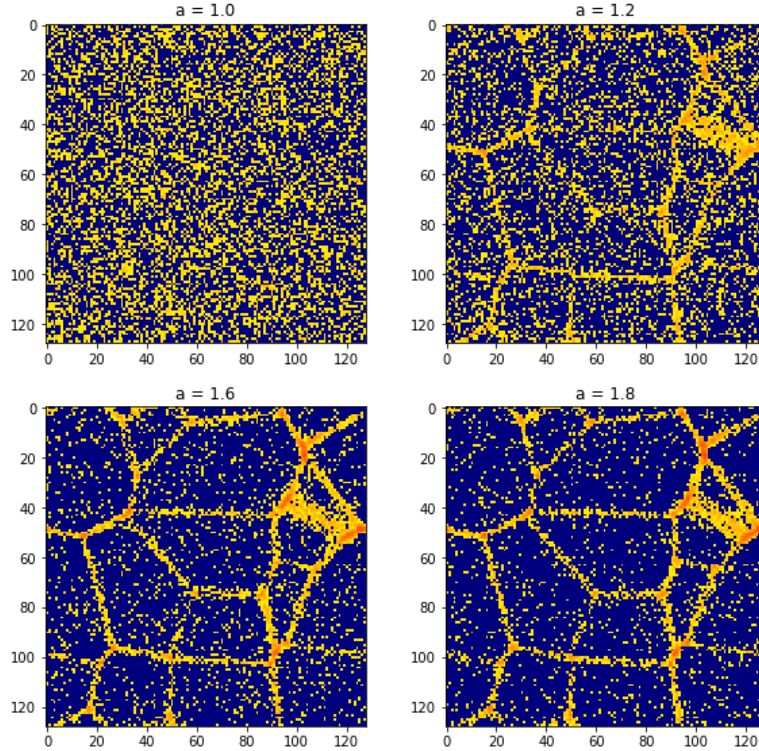


Figure 15: 4 Slices of the Voronoi kinematic model

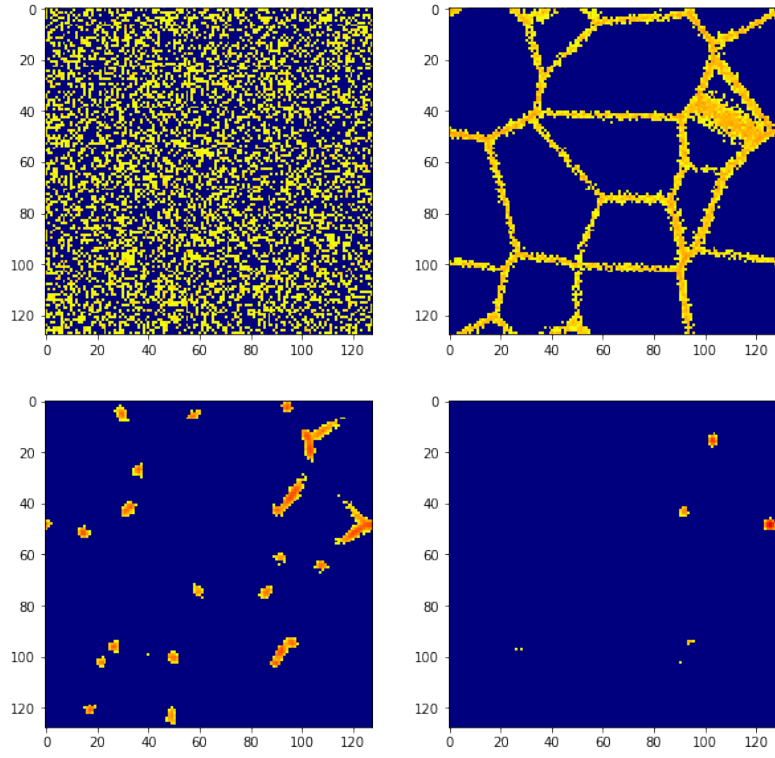


Figure 16: The 4 phases of the Voronoi kinematic model

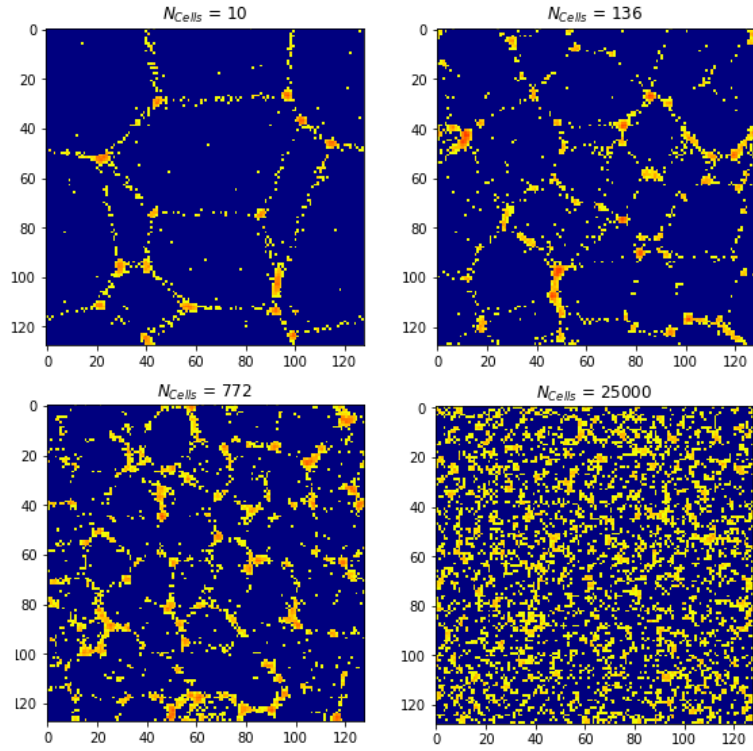


Figure 17: 4 Slices of the increasing cells model

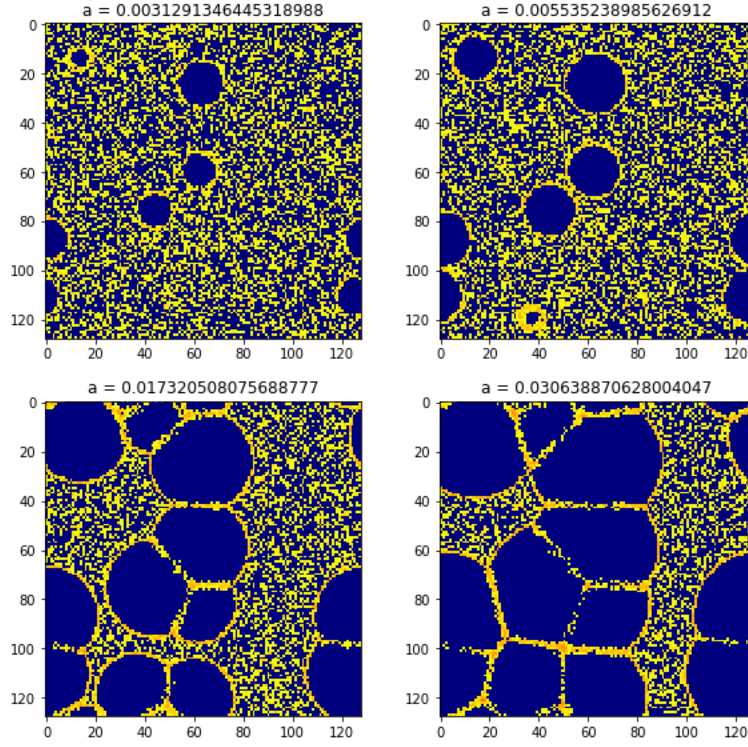


Figure 18: 4 Slices of the bubble explosion model

6.3.2 N-Body Simulations

The N-body simulation used was created by Bos (2016). The N-body simulation uses a Λ CDM powerspectrum as initial conditions with cosmological parameters $\Omega_M = 0.272$, $\Omega = 0.728$ and $h = 0.704$ in a $200 \times 200 \times 200$ Mpc³ box. From the simulation eight snapshots were taken at different cosmic times. The density fields were taken at redshift 0, 0.1, 0.25, 0.51, 1, 2, 3 and 3.8. The redshifts translate to the expansion factors: 1, 0.91, 0.8, 0.66, 0.5, 0.33, 0.25 and 0.21. The density fields were given on a grid with 512^3 pixels. The N-Body simulation is used as a model as it can provide results very close to using regular data. It is also used as a basis for trying to understand how the Fourier phases form structure as a function of cosmic time. In order to be able to compare structure formation σ_8 was also calculated at each redshift. Figure 19 shows the result of taking the mean of 1000 measurements of σ_8 . One slice was taken at 6 of the cosmic times and a part of the density fields along with the Fourier phases of the complete slice are shown in Figure 20 and 21.

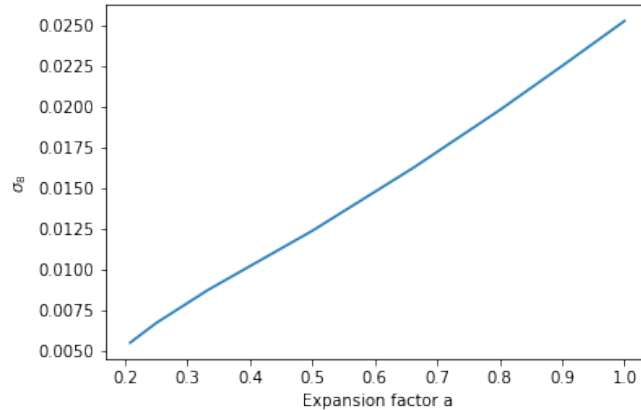


Figure 19: σ_8 as a function of expansion factor a , densities are defined as $\delta + 1 = \frac{\rho}{\bar{\rho}}$

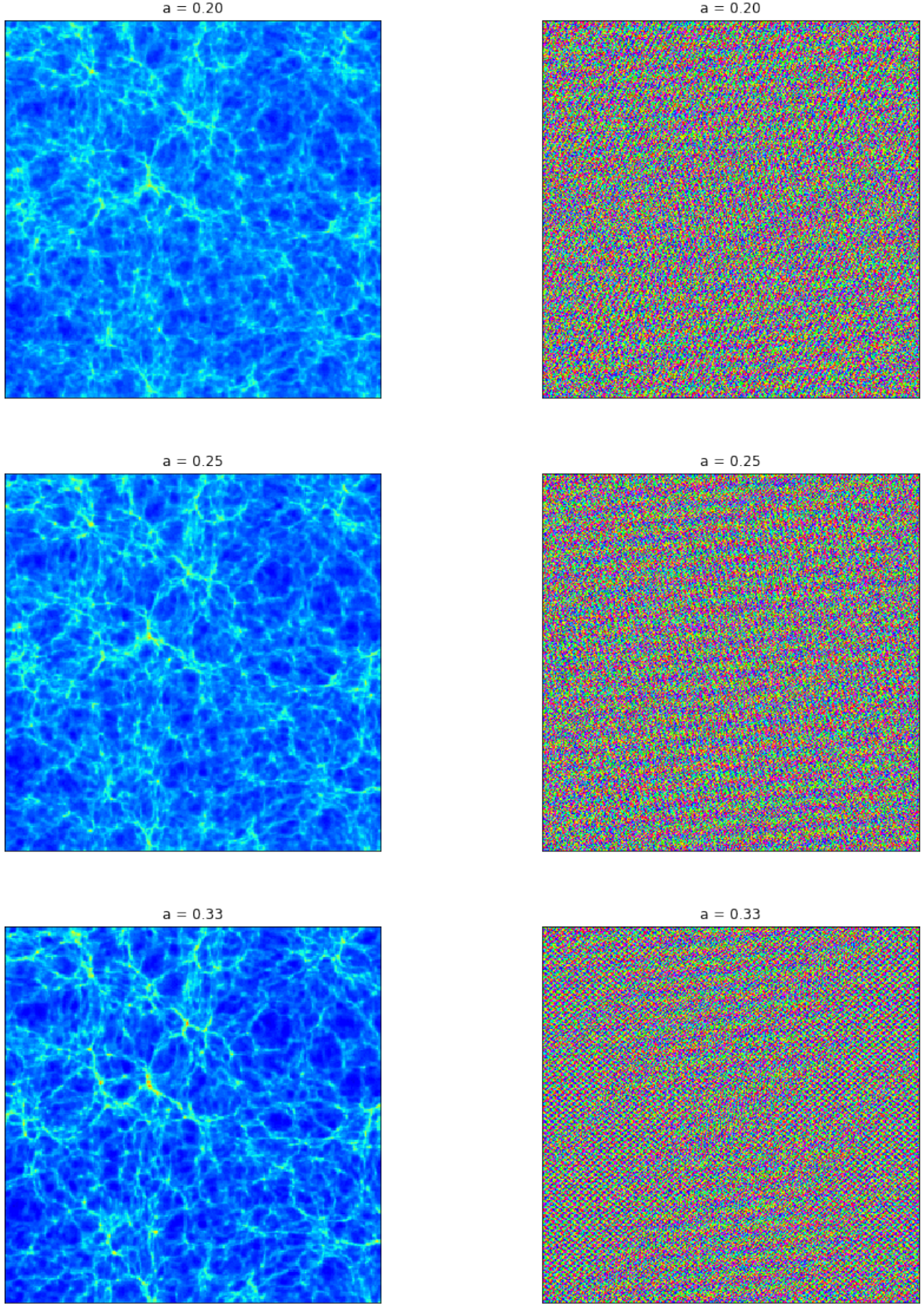


Figure 20: On the right 3 slices of the density fields defined by $\log(1 + \delta)$ showing a 100×100 Mpc part of the total slice, the left plots have the Fourier phases of the complete 200×200 Mpc slice. The expansion factor of each slice is found in the title of each slice.

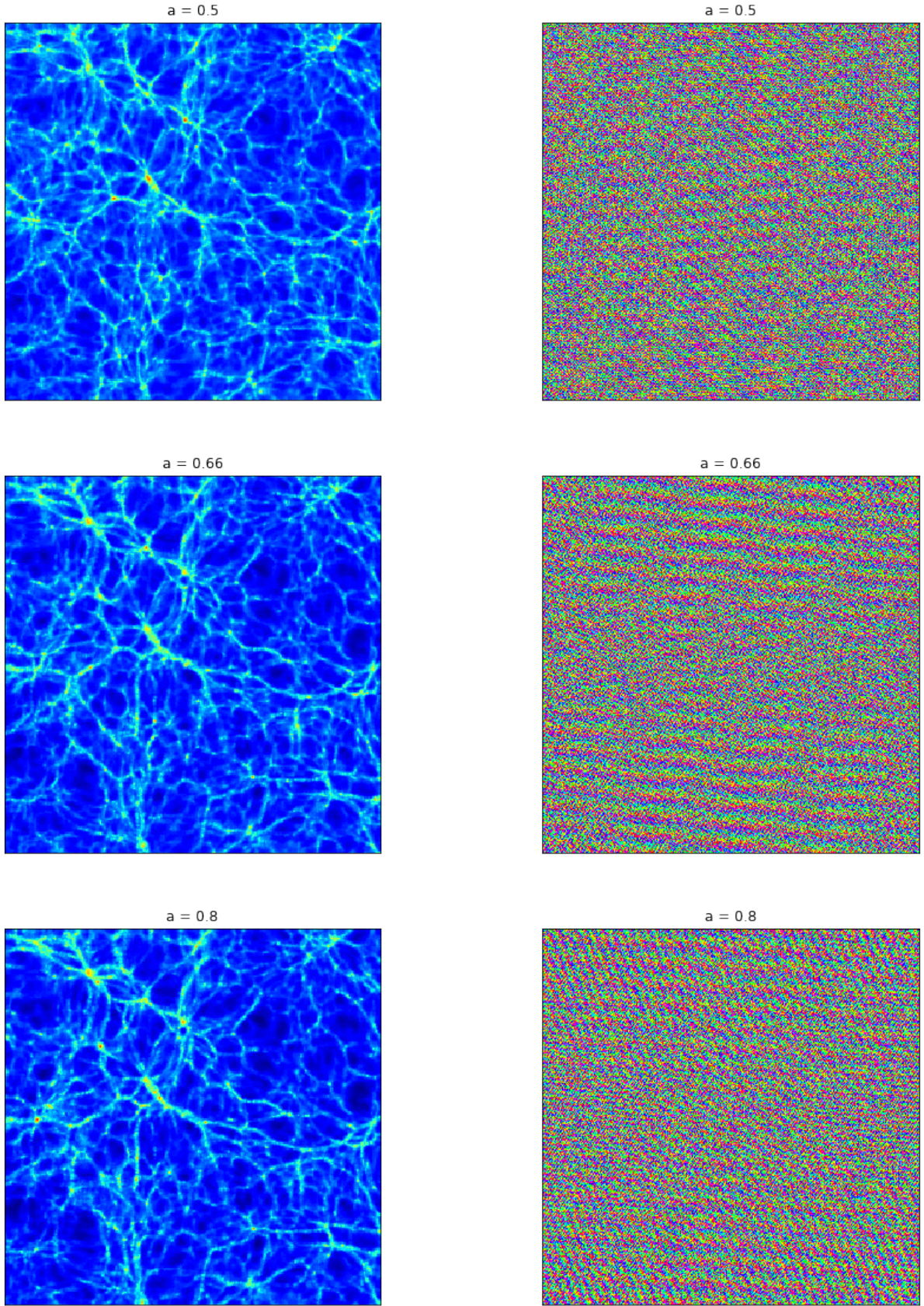


Figure 21: On the right 3 slices of the density fields defined by $\log(1 + \delta)$ showing a 100×100 Mpc part of the total slice, the left plots have the Fourier phases of the complete 200×200 Mpc slice. The expansion factor of each slice is found in the title of each slice.

7 Phase Invariants

When translating an image or grid in real space (assuming periodic boundary conditions) the Fourier amplitudes are unchanged, the translations are translated into the Fourier phases only. We previously showed this and the function $g(\vec{u})$ which is the Fourier transform of $f(\vec{x})$ translated by \vec{u} gives:

$$g(\vec{u}) = e^{i\vec{u} \cdot \vec{k}} \hat{f}_0(\vec{k}), \quad (53)$$

with $\hat{f}_0(\vec{k})$ being the Fourier transform of the untranslated function $f(\vec{x})$. In order to be able to properly analyse the phases a invariant needs to be used to make sure the framework gives equal results even after translations. To be able to easily see the effect of translations on the phases we define the phase gradient:

$$\vec{\nabla}_\phi \hat{f}(\vec{k}) = \vec{\nabla} \left[\arg \left(|\hat{f}(\vec{k})| e^{i\phi(\vec{k})} \right) \right] = \vec{\nabla} \phi(\vec{k}), \quad (54)$$

the phase gradient takes the derivatives of \vec{k} of the phase spectrum $\phi(\vec{k})$. When looking at the phase gradients before and after a translation it can be seen that the gradients are linearly dependent on how much the original image is translated:

$$\vec{\nabla}_\phi g(\vec{u}) = \vec{\nabla}(\vec{u} \cdot \vec{k}) + \vec{\nabla} \phi(\vec{k}) = \vec{u} + \vec{\nabla} \phi(\vec{k}), \quad (55)$$

the translation adds a constant term to the phase gradient. This effect be seen in Figure 22, here the phases of the image detailed in Section 6.1 is used to clearly show the effect of a translation. The phases are not randomly distributed and provide a clear mean value. Translating a phase distribution with a clear mean gradient results in an oscillation around π as this will be the value in the two boundaries conditions, the mean in the middle of the distribution or the mean close to 0 or 2π . This behaviour is seen in Figure 22. The phase gradient perpendicular to the translation is constant. The relation between translations and the gradient shows that it is important to find an invariant under translations. To describe Fourier phases in a robust framework an invariant is essential especially due to the fact that in cosmology the universe should be homogeneous and the choice of coordinates should not make a difference.

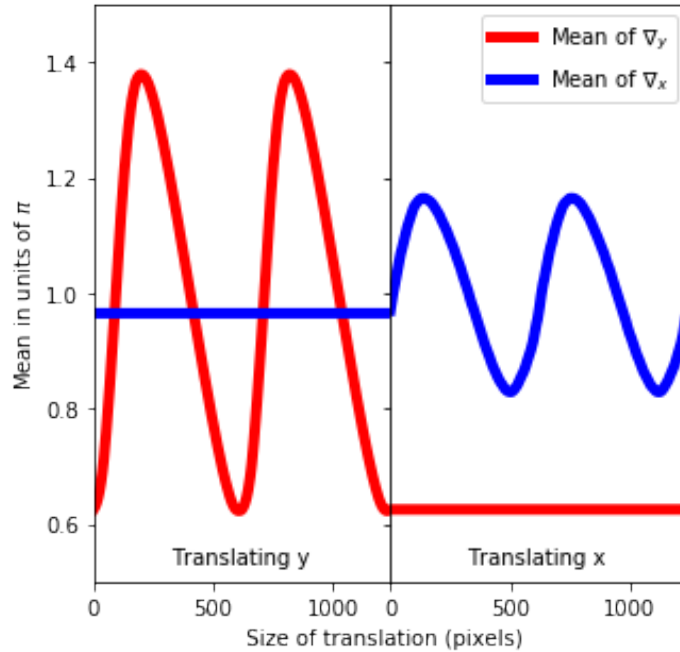


Figure 22: The relation between the mean of the gradient and the translations along x and y. The first half shows translations in the y direction and the second half shows translations in the x direction

7.1 The Taylor invariant

In the field of image recognition there has been a lot of research into creating invariants for images to for example add invisible watermarks as shown by Ruanaidh and Pun (1998). One particularly interesting invariant is the Taylor invariant. The Taylor invariant is a translation invariant form of the image, it does this by cancelling out the linear term added to the gradient of the Fourier phases. The Taylor invariant can be seen as a unique solution for the function $g(\vec{u})$ where:

$$g_{Taylor} = g(\vec{u}_0) = e^{i\vec{u}_0 \cdot \vec{k}} \hat{f}_0(\vec{k}) \quad (56)$$

where \vec{u}_0 is the unique translation which produces the Taylor invariant from the original function $f(\vec{k})$. One way of finding the Taylor invariant is proven by Lin and Brandt (1993). They prove that the Taylor invariant is the Fourier transform of a function with its center of mass located at the origin. This solution is of particular interest to us as this would provide a very easy way to calculate \vec{u}_0 as this would simply be the inverse of the coordinates of the center of mass.

7.1.1 The center of mass

Finding the center of mass of an image can be done by averaging the rows and columns and then calculating the center of mass using Equation 57. Using the coordinates with the mean of the related row/column results in the 1d coordinate of the center of mass.

$$x_{com} = \frac{1}{M_{tot}} \sum_{i=1}^n x_i m_i, \quad (57)$$

where x_{com} is the coordinate of the center of mass, M_{tot} is the total mass and x_i is the position of the particle with mass m_i and index i . The problem with using this technique is that it does not take periodic boundary conditions into account and the obtained center of mass is not invariant under translations in such a system. A simple thought experiment to understand this degeneracy is to imagine a circle first centered at the origin. The center of mass of the image will be in the center of the circle. If we now translate the center of the circle into a corner we get a situation in which each corner of the image now has a quarter circle. If we now again calculate the center of mass again it will again be at the origin since the mass is now distributed symmetrically in the corners. This center of mass is thus degenerate and has multiple solutions in a periodic system. From this we can conclude that a different method is needed to find the center of mass that does use the fact that the system has periodic boundary conditions.

7.1.2 A center of mass in a periodic system

A solution to this problem was proposed by Bai and Breen (2008). They talk about a method to calculate the center of mass in a system with periodic boundary conditions. It works by mapping the points to a circle. Then the center of mass is then calculated on the circle. The center of mass coordinates on the circle can then be mapped to the coordinate of the center of mass. Again the means of the rows/columns can be used turning it into two one dimensional problems. First the original coordinates are mapped to angles on the circle:

$$\theta_i = \frac{x_i}{x_{max}} 2\pi, \quad (58)$$

where x_i is the original coordinate, x_{max} is the sidelength of the periodic box and θ_i is the angular coordinate to which the points are mapped to. These angles can then be used to map the original points to points on a unit circle using:

$$k_i = \cos \theta_i \frac{x_{max}}{2\pi}, \quad (59)$$

and

$$l_i = \sin \theta_i \frac{x_{max}}{2\pi}, \quad (60)$$

here k_i and l_i are the x and y coordinate of the circle the points are mapped to. Then we can apply Equation 57 to both of these new coordinates to find k_{com} and l_{com} which will be the center of mass coordinates on the circle. At this point all that is left to do is reverse transform this back into the original coordinate:

$$\theta_{com} = \arctan2(l_{com}, k_{com}) + \pi, \quad (61)$$

and finally

$$x_{com} = x_{max} \frac{\theta_{com}}{2\pi}, \quad (62)$$

x_{com} is then the center of mass coordinate in the original coordinate system. Doing this for both x and y grants a unique solution to the center of mass in a system with periodic boundary conditions.

7.2 The Taylor invariant of a static image

The Taylor invariant looks to be a good way to achieve an invariant mathematically but it is always good to inspect how it behaves in a more realistic setting. To do this a simple image was chosen with some clearly visible structure in the phases. This image is translated to a new position and the periodic center of mass is then calculated of the translated image. The center of mass is then translated to be at the origin. Then the Fourier transform is taken of the image. The results of two iterations of this can be seen in Figure 23:

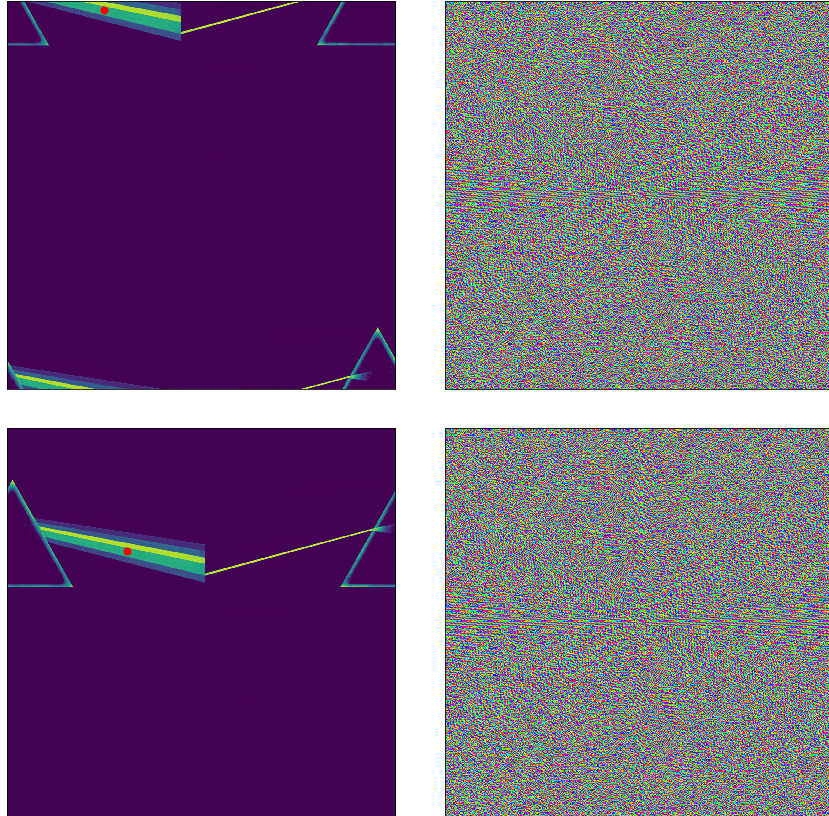


Figure 23: The left images show the randomly translated image along with a red dot at the calculated center of mass. The right images show the Fourier phases ($\phi(k_1, k_2)$) of the Taylor invariant using the circular colourmap described in Section 3.1.4. Image by Hardie, Thorgerson, and Powell (1973).

From the figure the two iterations of the Taylor invariant appear to be very similar and the center of mass is located at the same position for both randomly translated images. On first sight both phase images appear to be very similar. In order to see if the phases were truly equal we looked at the residuals between the two images defined as:

$$\text{Res} = |\phi_1(k_1, k_2) - \phi_2(k_1, k_2)| \quad (63)$$

where ϕ_1 is the phase spectrum of the first image and ϕ_2 is the phase spectrum of the second image. The result of this can be seen in Figure 24:

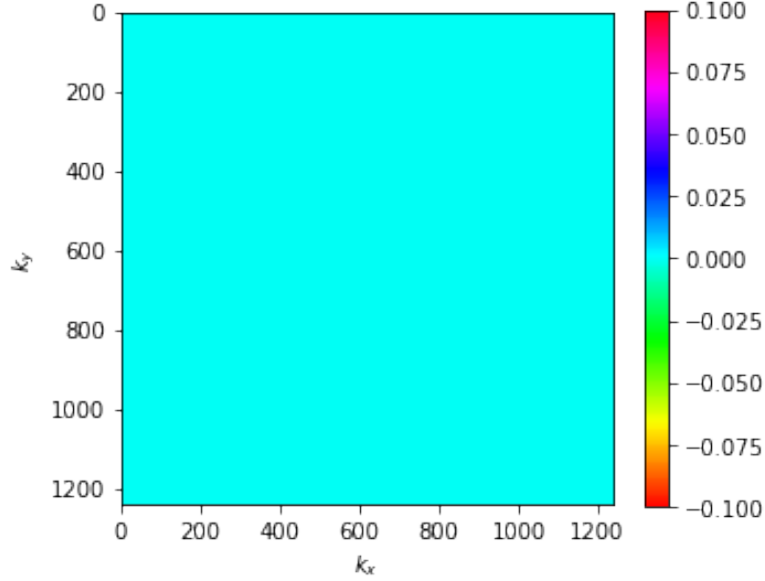


Figure 24: The residuals of the Taylor invariants in radians

The residuals are nicely close or equal to zero as expected and thus the Taylor invariant works. One thing to note is that the numerical Taylor invariant (when the translations are in units of bins) is not perfect. Round of errors can cause the center of mass to be at different locations around the actual center of mass (translated by one bin). In this case the residuals are not zero and there clearly is a linear part of the gradient still active (An example of this can be found in Figure 41 in Section A.2). Great care needs to be taken to be completely sure that the used center of mass is correct and that rounding the center of mass coordinates is done equally. Ideally it would be best to do the translations by interpolation as this will remove the possibility of rounding errors. The Taylor invariant appears to be a good framework for working with translation invariant Fourier phases, great care is needed to use it properly especially when there is a high probability of numerical errors.

8 Data analysis

In this section the Fourier phases of the different models will be analysed. First we look into the phase structure, what causes the phases structure and we provide a new quantity that is able to quantify the amount of phases structure. This quantity is then computed for a number of models in order to find what the quantity and thus the phase structure responds to in real space. The phases are also analysed using the method discussed by B. S. Ryden and Gramann (1991). Here we use the phase shifts to look at evolution of the phases. Both methods of analysing the Fourier phases are also tested for invariance under translations.

8.1 Fourier phases of the cosmic web

The simplest way to analyse the Fourier phases is to simply take the Fourier transform and then observe the phases. All Fourier transform will be taking of the field with densities defined by $\delta + 1$, this allows for easy transformations to a logarithmic density scale (like in Figure 25). An example of the Fourier phases of one slice can be seen in Figures 20, 21 and 25. The phases are very-non random and display clear structure. The phases display regularities and patterns that are close to wavelike. When looking and comparing multiple slices there is a variety in the amount of phase structure, sometimes the structure is very regular and wavelike and sometimes there does not appear to be any real dominant structure and the phases look more like noise. The slice in Figure 25 shows one of the more extreme cases.

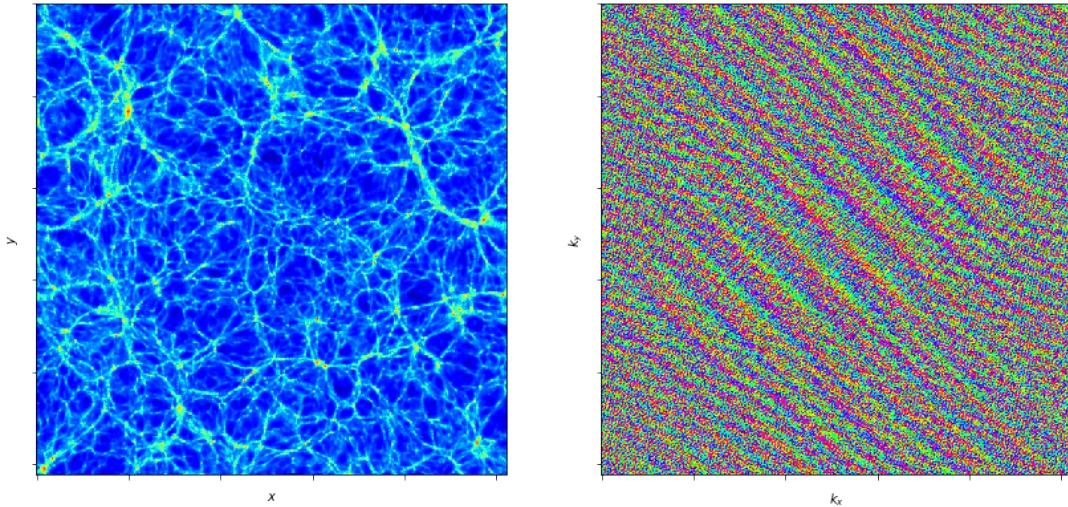


Figure 25: Left a slice of a 200 Mpc³ box at expansion factor 1 with logarithmic density scaling for visual purposes and right the phases of the Fourier transform ($\phi(k_1, k_2)$) of the slice in the left image with linear densities.

8.1.1 Origin of phase structure

To inspect the origin of the structure we threshold the original field shown in Figure 25 from below and above with a simple filter. from above this becomes:

$$\delta(\vec{x}) = \begin{cases} \delta(\vec{x}) & \delta(\vec{x}) \leq \delta_{threshold} \\ \delta_{min} & \delta(\vec{x}) > \delta_{threshold} \end{cases}, \quad (64)$$

and from below:

$$\delta(\vec{x}) = \begin{cases} \delta(\vec{x}) & \delta(\vec{x}) \geq \delta_{threshold} \\ \delta_{min} & \delta(\vec{x}) < \delta_{threshold} \end{cases}, \quad (65)$$

in both these expressions δ_{min} is the lowest value of the density field. This was not set at zero as this would cause problems with logarithmic plots. The results can be seen in Figure 26 and

27. Removing just a small part of the overdensities causes most original phase structure to be lost while a very huge part of the lower densities can be removed while still retaining most of the original phase structure. The overdensities dominate the phase structure. This provides an explanation for the wavelike patterns seen in the phase structure. If we look at the Fourier transform of a Dirac delta function ($f_\delta(\vec{x}) = \delta_D(\vec{x} - \vec{x}_0)$):

$$\hat{f}_\delta(\vec{k}) = \int_{-\infty}^{\infty} \delta_D(\vec{x} - \vec{x}_0) e^{i\vec{x} \cdot \vec{k}} d\vec{x} = e^{i\vec{x}_0 \cdot \vec{k}}, \quad (66)$$

we see that delta spikes cause waves in the Fourier phases. The overdensities are so dominant that they act like delta functions. The combination of a couple of these spikes create the wavelike patterns. The Fourier phase structure observed is a superposition of these delta spike induced waves.

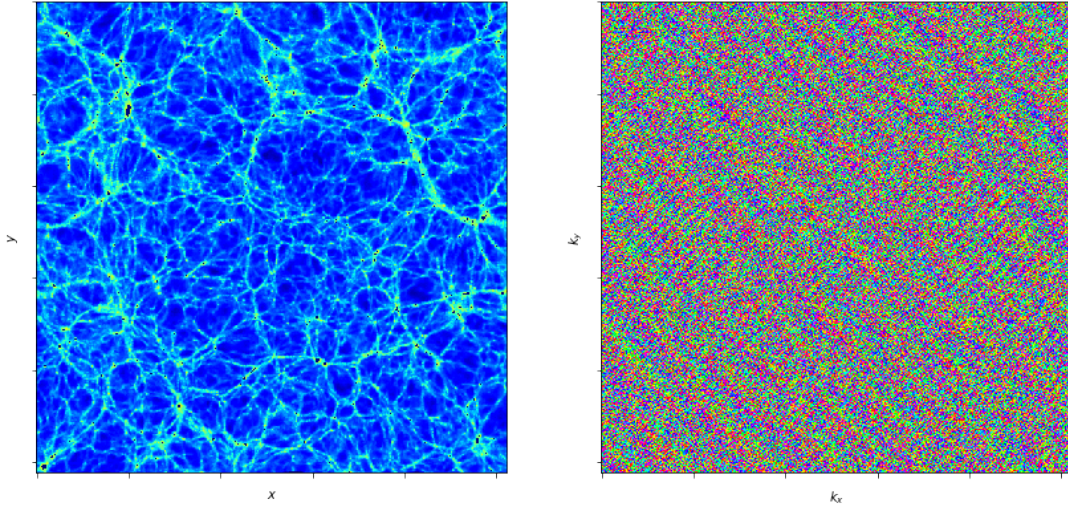


Figure 26: The left image is a slice of a 200 Mpc³ box at expansion factor 1 with logarithmic density scaling for visual purposes filtered by 64 until most of the structure in the phase diagram is removed. The dark blue spots show the filtered overdensities. The right image shows the phases of the linear scaled real field with the same filter

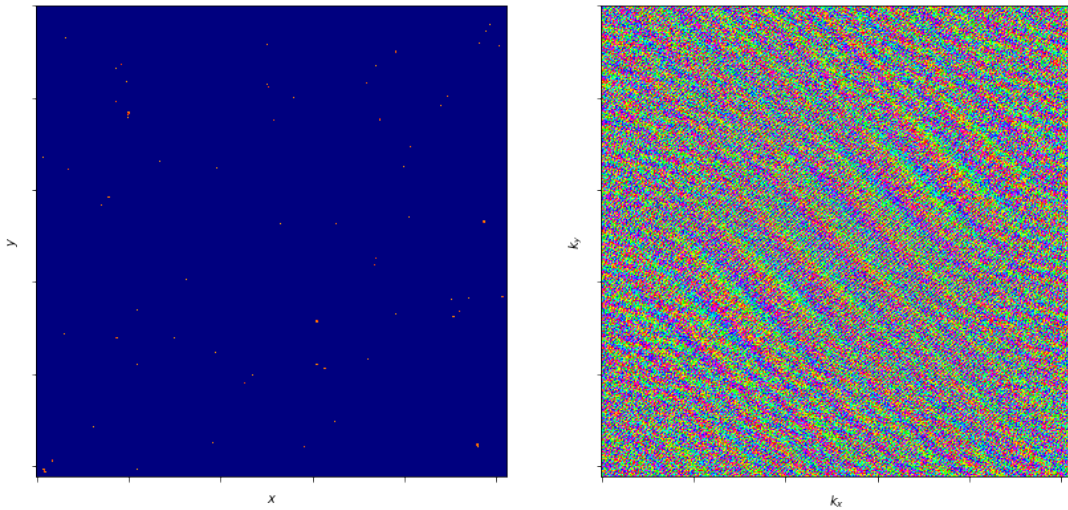


Figure 27: The left image is a slice of a 200 Mpc³ box at expansion factor 1 with logarithmic density scaling for visual purposes filtered by 65 until the point where the structure is about to disappear. The right image shows the phases of the linear scaled real field with the same filter

8.2 Quantifying the phase structure

The structure observed is a superposition of Fourier waves. This leads to the question: How to we quantify this structure? One observation that helps solve this problem can be found by applying the phase gradient to the Dirac delta induced waves:

$$\nabla_{\phi}(e^{i\vec{x}_0 \cdot \vec{k}}) = \vec{\nabla}(\vec{x}_0 \cdot \vec{k}) = \vec{x}_0, \quad (67)$$

the Dirac delta functions translate to a constant value in the phase gradient. A two dimensional view of this can be seen in Figure 28.

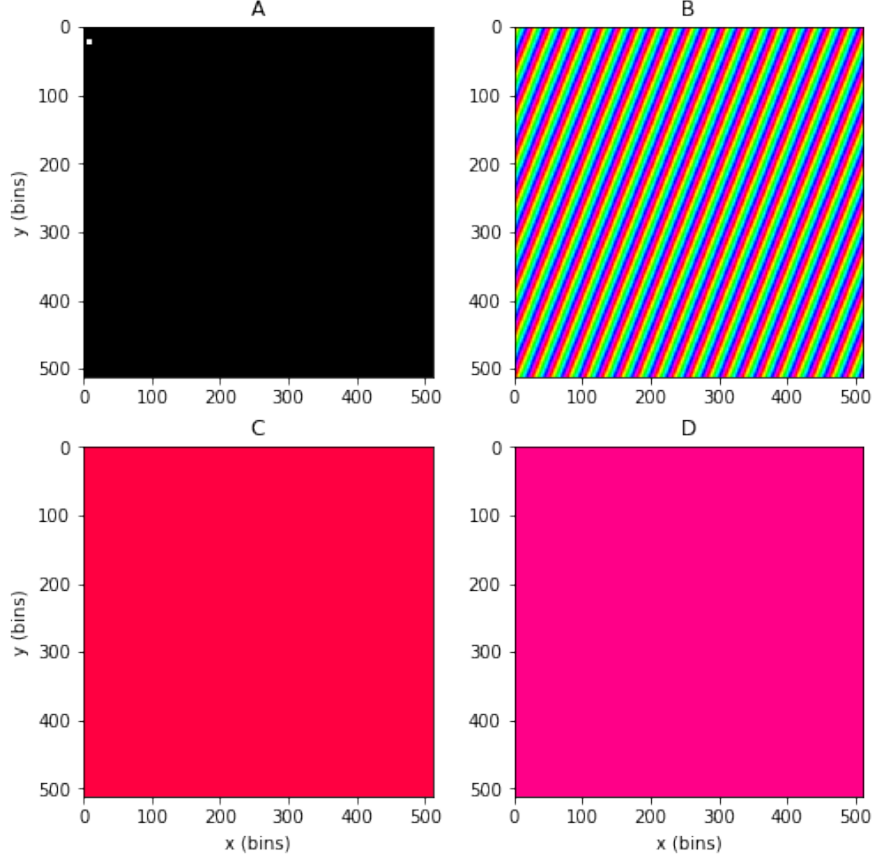


Figure 28: A two dimensional numerical Fourier transform of a delta function. (A) shows the 512x512 grid of zeros with one point being one acting like a Dirac delta. (B) shows the phases obtained from Fourier transforming the grid from (A). (C) and (D) show the phases gradients in the y(C) and x(D) directions.

We can use the constant gradient to quantify phase structure. We will assume that we have a Gaussian random field in which delta functions emerge as a function of expansion factor. At the start this will mean that each slice will have a very well defined mean due to the phases being uniformly distributed in between 0 and 2π . This random behaviour is also carried over into the gradient. The Dirac deltas will be uniformly distributed and cause Fourier phases to have varying slopes, since this depends on their positions (\vec{x}_0 in Equation 67). Once a delta function is added to a certain slice the mean value of the gradient will shift from π to the value of the constant slope seen in Equation 67. The mean of the gradient field was equal for all slices when they were randomly distributed but the individual means now take on a range of values as more and more Dirac delta functions are added. This causes the distribution of gradient means to slowly drift away from the original mean value, like a dissipation process were a narrow distribution becomes a normal distribution with a standard deviation that increases as a function of time. This standard deviation is what can quantify phase structure and from now on will be referred to as σ_{∇} . The mathematical definition of σ_{∇} can be found in Section 5.4.

8.3 Quantifying σ_{∇}

In order to find what σ_{∇} quantifies the models made with the Voronoi code were used. These models are detailed in Section 6.3. For each of the models σ_{∇} is calculated for three coordinates (x, y and $\sqrt{x^2 + y^2}$). $\sigma_{1.6}$ and the percentage of empty cells is also calculated for each time step. $\sigma_{1.6}$ is a small modification of σ_8 where the sum is simply done over each density bin with each bin having a size of approximately 1.6 Mpc. $\sigma_{1.6}$ relates closely to the amount of structure and is also a good tracer of extreme overdensities. The percentage of empty cells is used as a very simple way to have some trace of underdensities. The results can be seen in Figure 29, 30, 31 and 32.

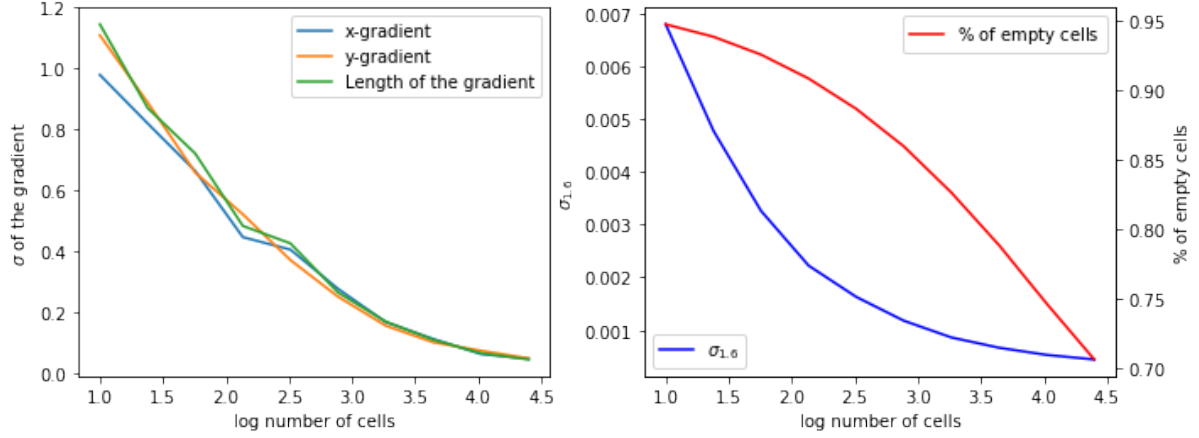


Figure 29: In the left plot σ_{∇} as a function of number of cells/nodes. The right plot shows $\sigma_{1.6}$ and the percentage of empty cells as a function of the number of cells.

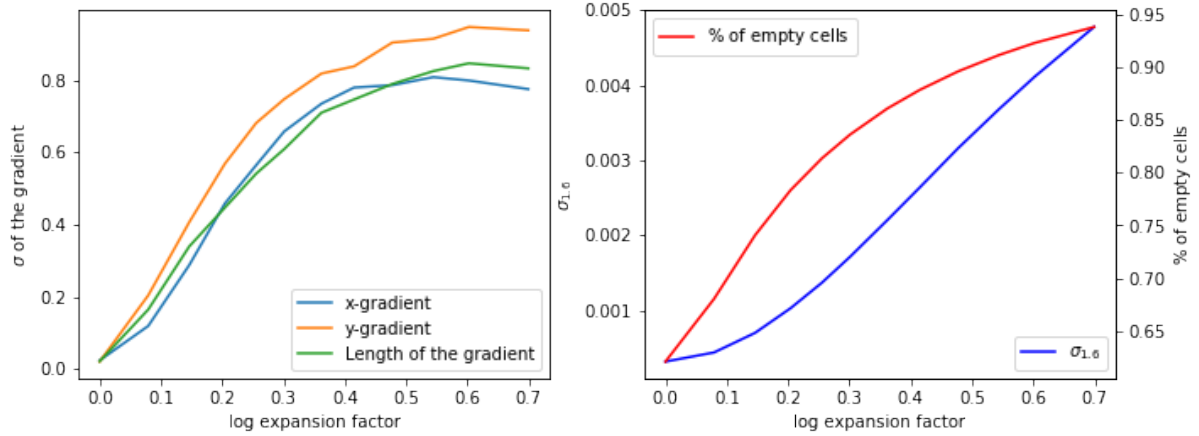


Figure 30: In the left plot σ_{∇} as a function of expansion factor. The right plot shows $\sigma_{1.6}$ and the percentage of empty cells as a function of the expansion factor.

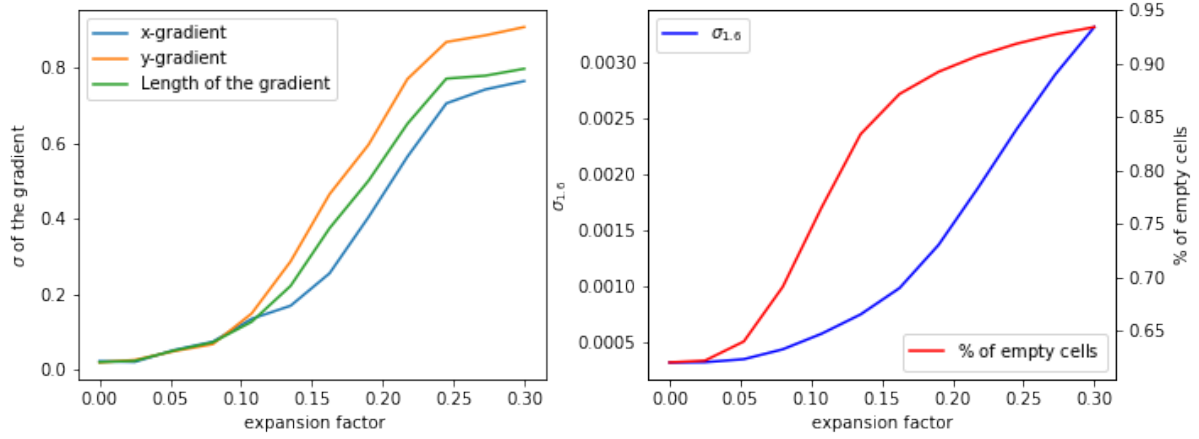


Figure 31: In the left plot σ_{∇} as a function of expansion factor. The right plot shows $\sigma_{1.6}$ and the percentage of empty cells as a function of the expansion factor.

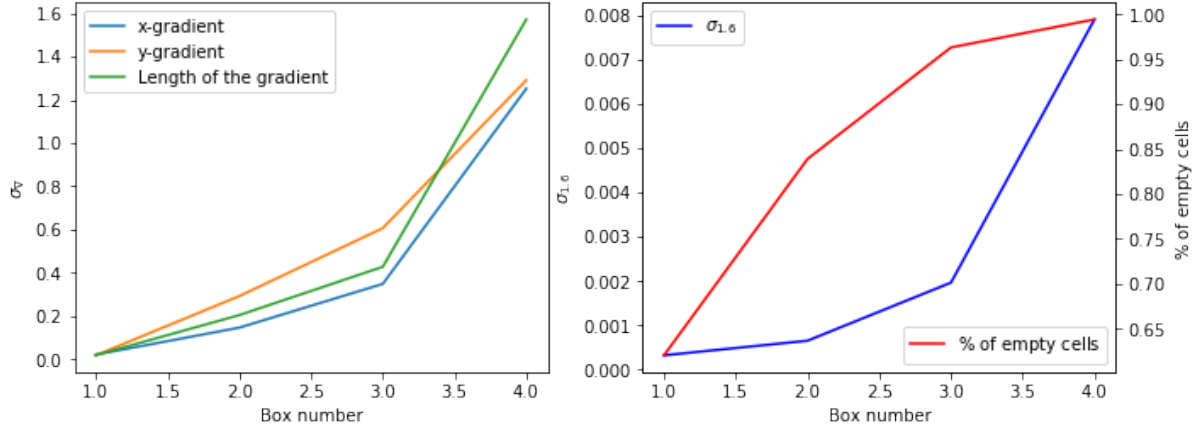


Figure 32: σ_{∇} for 4 limiting cases during cosmic structure formation. Starting of with random (Box number 1) then all particles are put in the walls (Box number 2) then all particles are put in filaments (Box number 3) and finally all particles are put in clusters (Box number 4).

Figure 29 shows the results of increasing the number of Voronoi cells. Here we see that there appears to be a close relation between $\sigma_{1.6}$ and σ_{∇} . All of these boxes were taken at expansion factor 5 and were thus fully evolved. Due to the constant amount of particles what is seen also has major numerical artifacts. Adding more cells does not only add more overdense regions but it also greatly reduces how overdense the regions eventually become. The final state of the box (with 25000 cells) does not have many very dense regions spread over the box but is instead more like a random field and the cells are completely unresolved. Due to the field being completely evolved we do still have overdense regions that dominate. This causes σ_{∇} to have behaviour close to $\sigma_{1.6}$ and mostly traces the overdense regions.

Figure 30 shows the results of simply allowing the universe to evolve as a function of expansion factor starting from an initially Gaussian field. The box has 25 evolving cells. The behaviour here is different from the results from Figure 29 as σ_{∇} now evolves with little to no direct relation to $\sigma_{1.6}$. The behaviour can now instead be traced to the percentage of empty cells. Due to the number of particles being relatively low the first structures that form are not overdense region but are instead empty regions. The empty regions are thus the dominant extreme in the box and the phase correlations will mainly be caused by very underdense regions.

Figure 31 shows the results of the bubble explosion model and further confirms this hypothesis. Here we see that the behaviour of σ_{∇} is not linked to strictly one of the other quantities

but instead lies in between the two quantities. This can be explained by the fact that this model creates density shockwaves that leave behind truly empty areas. It forms overdense structure (the shockwaves) and very underdense structure at the same time and both extremes now impact the behaviour of σ_{∇} .

Figure 32 shows the result of the hyperevolution model. here σ_{∇} again follows closer to $\sigma_{1.6}$. The boxes all have lot's of structure already formed (except of course the first point) and this will thus be the dominant component in the phases. This also confirms the observations found for Figure 29 as both experiments had boxes with dominant structure and in both σ_{∇} has behaviour close to $\sigma_{1.6}$

The phases need to be heavily correlated in order to allow for these very extreme regions in real space. σ_{∇} appears to be a tracer of not just the very overdense regions but also for very underdense regions. For regular cosmological simulations the underdense region can only be a maximum of 1 under the average as $\delta \geq -1$. The overdensities can be several orders of magnitude above the average causing them to be far more dominant. In these simple simulations the overdensities are far better behaved, the low number of particles caused the overdensities to reach values much lower then usual N-Body simulations, which allows for both under and overdensities to leave their mark in σ_{∇} . This does give insight into the true nature of σ_{∇} namely that it traces the extremes in the cosmic density field.

8.3.1 σ_{∇} for N-body simulations

The value of σ_{∇} was also calculated for the cosmological simulation mentioned in section 6.3.2 for each expansion factor. The results can be seen in Figure 33:

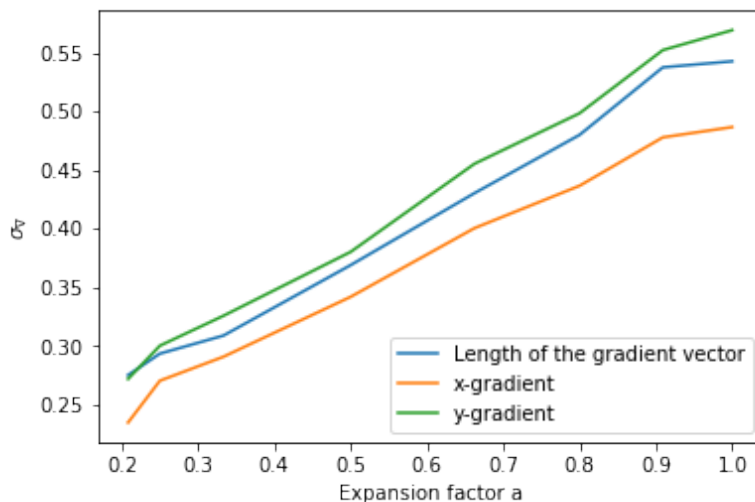


Figure 33: σ_{∇} as a function of expansion factor. The value is shown for the x direction, the y direction and for the length of the gradient.

In Figure 33 σ_{∇} follows a near linear increase with expansion value and again has behaviour very close to σ_8 that was plotted in Figure 19. For a simulation with a much larger number of particles it again traces σ_8 and is much closer related to the extreme overdensities. A different way of displaying σ_{∇} is to display it as error-bars to the mean value of the gradient as can be seen in Figure 34:

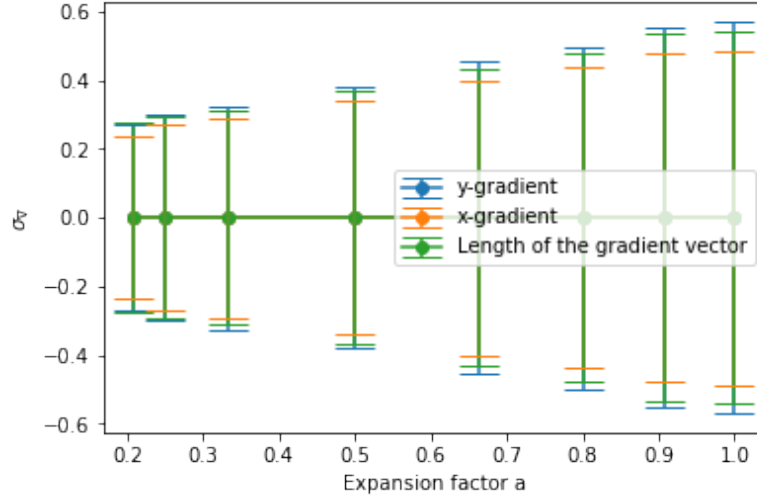


Figure 34: A plot displaying σ_{∇} as errorbars to the mean value of the gradient. The mean value of all the mean value was subtracted from the plotted points and the points show the deviation of the total mean (even though this is barely noticeable).

While this may not be the nicest way to actually analyse σ_{∇} it does give insight into what the quantity is and it shows how stable the mean value is. It also shows how strong the increase of σ_{∇} actually as it's still very apparent even when compared to the mean value.

8.4 Phase shifts to track cosmic structure formation

B. S. Ryden and Gramann (1991) discuss a different formalism that uses the phases to analyse cosmological simulations. The universe starts off as a Gaussian density field which implies that the phases are distributed randomly and there is no information in the phases. As time evolves the phases must start drifting away from their initially random phases in order to form the structures that can be seen in Figure 25. To quantify the shift from the original phase the quantity $|\Delta\phi|$ is defined as:

$$|\Delta\phi_i(\vec{k})| = |\phi_i(\vec{k}) - \phi_0(\vec{k})|, \quad (68)$$

where ϕ_0 is the phase of the initially Gaussian field and ϕ_i is the value of the phase at a later point in cosmic time. This can then be further generalised and $|\Delta\phi_i|$ can be made as a function of $|\vec{k}|$ which is what is done in the original paper. How the powerspectrum evolves after the linear phase has been studied in great detail with contributions by for example Peacock and Dodds (1996) and Smith et al. (2003) and is a heavily researched topic. Looking at when and how quickly certain phases start drifting away from Gaussianity can give insight in the evolution of the universe after the phase of linear growth. The results of applying this analysis on an N-body simulation with 8 different expansion factors, taking the first one as ϕ_0 , can be found in Figure 35:

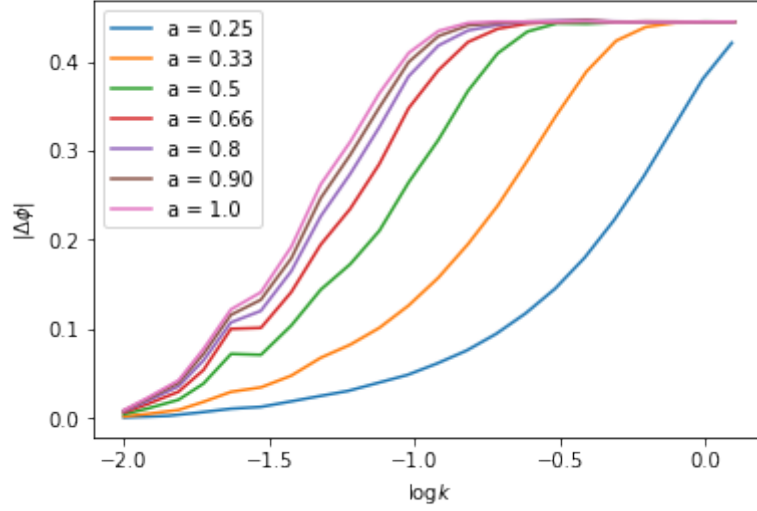


Figure 35: $|\Delta\phi|$ as a function of $\log k$ for 7 different expansion factors. $|\Delta\phi|$ is in units of π and k is in units of Mpc^{-1} . The data was rebinned into 25 equally spaced bins in $\log k$ and in each bins the average $\log k$ and $|\Delta\phi|$ were calculated

$|\Delta\phi|$ tends to $\pi/2$ for large k . The side effect of using the absolute shift is that $|\Delta\phi|$ can only have values in between 0 and π . The degeneracy in the shift comes from the fact that shift any above π or below $-\pi$ can also be expressed as lying in between π and $-\pi$. This causes the absolute range to only be in between 0 and π . The phases are random at first which means that the final difference in phase will also be distributed randomly over the range. Taking the average of the range of $|\Delta\phi|$ will give $\pi/2$ which implies that the phases will have completely drifted away from random when $|\Delta\phi|$ reaches $\pi/2$. The shifts for higher wavenumbers which translate to small scales also reach their non-random state earlier than the shifts for lower wavenumbers which is also what is expected since structure forms earlier at smaller scales. B. S. Ryden and Gramann (1991) look at which wavenumber the curve approaches $\pi/4$ in order to quantify at which expansion factor the phases drift from randomness, while this is a valid way of quantifying the observed curves it can be subject to noise. One way of quantifying the shape of the curves in a more detailed way is by fitting them with a simple model that can give insight in the slope of the curve and where it is located, one function that can fit this type of curve and gives information about the slope and position is the S curve:

$$f(x) = \frac{A}{1 + e^{-\frac{x-x_0}{x^*}}}, \quad (69)$$

where A is the value it goes to for large x , x_0 is an offset and x^* is a scale length which sets the slope of the S curve. First all the values were fitted but it was found that for the first couple of expansion values A would be a lot higher than $\pi/2$ which would be nonphysical, this is caused by the fact that we are missing the top half of the distribution for these expansion factors. As this is a numerical effect we are justified to set A to the correct value. A was set to the value fitted for the larger expansion factors which was approximately $\pi/2$ namely $\approx 0.453\pi$. This value was found for all the larger expansion factors which justifies its use. Both the offset and the scale length were fitted as a function of $\log a$. For the scale length a simple power law model proved to be a good fit. The offset was more difficult to fit and a broken power law found also used by Smith et al. (2003) was used:

$$y(x) = \frac{ax^3}{1 + bx + cx^{3-d}}. \quad (70)$$

The results of the fits can be found in Figure 36:

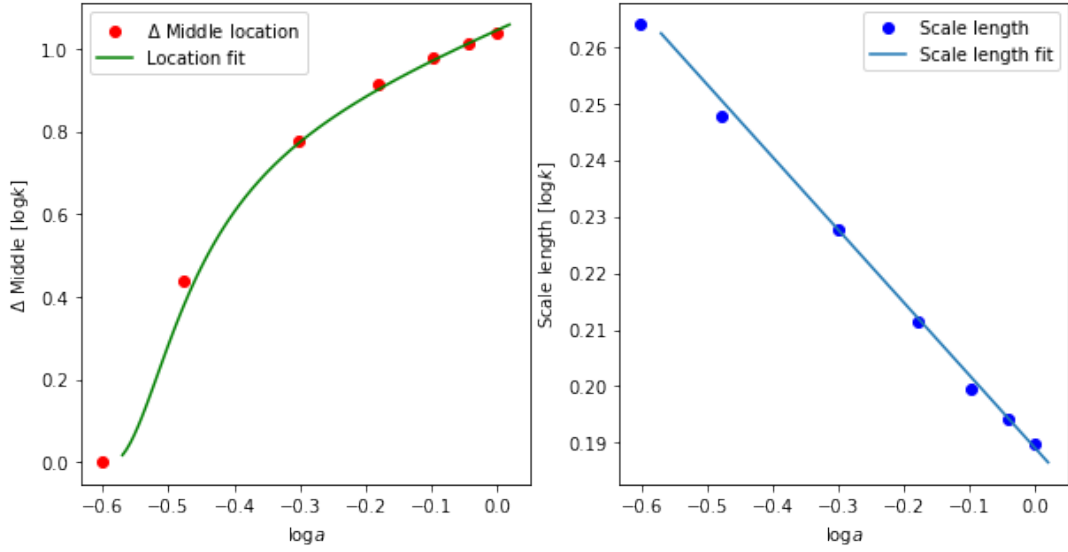


Figure 36: The results of the scale length and the offset of the fitted S-curves to the data in Figure 35. The left plot shows the offset and the right plot shows the scale length. The left image shows the result of the fit using Equation 70 and the left image shows the fit of a power law.

The offset has a more complex behaviour and needs to be compared with results from different simulations in order to possibly find a better way to describe it's behaviour. The scale length can be fit very well to a power law. If this power law behaviour would be intrinsic to this method of observing the phases it could prove to be a nice way to compare different simulations in a very detailed but still simple way. The result of the power law fit gave:

$$x^*(a) = (1.545 \pm 0.002) \cdot a^{-0.124 \pm 0.002}. \quad (71)$$

The broken power-law fit did not give convergent standard deviations on the fit parameters and a better fitting function is needed in order to be able to compare it with different results. While the fit for the offset is not perfect it does show that both quantities appear to follow a smooth predictable behaviour which could easily be compared to different simulations or observations.

8.5 The cosmic Taylor invariant

The Taylor invariant was previously introduced as a translation invariant way to analyse the Fourier phases. A translation invariant is naturally a very import framework to work in when performing analysis but it is interesting to see the effect of applying the Taylor invariant to the various methods used in this thesis. The methods used may intrinsically be invariant under translations which would greatly add to their reliability. For both quantities three version of translations were tested. For one the image is untranslated, for one the image is translated to it's center of mass (creating the Taylor invariant) for each expansion factor and the Taylor invariant is calculated for the first expansion factor and the translation is kept constant for all other expansion factors.

8.5.1 The invariance of σ_{∇}

The original calculation of σ_{∇} found in Figure 33 was done by with the original coordinate system. Below are the result for using the Taylor invariant before every Fourier transform. (Figure 37) and for a image translated to a constant point for each slice (Figure 38). The range of the plot is almost equal for all three plots and all small fluctuations could simply originate from numerical errors. In this case the most well behaved plot appears to be Figure 33. Figure 38 with the random initial translation should in theory be no different from Figure 33 since they technically

do the same thing as for both the translation is constant. The original Figure 37 actually appears to be the slight less well behaved, since this coincides with taking the Taylor invariant at each slice this might actually be worse for the analysis. Calculating the invariant for all slices causes the coordinate system to be different for every slice and this might causes small problems in the final result. The error due to this effect does appear to be quite small. In general σ_{∇} appears to be close to invariant under translations but a more thorough analysis should be done with a larger set of data in order to confirm this further.

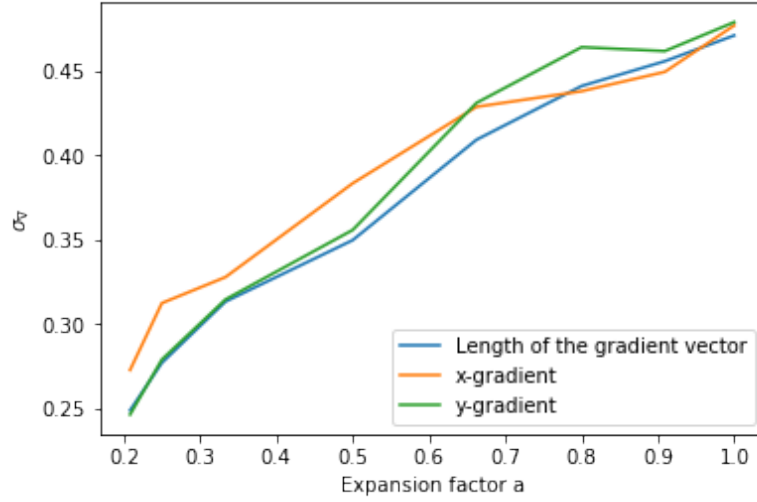


Figure 37: The result of calculating σ_{∇} and taking the Taylor invariant before every Fourier transform (for every slice). σ_{∇} is shown as a function of expansion factor

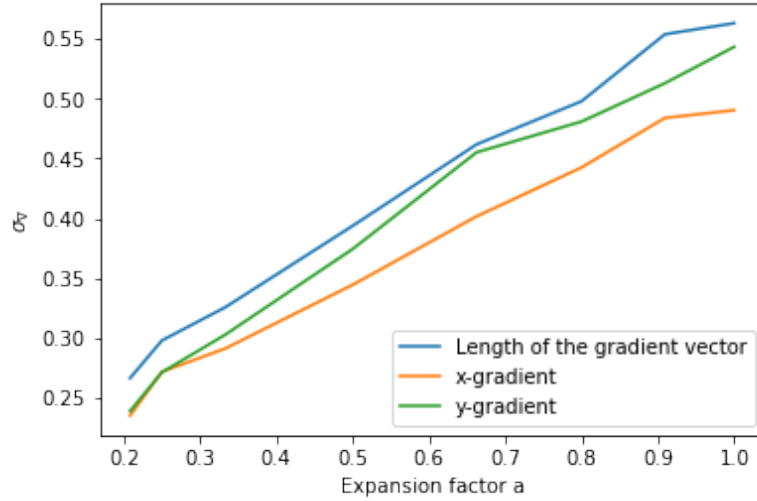


Figure 38: σ_{∇} as a function of expansion factor after a random initial translation

8.5.2 The invariance of $|\Delta\phi|$

The original box taken for Figure 35 is the raw density box. I does not use the Taylor invariant and it is untranslated. Here again the starting box was translated and in this case the translation was for the Taylor invariant for the lowest expansion factor (Figure 39). Also the Taylor invariant was calculated for each expansion factor (Figure 40). Figure 39 and Figure 35 both appear to be very well behaved and their behaviour is very comparable. The big difference is seen in Figure 40 which has a very odd shape and has a lot more noise. This is also the only plot where lines

cross. This method of analysis works by comparing phases of the same wavenumber (and in this case the fully general \vec{k}). In a perfect world the center of mass and thus the Taylor invariant would be conserved over time but this is not what is seen. The size of the translation to acquire the Taylor invariant changes in between expansion factors and this causes problems for this kind of method. The translations causes the original phases to not always exactly co-inside with the original phases. This is a relatively minor effect but it does create huge nonphysical artifacts like the bump seen before it flattens out. For this method it is important that the coordinate system does not change when evolving the box. Besides that it appears to be invariant under translations when the translation is constant during the entire analysis.

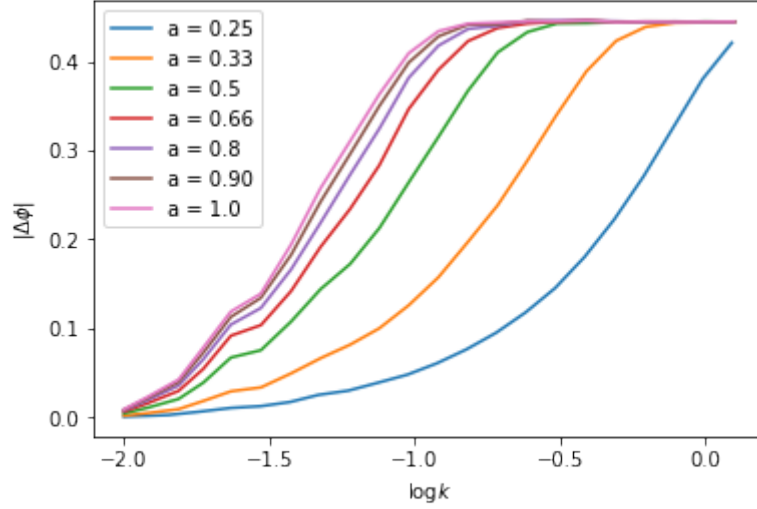


Figure 39: σ_{∇} as a function of expansion factor without using the Taylor invariant

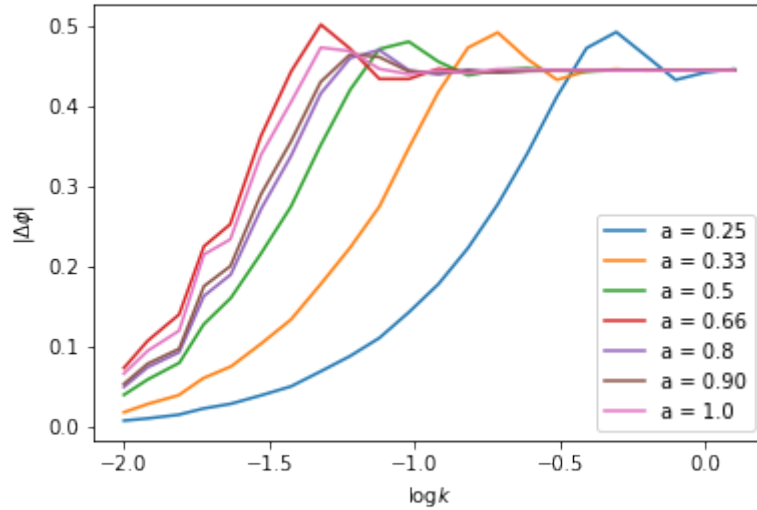


Figure 40: σ_{∇} as a function of expansion factor after a random translation

9 Conclusion

The goal of this thesis was to create an invariant basis to analyse Fourier phases and to find a way to quantitatively describe the structure found in the Fourier phases. The Taylor invariant is found as a translation invariant way to analyse the Fourier phases. The Taylor invariant is found by shifting the center of mass to the origin. The Taylor invariant was tested and found to work for simple images but great care needs to be taken as small errors in determining the Taylor invariant can be detrimental to the results.

The wavelike structure found when looking at the Fourier phases appears to originate mostly from the very overdense regions. In order to quantify the wavelike structure found in the phases the quantity σ_{∇} is defined. σ_{∇} makes use of the fact that means of the phase gradient different 2D slices start drifting from the normal average as the waves that form have in constant value in the gradient. σ_{∇} is found to be a tracer of extreme regions in space, both under and overdense regions. For cosmological simulations it follows mostly overdensities as they diverge further away from the mean density than the underdensities.

A different way of analysing the phases is by looking at how the initial phases shift away from random. The phases shift away from the initial conditions smoothly starting with the phases at large values of k . The curves obtained at each redshift were able to be fitted by S-curves. The parameters of the curve create smooth curves which can easily be compared in between different simulations. For a Λ CDM simulation the slope of the S-curve follow an almost perfect power law with parameters: $x^*(a) = (1.545 \pm 0.002) \cdot a^{-0.124 \pm 0.002}$.

Both σ_{∇} and the shifts were tested for invariance and appear to be independent of initial coordinate system. As long as the coordinate system is kept the same during the evolving of the box. Using different coordinate systems during one iteration of the analysis does cause problems but this is to be expected.

Both methods of analysis have a smooth behaviour and this allows for the results to be compared for different simulations. The analyses have to be applied to a wider range of simulations to confirm this and to gain more insight into what the results imply in real space.

References

- Adler, R. J. (1981). *The Geometry of Random Fields*.
- Aragón-Calvo, M. A. et al. (2010). “The Spine of the Cosmic Web”. In: *ApJ* 723, pp. 364–382. DOI: 10.1088/0004-637X/723/1/364. arXiv: 0809.5104.
- Bai, Linge and David Breen (2008). “Calculating Center of Mass in an Unbounded 2D Environment”. In: *Journal of Graphics Tools* 13.4, pp. 53–60. DOI: 10.1080/2151237X.2008.10129266.
- Bardeen, J. M. et al. (1986). “The statistics of peaks of Gaussian random fields”. In: *ApJ* 304, pp. 15–61. DOI: 10.1086/164143.
- Bond, J. R., L. Kofman, and D. Pogosyan (1996). “How filaments of galaxies are woven into the cosmic web”. In: 380, pp. 603–606. DOI: 10.1038/380603a0. eprint: astro-ph/9512141.
- Bos, Evert Gerardus Patrick (2016). “Clusters, voids and reconstructions of the cosmic web”. English. PhD thesis. University of Groningen. ISBN: 978-90-367-9282-0.
- Cautun, M. et al. (2014). “Evolution of the cosmic web”. In: *MNRAS* 441, pp. 2923–2973. DOI: 10.1093/mnras/stu768. arXiv: 1401.7866.
- Chiang, L.-Y. and P. Coles (2000). “Phase information and the evolution of cosmological density perturbations”. In: *MNRAS* 311, pp. 809–824. DOI: 10.1046/j.1365-8711.2000.03086.x. eprint: astro-ph/9905250.
- Coles, P. (2005). “Phase Correlations and Topological Measures of Large-scale Structure”. In: *ArXiv Astrophysics e-prints*. eprint: astro-ph/0502088.
- Coles, P. and L.-Y. Chiang (2000). “Characterizing the nonlinear growth of large-scale structure in the Universe”. In: *Nature* 406, pp. 376–378. DOI: 10.1038/35019009. eprint: astro-ph/0006017.
- Davis, M. and P. J. E. Peebles (1983). “A survey of galaxy redshifts. V - The two-point position and velocity correlations”. In: *ApJ* 267, pp. 465–482. DOI: 10.1086/160884.
- Dicke, R. H. et al. (1965). “Cosmic Black-Body Radiation.” In: *ApJ* 142, pp. 414–419. DOI: 10.1086/148306.
- Eisenstein, D. J. (1997). “An Analytic Expression for the Growth Function in a Flat Universe with a Cosmological Constant”. In: *ArXiv Astrophysics e-prints*. eprint: astro-ph/9709054.
- Guth, A. H. (1981). “Inflationary universe: A possible solution to the horizon and flatness problems”. In: 23, pp. 347–356. DOI: 10.1103/PhysRevD.23.347.
- Hardie, George, Storm Thorgerson, and Aubrey Powell (1973). *The Dark Side of the Moon*. Album by Pink Floyd: David Gilmour, Roger Waters, Rick Wright and Nick Mason.
- Harrison, E. R. (1970). “Fluctuations at the Threshold of Classical Cosmology”. In: 1, pp. 2726–2730. DOI: 10.1103/PhysRevD.1.2726.
- Icke, V. and R. van de Weygaert (1991). “The galaxy distribution as a Voronoi foam”. In: 32, pp. 85–112.
- Komatsu, E. et al. (2003). “First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Tests of Gaussianity”. In: *ApJS* 148, pp. 119–134. DOI: 10.1086/377220. eprint: astro-ph/0302223.
- Lin, Feng and Robert D. Brandt (1993). “Towards absolute invariants of images under translation, rotation, and dilation”. In: *Pattern Recognition Letters* 14.5, pp. 369–379. ISSN: 0167-8655. DOI: [https://doi.org/10.1016/0167-8655\(93\)90114-S](https://doi.org/10.1016/0167-8655(93)90114-S). URL: <http://www.sciencedirect.com/science/article/pii/016786559390114S>.
- Linde, A. D. (1982). “A new inflationary universe scenario: A possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems”. In: *Physics Letters B* 108, pp. 389–393. DOI: 10.1016/0370-2693(82)91219-9.
- Lorimer, D. R. and M. Kramer (2004). *Handbook of Pulsar Astronomy*.
- Peacock, J. A. and S. J. Dodds (1996). “Non-linear evolution of cosmological power spectra”. In: *MNRAS* 280, pp. L19–L26. DOI: 10.1093/mnras/280.3.L19. eprint: astro-ph/9603031.

- Peebles, P. J. E. (1980). *The large-scale structure of the universe*.
- Penzias, A. A. and R. W. Wilson (1965). “A Measurement of Excess Antenna Temperature at 4080 Mc/s.” In: *ApJ* 142, pp. 419–421. DOI: 10.1086/148307.
- Planck Collaboration, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, et al. (2014). “Planck 2013 results. XXIV. Constraints on primordial non-Gaussianity”. In: 571, A24, A24. DOI: 10.1051/0004-6361/201321554. arXiv: 1303.5084.
- Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, et al. (2016). “Planck 2015 results. XIII. Cosmological parameters”. In: 594, A13, A13. DOI: 10.1051/0004-6361/201525830. arXiv: 1502.01589.
- Romberg, W. (1955). “Vereinfachte numerische Integration”. In: *Kgl.Norske Vid.Selsk.Forsk.* 28.7. Cited By :101, pp. 30–36.
- Ruanaidh, Joseph J.K.Ö. and Thierry Pun (1998). “Rotation, scale and translation invariant spread spectrum digital image watermarking” This work was supported by the Swiss National Science Foundation (grant no. 5003-45334).” In: *Signal Processing* 66.3, pp. 303–317. ISSN: 0165-1684. DOI: [https://doi.org/10.1016/S0165-1684\(98\)00012-7](https://doi.org/10.1016/S0165-1684(98)00012-7). URL: <http://www.sciencedirect.com/science/article/pii/S0165168498000127>.
- Ryden, B. (2003). *Introduction to cosmology*.
- Ryden, B. S. and M. Gramann (1991). “Phase shifts in gravitationally evolving density fields”. In: *ApJL* 383, pp. L33–L36. DOI: 10.1086/186234.
- Shandarin, S. F. and Ya. B. Zeldovich (1989). “The large-scale structure of the universe: Turbulence, intermittency, structures in a self-gravitating medium”. In: *Rev. Mod. Phys.* 61 (2), pp. 185–220. DOI: 10.1103/RevModPhys.61.185. URL: <https://link.aps.org/doi/10.1103/RevModPhys.61.185>.
- Smith, R. E. et al. (2003). “Stable clustering, the halo model and non-linear cosmological power spectra”. In: *MNRAS* 341, pp. 1311–1332. DOI: 10.1046/j.1365-8711.2003.06503.x. eprint: astro-ph/0207664.
- Totsuji, H. and T. Kihara (1969). “The Correlation Function for the Distribution of Galaxies”. In: 21, p. 221.
- van de Weygaert, R. (2002). “Froth across the Universe Dynamics and Stochastic Geometry of the Cosmic Foam”. In: *ArXiv Astrophysics e-prints*. eprint: astro-ph/0206427.
- (2003). “The cosmic foam: stochastic geometry and spatial clustering across the universe”. In: *Statistical Challenges in Astronomy*. Ed. by E. D. Feigelson and G. J. Babu, pp. 175–196.
- (2016). *Lecture Notes: Cosmic Structure Formation*.
- van de Weygaert, R. and J. R. Bond (2008). “Observations and Morphology of the Cosmic Web”. In: *A Pan-Chromatic View of Clusters of Galaxies and the Large-Scale Structure*. Ed. by M. Plionis, O. López-Cruz, and D. Hughes. Vol. 740. Lecture Notes in Physics, Berlin Springer Verlag, p. 24. DOI: 10.1007/978-1-4020-6941-3_11.
- van de Weygaert, R. and V. Icke (1989). “Fragmenting the universe. II - Voronoi vertices as Abell clusters”. In: 213, pp. 1–9.
- Voronoi, Georges (1908). “Nouvelles applications des paramètres continus à la théorie des formes quadratiques. Deuxième mémoire. Recherches sur les paralléloèdres primitifs.” In: *Journal für die reine und angewandte Mathematik* 134, pp. 198–287. URL: <http://eudml.org/doc/149291>.
- Zel’dovich, Y. B. (1970). “Gravitational instability: An approximate theory for large density perturbations.” In: 5, pp. 84–89.
- Zeldovich, Y. B. (1972). “A hypothesis, unifying the structure and the entropy of the Universe”. In: *MNRAS* 160, 1P. DOI: 10.1093/mnras/160.1.1P.

A Appendix

A.1 Numerical Mathematical tools

In order to be able to analyze the data certain mathematical operators are needed to be able to numerically analyse the data. In this section the method of integration and derivatives are derived and described.

A.1.1 Romberg integration

For normalization of P_{cdm} a numerical integrator is needed. Due to the nature of the integral (having lots of narrow structure) Romberg integration was chosen as the method of integration (Romberg (1955)). Romberg integration is an implementation of trapezoid integration that allows for integration up to a set error which can reduce computation time. It does this by calculating the integral at better and better intervals (starting with the entire integration range) and extrapolating the possible error. When the error is below the given limit it stops and returns the value of the integral. In python the package `scipy` has a function that implements this namely `scipy.integrate.romberg` and this function was used to calculate integrals.

A.1.2 Numerical Gradient

The numerical gradient is constructed by looking at the official definition of the derivative:

$$\frac{\partial f}{\partial x} = \lim_{d \rightarrow 0} \frac{f(x) - f(x + d)}{d} \quad (72)$$

From this a very simple approximation can be made for the derivative when there only is a discrete set of x and y values. In this case it is impossible to take the limit to 0 d is instead chosen as a variable that can be used to set the scale at which the derivative is taken and for this thesis it is set at 2. If the definition is now expanded to two dimensions with d=2 we get:

$$\nabla_x f(n, m) = \frac{f(n + 1, m) - f(n - 1, m)}{2} \quad (73)$$

$$\nabla_y f(n, m) = \frac{f(n, m + 1) - f(n, m - 1)}{2} \quad (74)$$

Where n and m are the pixel x and y coordinates of the 2d Fourier phase image. d = 2 is chosen instead of d = 1 to smooth out some of the small scale noise.

A.2 Residual due to round of errors

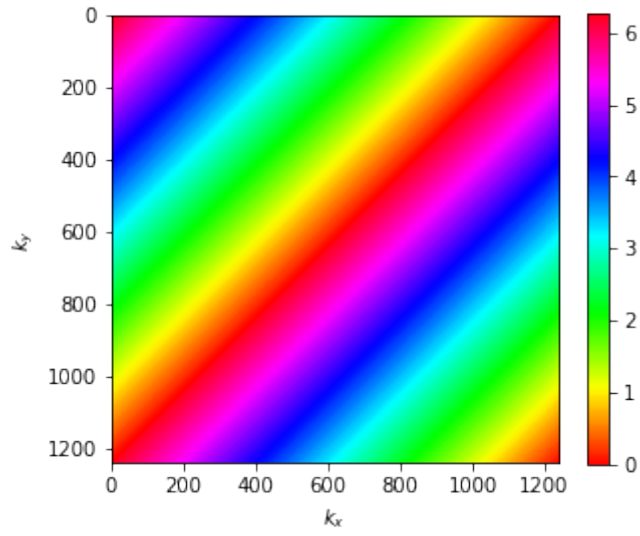


Figure 41: The result of calculating the residuals between two Taylor invariants of the same image when the image is translated to the center of mass with a rounding error. In this case the one of the Taylor invariants is shifted by one in both the x and y direction with respect to the other creating a diagonal wave pattern.

A.3 Voronoi 3D Boxes

A.3.1 Voronoi kinematic model experiment

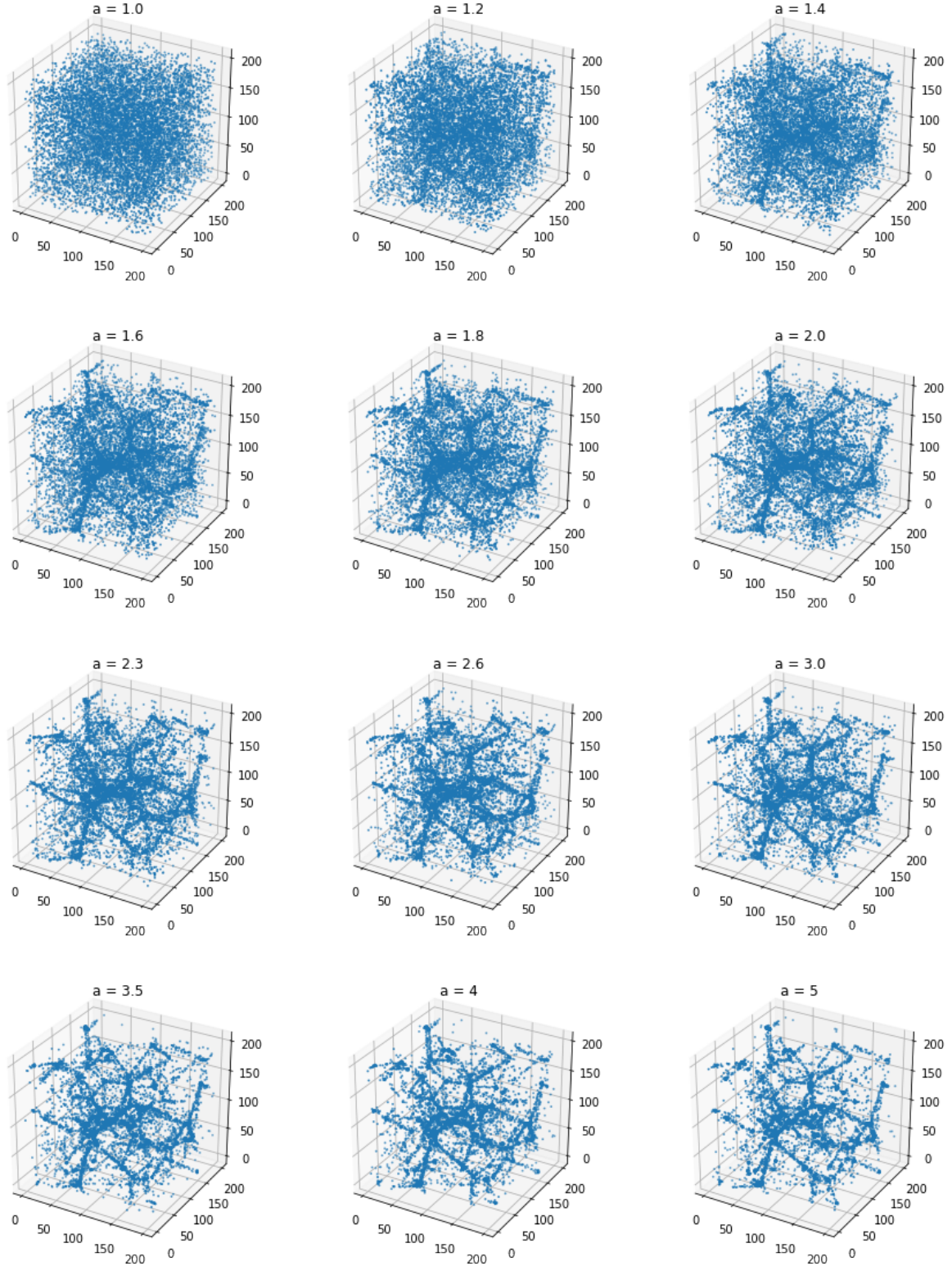


Figure 42: The 12 boxes used to inspect σ_V and allowing the box to evolve according to a Voronoi kinematic model. 10000 particles of the 1000000 are plotted

A.3.2 Voronoi elements experiment

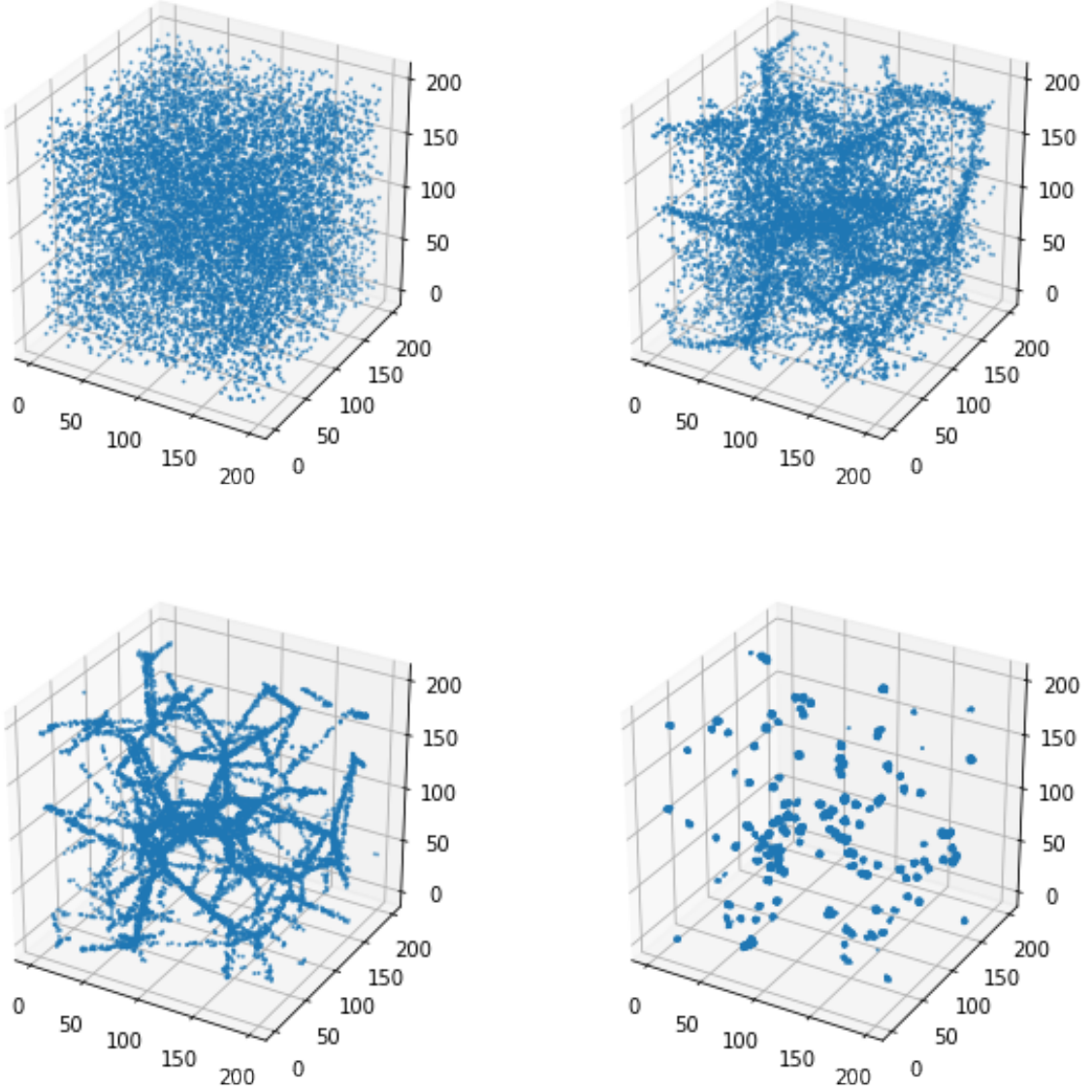


Figure 43: The 4 boxes used to inspect σ_{∇} in the 4 most extreme cases. The first box is the initial field, the second box has all particles in the walls, the third box has all particles in filaments and the final plot has all particles in clusters. 10000 particles of the 1000000 are plotted

A.3.3 Increasing cells experiment

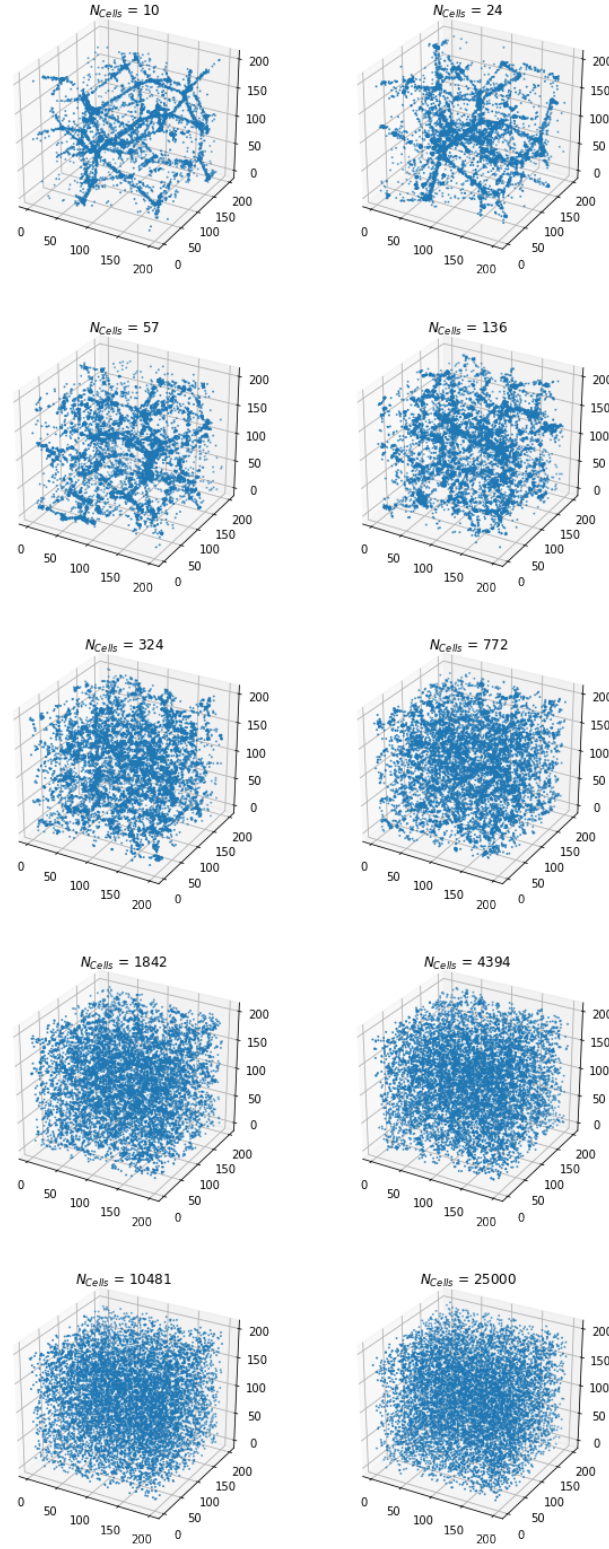


Figure 44: The 10 boxes used to inspect σ_{∇} with an increasing number of Voronoi cells in each box. 10000 particles of the 1000000 are plotted

A.3.4 Explosion model experiment

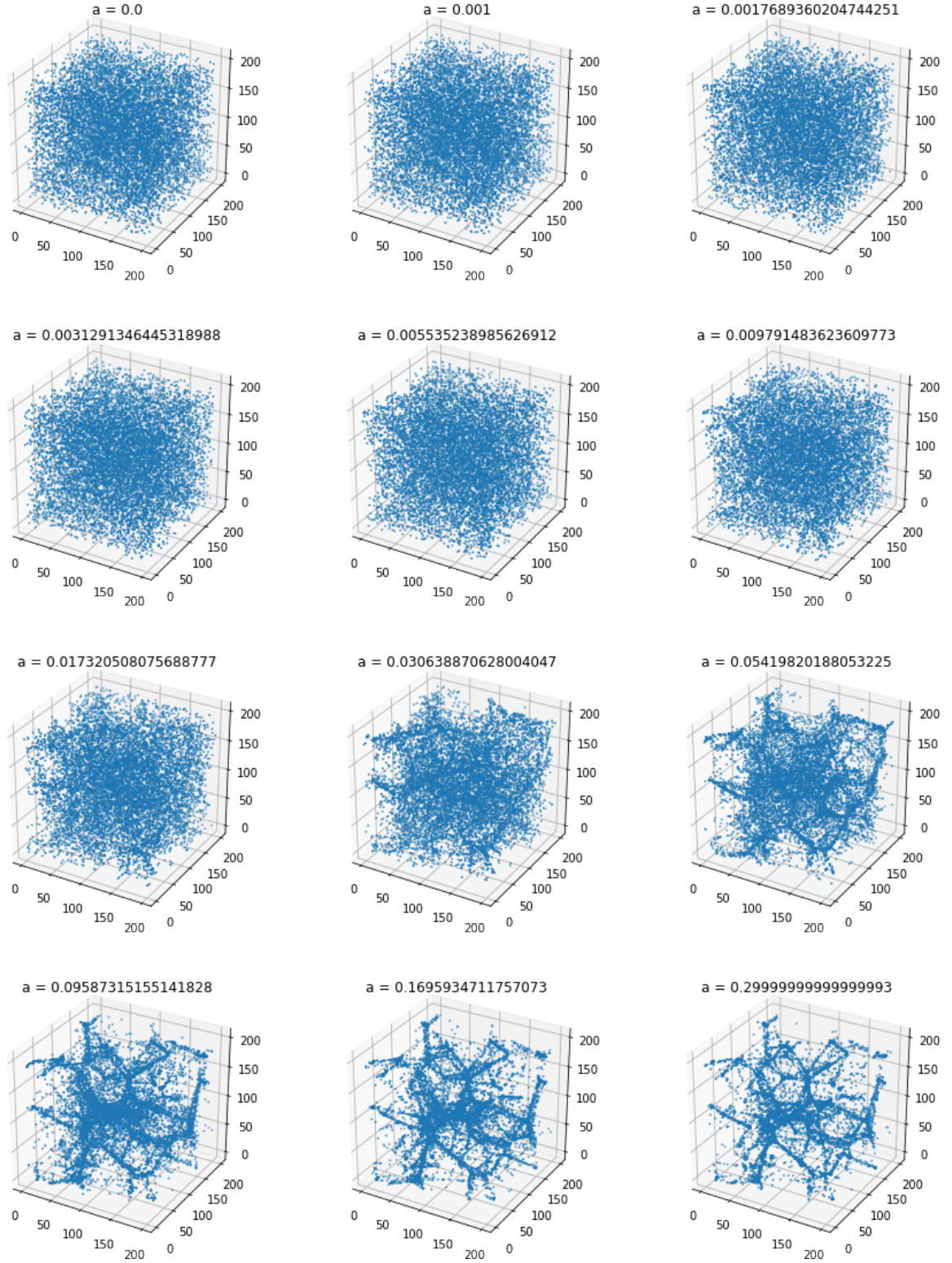


Figure 45: The 12 boxes used to inspect σ_{∇} and allowing the box to evolve according to a bubble explosion model. Here a is the avg distance between particle and wall. 10000 particles of the 1000000 are plotted

A.4 Voronoi Slices

A.4.1 Voronoi kinematic model experiment

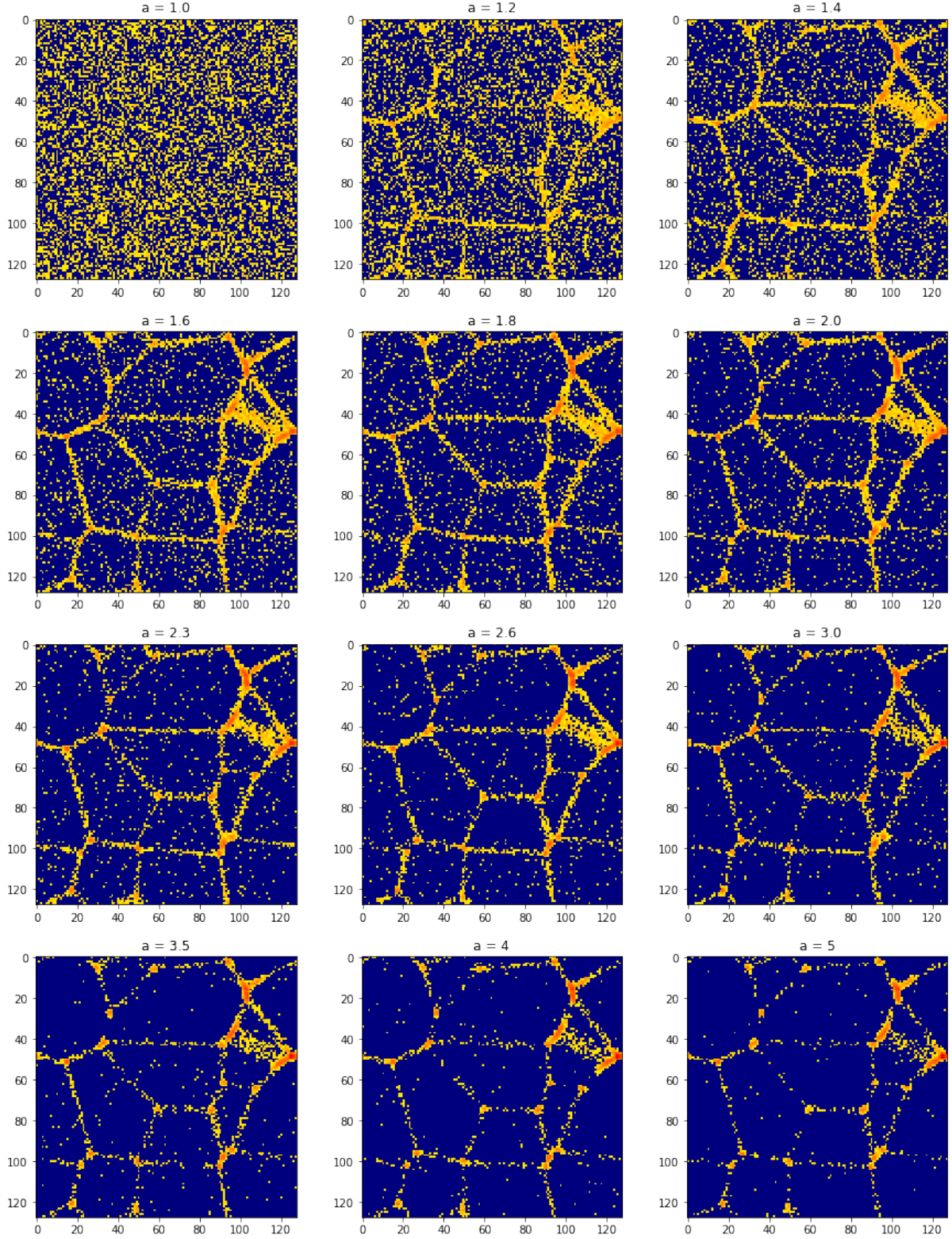


Figure 46: A slice out of the density fields of the 12 boxes used to inspect σ_{∇} and allowing the box to evolve according to a Voronoi kinematic model.

A.4.2 Voronoi elements experiment

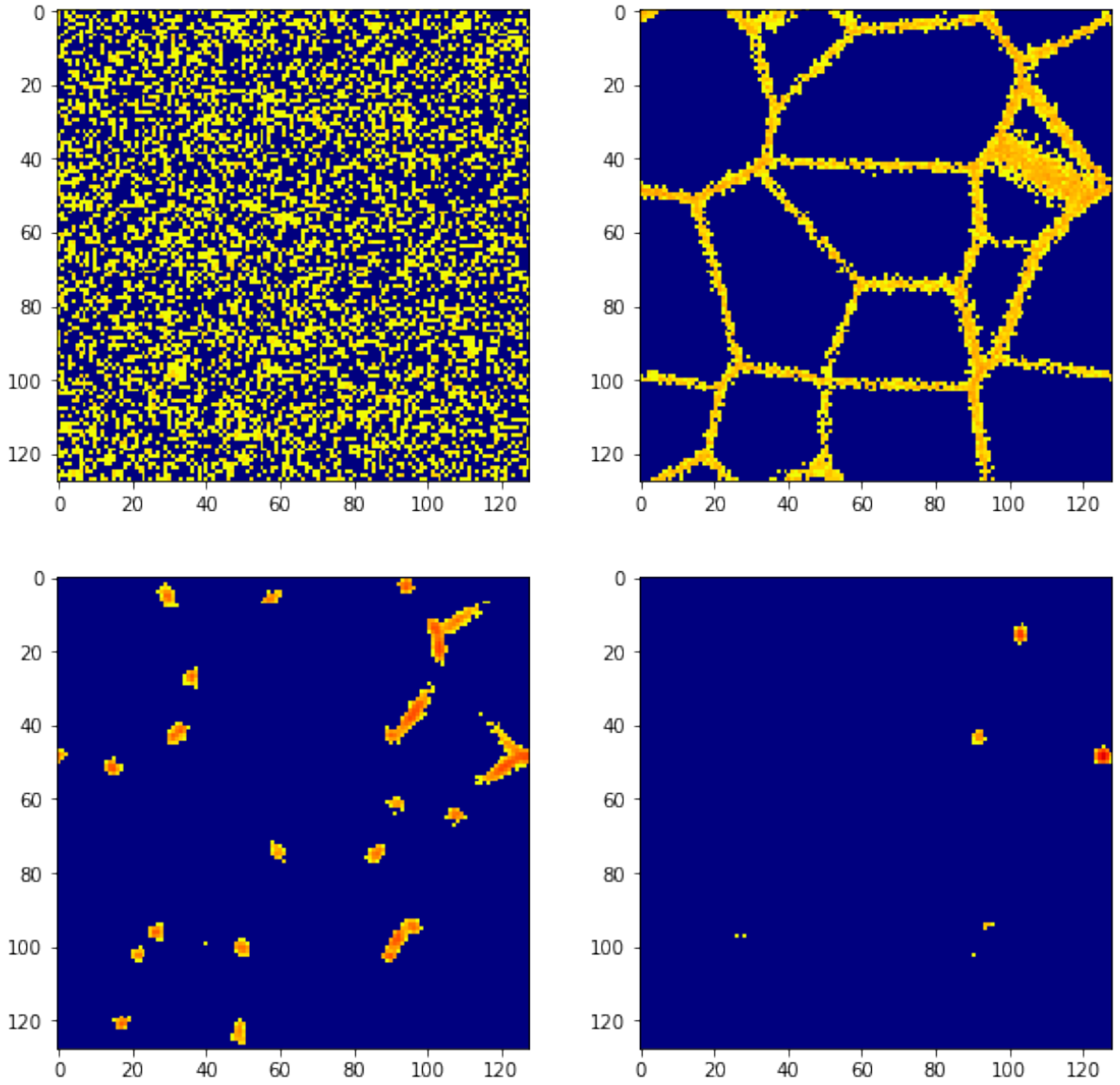


Figure 47: A slice out of the density fields from the 4 boxes used to inspect σ_{∇} In the 4 most extreme cases. The first box is the initial field, the second box has all particles in the walls, the third box has all particles in filaments and the final plot has all particles in clusters.

A.4.3 Increasing cells experiment

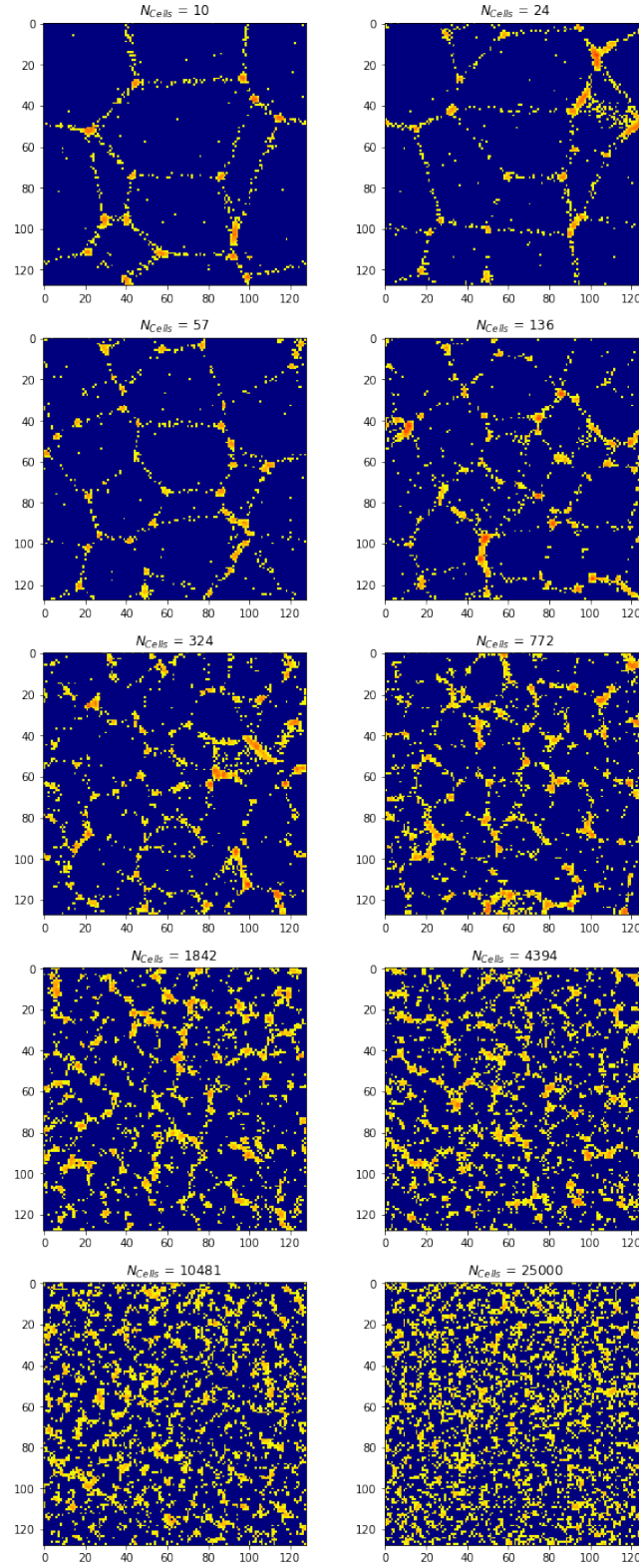


Figure 48: A slice of the density fields of the 10 boxes used to inspect σ_{∇} with an increasing number of Voronoi cells in each box.

A.4.4 Explosion model experiment

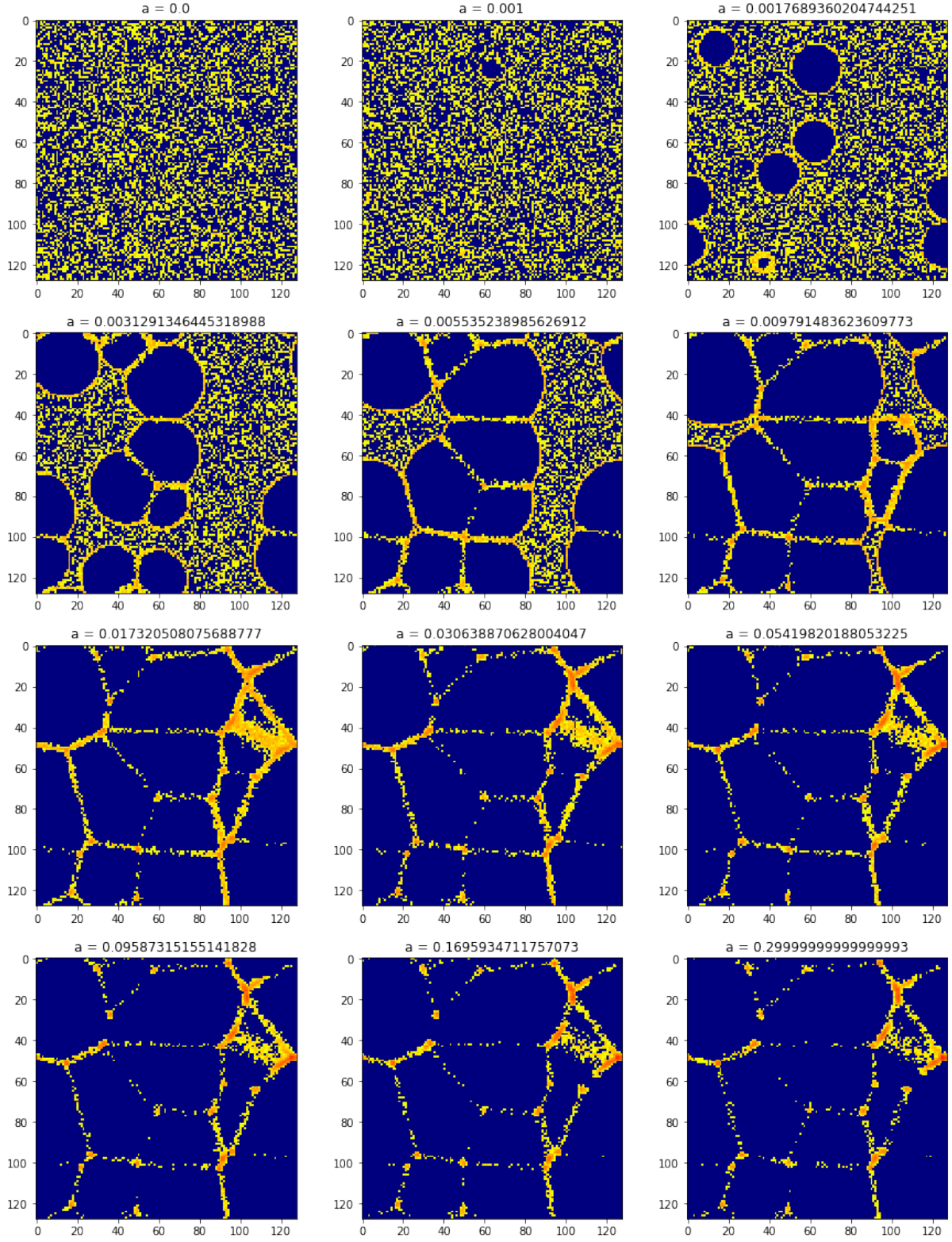


Figure 49: A slice out of the density fields of the 12 boxes used to inspect σ_V and allowing the box to evolve according to a bubble explosion model. Here a is the avg distance between particle and wall.