

Department of Physics and Astronomy

University of Heidelberg

Master thesis

in Physics

submitted by

Davide Mapelli

born in Milan

2024

Eigenvalue spectrum
of random autocorrelation matrices
of classical particle ensembles

This Master thesis has been carried out by Davide Mapelli

at the

Institute of Theoretical Physics, Heidelberg

under the supervision of

Prof. Dr. Matthias Bartelmann

and

Prof. Dr. Tilman Enss

Abstract

Kinetic Field Theory (KFT) is a very general tool that fully describes the behavior of ensembles of classical particles. This work focuses on KFT applied to the formation of large-scale structures in the universe, such as galaxy clusters, filaments, and walls. Its formalism gives rise to a generating functional that depends on density and momentum covariance matrices, which are sensible to the initial conditions of particle distribution in phase space. In more detail, the eigenvalues spectrum of these random momentum and density autocorrelation matrices are investigated with both analytical and numerical methods. Mainly methods from Random Matrix Theory (RMT) are used in this work, together with basic tools from probability theory. Specifically, approximated expressions for the high and low-density limits are provided, and their applicability to KFT is discussed.

Zusammenfassung

Die kinetische Feldtheorie (KFT) ist ein sehr allgemeines Werkzeug, welches das Verhalten von Ensembles klassischer Teilchen vollständig beschreibt. Diese Arbeit konzentriert sich auf die Anwendung der KFT auf die Bildung großskaliger Strukturen im Universum, wie Galaxienhaufen, Filamente und Mauern. Ihr Formalismus führt zu einem erzeugenden Funktional, das von Dichte- und Impulskovarianzmatrizen abhängt, die empfindlich von den Anfangsbedingungen der Teilchenverteilung im Phasenraum abhängen. Im Detail wird das Eigenwertspektrum dieser zufälligen Impuls- und Dichteautokorrelationsmatrizen sowohl mit analytischen als auch numerischen Methoden untersucht. In dieser Arbeit werden hauptsächlich Methoden aus der Random-Matrix-Theorie (RMT) verwendet, zusammen mit grundlegenden Techniken der Wahrscheinlichkeitstheorie. Speziell werden näherungsweise Ausdrücke für die Grenzen hoher und niedriger Dichte bereitgestellt, und ihre Anwendbarkeit auf die KFT wird diskutiert.

Contents

Introduction	1
1 Cosmology fundamentals	3
1.1 Basis of cosmology	3
1.1.1 Λ Cold Dark Matter model	5
1.1.2 Linear structure formation	5
1.2 Zel'dovich approximation	7
1.3 Zel'dovich tensor and its eigenvalues	8
2 Probability and random fields	11
2.1 Probability theory	11
2.1.1 Multiple random variables	12
2.1.2 Central limit theorem	13
2.2 Gaussian random fields	14
3 Kinetic Field Theory	18
3.1 Initial phase-space distribution	19
3.2 Covariance matrices	21
3.3 Generating functional	24
3.3.1 Density operator	24
3.3.2 Free generating functional	26
4 Random Matrix Theory	29
4.1 Matrix basis	29
4.2 Gaussian random matrices	30
4.2.1 Coulomb gas	33
4.2.2 Resolvent	34
4.3 Wishart-Laguerre matrices	35
4.4 Euclidean random matrices	36
4.4.1 High density expansion	37
4.4.2 Low density expansion	37
4.4.3 Considerations	38

5 Applications and results	42
5.1 Choice of correlations	42
5.2 Eigenvalues distribution of $C_{\delta\delta}$	44
5.2.1 High density limit	44
5.3 Eigenvalues distribution of \mathbf{C}_{pp}	49
5.3.1 Two particles	49
5.3.2 N particles	51
5.3.3 Individual probability distribution	55
5.4 Generating functional	57
5.4.1 Moments expansion	57
6 Conclusions	59
A Spherical Bessel functions	61
B Gaussian integrals	63
C Nearest neighbours	64
D Eigenvalues spectrum of the Zel'dovich tensor	65
E Vandermonde determinant	71
F Additional calculations	72
G Numerical considerations	74
Bibliography	81

Introduction

Glimpsing the Cosmos has always intrigued mankind. Unfortunately, this has only been possible in a proper way from the last century on, due to advanced technologies that allowed the scientific community to collect cosmic data. Thanks to the pioneering work of theorists like Einstein, Hubble, Friedmann, and many others, we started to unravel the mysteries of the Universe and gave the chance for the data to be placed in a proper theoretical context. One of the biggest challenges of Cosmology is to understand the processes that lead to the formation of observable structures that we can observe, such as voids, filaments and galaxy clusters. The latter have been observed by a galaxy survey conducted at $2 \mu\text{m}$ wavelength (the 2MASS survey [1]), [2,3]. The field of cosmic structure formation proposes to formulate a theory that could explain what allowed the *emergence* of these kinds of structures from an isotropic and homogeneous configuration. We can consider the Cosmic Microwave Background (CMB) as a Gaussian random field [4], whose temperature fluctuations, of magnitude $\approx 10^{-5}$, represent the density fluctuations roughly 380,000 years after the Big Bang [5]. Following the most accepted cosmology theory, namely the Λ Cold Dark Matter (Λ -CDM) model, we expect these structures to be formed from these initial inhomogeneities, all the way throughout time, until today. The arising of large-scale structures from Gaussian fluctuations has been proven to be a consequence of the gravitational collapse of over-dense regions in the Universe using the Zel'dovich approximation [6,7].

According to the cosmological standard model, what connects the amplitude of the initial CMB state and the structures we observe today should have been mediated by something that must not interact electromagnetically [8], hence the name, Dark Matter. Otherwise the structures arising from CMB temperature fluctuations would be two orders of magnitude larger. One of the main approaches is to settle the ideal hydrodynamics equations and consider them on an expanding Universe, together with Newtonian gravity. Then, one can perturbatively solve these equations and achieve an analytical formalism that describes the growth of structure formation [9]. Notwithstanding the great results achieved by this method, it turned out that this breaks down as soon as, in the fluid description of dark matter, matter flow converges. If the streams cross, a multi-valued velocity field arises, giving rise to the so-called *shell-crossing* problem. In this sense, Kinetic Field Theory (KFT) proposes to analytically pursue this logic, with the big advantage that particle trajectories naturally do not cross in phase space, avoiding the *shell-crossing* problem by definition. KFT should not be understood as a *ad hoc* cosmological theory, but is rather a

more general theory, developed for mainly study glasses, fluctuation-dissipation theorems, and ergodic-non-ergodic transition [10–13]. It has been adapted by Bartelmann *et al.* to large-scale structures in the Universe [14–17] and describes structures as an ensemble of correlated classical particles obeying Hamiltonian dynamics. All the dynamics information is enclosed in a generating functional, structurally similar to the one of statistical quantum field theories.

In this framework, the random momentum and density autocorrelation matrices, that rise in the generating functional of KFT, and that are sensible to the initial conditions of particle distribution in the phase space, can be seen as Euclidean Random Matrices (ERMs). They are a special class of random matrices [18, 19], which can be studied with the formalism of Random Matrix Theory (RMT), i.e. a theoretical-physical tool that has been developed thanks to the work of Wishart [20] and Wigner [21, 22]. The goal of RMT is to calculate the eigenvalues distribution and other statistical properties of ensembles of matrices whose elements are random variables. In a sense, RMT is a modified version of the more known central limit theorem for random variables, applied to random matrices. This thesis aims to use the methods developed to compute the eigenvalues distribution for ERMs and apply them to the covariance matrices of KFT. In particular the high and the low-density limit will also be considered.

This thesis begins in Chapter 1 with a review of the cosmological background used to build the KFT formalism. In particular the Friedmann equations, the Λ -CDM model, the Zel'dovich approximation, and the process that leads to structure formation are introduced. The latter is studied also from an RMT perspective.

In Chapter 2 an introduction to standard and multivariate probability theory is presented, the goal is to recap the information and settle the formalism largely used by RMT and KFT. A special focus is given to the Gaussian random fields.

With Chapter 3 KFT is finally introduced, with a focus on the derivation of the elements that will be treated subsequently: the generating functional and the covariance matrices.

Chapter 4 is devoted to giving a self-consistent introduction to the Random Matrix Theory formalism. It is not meant as a full review of the topic, but rather specifically targeted at the scope of this thesis work.

In Chapter 5, after reviewing the possible matter density contrast power spectra, the eigenvalues distribution of $C_{\delta\delta}$ and \mathbf{C}_{pp} covariance matrices of KFT in the high and low-density limit are studied. It follows a short analysis of the single eigenvalue distribution and its applicability in the generating functional of KFT.

Chapter 6 concludes this thesis with a discussion of the obtained results, possible improvements of the framework, and future outcomes of the research.

Chapter 1

Cosmology fundamentals

Before proceeding in more detail with the mathematical formalism of Kinetic Field Theory and presenting the essential topics of this thesis, we need to give a review of the conceptual framework where KFT takes place. In particular, will be described the main equations that rule the Universe as we know it, focusing on the theory that leads to linear cosmic structure formation, such as the Zel'dovich approximation. It explains the rise of large-scale structures, which will be then studied under a Random Matrix Theory approach.

1.1 Basis of cosmology

Cosmology is described in the framework of General Relativity, since, out of the four interactions (strong, weak, electromagnetic and gravitational), the only relevant force capable of making interact far away and neutral matter is gravity. General Relativity describes space-time as a four-dimensional manifold whose metric tensor $g_{\mu\nu}$ is a dynamical field that defines the line element

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu. \quad (1.1)$$

The components of the metric, are related to the energy-matter distribution of space-time by Einstein's field equations

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} + \Lambda g_{\mu\nu}. \quad (1.2)$$

In the right-hand-side of the last equation, rests the energy content, expressed by the energy-momentum tensor

$$T_{\mu\nu} = (\rho c^2 + p) u_\mu u_\nu + p g_{\mu\nu}, \quad (1.3)$$

for fluids in thermodynamical equilibrium, and a *cosmological constant* Λ term added to match the accelerating expansion of the Universe.

Cosmology is based on two further principles, the so-called *cosmological principle*. Is indeed assumed that the Universe is spatially homogeneous and isotropic over sufficiently large scales. This means that any observation made in the Universe should not depend on the direction in

which the observation is performed nor on the location from where the observation is done. Applying the cosmological principle to the line element (1.1), it turns into the Friedmann-Lemaître-Robertson-Walker (FLRW) metric

$$ds^2 = c^2 dt^2 - a^2(t) [dr^2 + f_K^2(r)(d\theta^2 + \sin^2 \theta d\varphi^2)], \quad (1.4)$$

in spherical coordinates, and where the function $f_K(r)$, which describes the curvature of space, is of the form

$$f_K(r) = \begin{cases} \frac{\sin \sqrt{K}r}{\sqrt{K}} & K > 0, \text{ spherical;} \\ r & K = 0, \text{ flat;} \\ \frac{\sinh \sqrt{-K}r}{\sqrt{-K}} & K < 0, \text{ hyperbolic.} \end{cases} \quad (1.5)$$

namely the only three possibilities that allow the curvature to be the same all across the space. In Eq. (1.4), $a(t)$ is the dimensionless *scale factor* and changes the distances over the time. It is set at unity nowadays $a_0 = a(t_0) = 1$.

The differential equations describing the evolution of the only free parameter $a(t)$ are known as Friedmann equations, they can be derived from Einstein's field equation, imposing the FLRW metric and considering the energy-momentum tensor (1.3)

$$\begin{aligned} \left(\frac{\dot{a}}{a}\right)^2 &= \frac{8\pi G}{3}\rho - K\frac{c^2}{3a^2} + \Lambda\frac{c^2}{3}, \\ \frac{\ddot{a}}{a} &= -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right) + \Lambda\frac{c^2}{3}. \end{aligned} \quad (1.6)$$

These equations describe the relation between the scale factor $a(t)$, the mass density ρ , the pressure p and the cosmological constant Λ . The two Friedmann equations can be combined into a continuity equation

$$\dot{\rho} + 3H\left(\rho + \frac{p}{c^2}\right) = 0 \quad (1.7)$$

that describes the evolution of the energy density. The pressure and the density are related by the following *equation of state*

$$p = \omega\rho c^2, \quad \omega = \begin{cases} 0 & \text{non-relativistic matter;} \\ \frac{1}{3} & \text{radiation;} \\ -1 & \Lambda. \end{cases} \quad (1.8)$$

Solving the continuity equation for a constant ω , it can be obtained

$$\rho_X(t) = \rho_{X,0}a^{-3(1+\omega)}, \quad (1.9)$$

for any component X , that could be, in our case matter or radiation, since the curvature vanishes $K = 0$ because of measurements [23]. Taking into account matter, radiation and the cosmological constant, the first of the Eqs. (1.6) becomes

$$\frac{H^2}{H_0^2} := E^2(a) = \Omega_{r,0}a^{-4} + \Omega_{m,0}a^{-3} + \Omega_{\Lambda,0}, \quad (1.10)$$

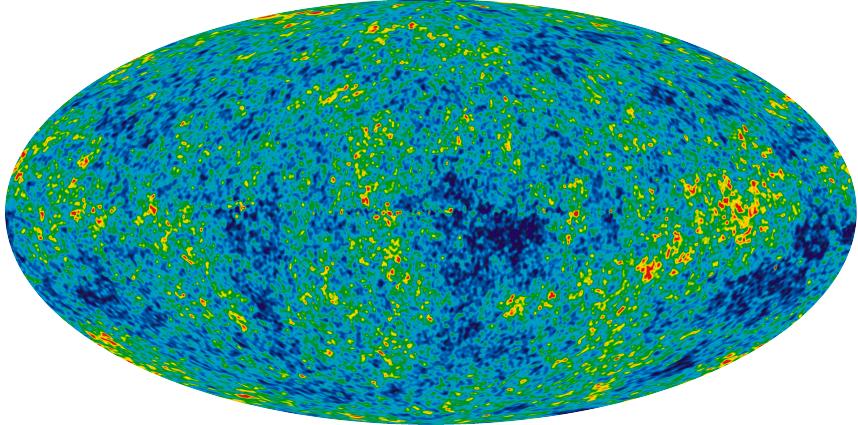


Figure 1.1: Wilkinson Microwave Anisotropy Probe heat map of temperature fluctuations in the cosmic microwave background

where we defined $H(t) = \frac{\dot{a}(t)}{a(t)}$ and $H_0 = H(a = 1)$ the Hubble constant. The density parameters are defined through

$$\Omega_X = \frac{8\pi G}{3H^2}\rho_X, \quad \Omega_\Lambda = \frac{\Lambda c^2}{3H^2}, \quad (1.11)$$

and in a flat Universe obey to

$$\sum_{x=m,r,\Lambda} \Omega_x = 1. \quad (1.12)$$

1.1.1 Λ Cold Dark Matter model

The role of the standard model of cosmology is played by the Λ -Cold Dark Matter (Λ -CDM) model. It is a mathematical model of that accepts the Big Bang Theory to explain the beginning of the Universe and with three major components: ordinary matter, a cosmological constant Λ and the Cold Dark Matter (CDM). The latter is the postulated form of matter that can better fit observational data. The adjective *cold* refers to the fact that is not relativistic, the *dark* term means that it does not interact electromagnetically. There are different candidates to be CDM particles, but this problem is not investigated any further in this thesis. One of the main validation for this theory, is the observation of the early Universe temperature fluctuations of the Cosmic Microwave Background 1.1, compared to the nowadays energy fluctuations. While the first are of the order $\approx 10^{-5}$, the presence of large-scale structures in the late Universe, push the energy density fluctuations to order $\approx 10^6$, as measured by the 2MASS survey [1]. This evolution from small fluctuations to large must have carried out by a mediator that does not interact electromagnetically [8], hence the CDM. The understanding of how the small temperature fluctuations transformed into the Universe we can observe nowadays is carried out by the field of *Cosmic Structure Formation*.

1.1.2 Linear structure formation

Focusing on the formation of linear structures, we will derive the equations that will lead to the Zel'dovich approximation and to KFT. Before that, let us define the *comoving frame* that will be

useful in the next steps. Indeed, since we observe (and assume) an expanding Universe, it is important to define comoving coordinates, for objects that move with the cosmic flow

$$\vec{q}_{\text{phys}}(t) = a(t)\vec{q}_{\text{co}}. \quad (1.13)$$

where the scale factor $a(t)$ is the same appearing in the FLRW metric (1.4). In this way is now possible to describe particles dynamics in an expanding Universe, only considering the peculiar motion with respect to the background. The physics velocity changes accordingly

$$\vec{v}_{\text{phys}} = \frac{d\vec{q}_{\text{phys}}(t)}{dt} = \frac{da}{dt}\vec{q}_{\text{co}} + a\frac{d\vec{q}_{\text{co}}}{dt} = H\vec{q}_{\text{phys}}(t) + a\vec{u}, \quad (1.14)$$

where \vec{u} is called *peculiar velocity* and describes the inherent motion of objects relative to the cosmological frame. All the differential operator changes as well

$$\begin{aligned} \vec{\nabla}_{\text{phys}} &= \frac{\vec{\nabla}_{\text{co}}}{a} \\ \partial_t f|_{\text{phys}} &= \partial_t f|_{\text{co}} - H\vec{q}_{\text{co}} \cdot \vec{\nabla}_{\text{co}} f. \end{aligned} \quad (1.15)$$

At this point is possible to consider the equations that describe the behaviour of the matter distribution, as a function of the density contrast

$$\delta(\vec{x}, t) = \frac{\rho(\vec{x}, t) - \bar{\rho}(t)}{\bar{\rho}(t)}, \quad \rho(\vec{x}, t) = \bar{\rho}(t)(1 + \delta(\vec{x}, t)), \quad (1.16)$$

namely the continuity, the Poisson and Euler's equations, and write them in comoving coordinates

$$\begin{aligned} \partial_t \delta + \vec{\nabla} \cdot [(1 + \delta)\vec{u}] &= 0, \\ \partial_t \vec{u} + 2H\vec{u} + (\vec{u} \cdot \vec{\nabla})\vec{u} &= -\frac{\vec{\nabla}\Phi}{a^2}, \\ \nabla^2\Phi &= 4\pi G\bar{\rho}a^2\delta. \end{aligned} \quad (1.17)$$

Linearising the system, namely dropping all terms of higher than first order, one obtained the linear evolution equation

$$\ddot{\delta} + 2H\dot{\delta} = 4\pi G\bar{\rho}\delta. \quad (1.18)$$

It can be further modified, replacing the time derivative with the derivative with respect to a

$$\begin{aligned} \partial_t &= aH\partial_a, \\ \partial_t^2 &= (aH)^2 \left[\partial_a^2 + \left(\frac{1}{a} + \frac{E'}{E} \right) \partial_a \right], \end{aligned} \quad (1.19)$$

resulting in

$$\delta'' + \left(\frac{3}{a} + \frac{E'}{E} \right) \delta' = \frac{3}{2} \frac{\Omega_m}{a^2} \delta, \quad (1.20)$$

where we used the fact that $4\pi G\bar{\rho} = \frac{3}{2} \frac{H_0^2 \Omega_{m,0}}{a^3}$. Since the equation (1.20) is an ordinary differential equation, homogeneous in δ , it is allowed to pick a solution of the form

$$\delta(a) = \delta_0 D_+(a) \quad (1.21)$$

where δ_0 is some amplitude at the reference scale factor a_0 . At early time in the Universe, $a \ll 1$, and the approximation $\Omega_m \simeq 1$ can be done, the structure growth equation looks like

$$D_+'' + \frac{3}{2a} D_+' = \frac{3}{2a^2} D_+. \quad (1.22)$$

since, during the matter dominated epoch:

$$\frac{E'}{E} = \frac{1}{2} \frac{(E^2)'}{E^2} = -\frac{3}{2a} \frac{\Omega_{m,0}}{\Omega_{m,0} + a^3 \Omega_{\Lambda,0}} \simeq -\frac{3}{2a}. \quad (1.23)$$

Finally, we can use the ansatz $D_+ = a^n$ for Eq. (1.22) to find that there are two solutions $n_{\pm} = (1, -3/2)$, meaning that the density contrast grows linearly with the scale factor a at early times

$$D_+(a) = a. \quad (1.24)$$

1.2 Zel'dovich approximation

The goal of this section is to describe the dynamics of a test particle of mass m in an expanding space-time. The Hamiltonian phase-space trajectories read [24]

$$x(t) = G(t, 0)x^{(i)} - \int_0^t dt' G(t, t') \begin{pmatrix} 0 \\ m \vec{\nabla} \varphi \end{pmatrix}, \quad (1.25)$$

where the potential φ satisfies the Poisson equation

$$\nabla^2 \varphi = \frac{3a}{2m^2}. \quad (1.26)$$

We needed to define the Green's function

$$G(t, t') = \begin{pmatrix} \mathbb{1}_3 & g_H(t, t') \mathbb{1}_3 \\ 0_3 & \mathbb{1}_3 \end{pmatrix}, \quad (1.27)$$

and the Hamiltonian propagator

$$g_H(t, t') = \int_{t'}^t \frac{d\bar{t}}{m(\bar{t})}. \quad (1.28)$$

The main idea is to solve the spatial component of the integral equation (1.25) iteratively, using some inertial trajectories.

$$\vec{q}(t) = \vec{q}^{(i)} + g_H(t, 0)\vec{p}^{(i)} - \int_0^t dt' g_H(t, t')m\vec{\nabla}\phi, \quad (1.29)$$

with now the free trajectory

$$\vec{q}(t) = \vec{q}^{(i)} + b(t)\vec{u}^{(i)}, \quad (1.30)$$

depending on some unknown function $b(t)$. Then the Jacobian of the previous "transformation" from initial and final positions reads

$$\mathcal{J} = \frac{\partial \vec{q}(t)}{\partial \vec{q}^{(i)}} = \mathbb{1}_3 + b(t)D^2\psi, \quad (1.31)$$

where the velocity $\vec{u}^{(i)} = \vec{\nabla}\psi$ can be considered as the momentum if the mass is set to unity at the initial time and hence can be expressed as the gradient of some velocity potential ψ . In particular, the hessian matrix of the velocity field $D^2\psi$, known as Zel'dovich or deformation tensor, will be used later. Replacing the Hamiltonian propagator as

$$g_H(t, t') \rightarrow t - t', \quad (1.32)$$

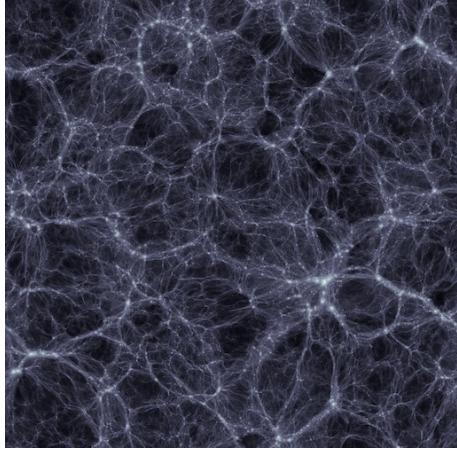


Figure 1.2: The image from [25] depicts the cosmic web as simulated in an N-body simulation of the Λ -CDM cosmology at $z = 0$. It spans a width of $250 h^{-1}$ Mpc and illustrates the logarithmic overdensity $\log(1 + \delta)$. According to Zel'dovich's first-order perturbation theory, the cosmic matter distribution is expected to give rise to structures such as pancakes, filaments, halos, and voids.

and the no-more unknown function $b(t) = t$, we obtain the new trajectories in real space

$$\tilde{q}(t) = \vec{q}^{(i)} + t\vec{u}^{(i)} - \int_0^t dt' g_H(t, t') m \vec{\nabla} \varphi, \quad (1.33)$$

where now the potential $\varphi = \phi + 4\pi G \bar{\rho} a^2 (t+1) \psi$ obeys to

$$\nabla^2 \varphi = A_\varphi \delta^{(\text{nl})} = 4\pi G \bar{\rho} a^2 \left(\delta - (t+1)\delta^{(i)} \right). \quad (1.34)$$

Having introduced the free reference motion (1.30), we obtain a good approximation to solve Eq. (1.29) iteratively.

1.3 Zel'dovich tensor and its eigenvalues

This section gives a motivation of why one could be in principle interested in the eigenvalues distribution of a random matrix. In particular, we analyze the spectrum distribution of the Zel'dovich tensor $D^2\psi$ in Eq. (1.31) and see what important consequences it brings to the understanding of the large-scale structures in the Universe. This has been first considered in Ref. [6, 26] in the context of the cosmic flow in Zel'dovich approximation.

If we consider the Zel'dovich approximation

$$\tilde{q}(t) = \vec{q}^{(i)} + t\vec{\nabla}\psi(\vec{q}), \quad (1.35)$$

using mass conservation $\rho(t)d^3q = \rho^{(i)}d^3q^{(i)}$, we have

$$\rho(t) = \frac{\rho^{(i)}}{\det \mathcal{J}} = \frac{\rho^{(i)}}{[1 + t\lambda_1(\vec{q})][1 + t\lambda_2(\vec{q})][1 + t\lambda_3(\vec{q})]}, \quad (1.36)$$

with the Jacobian $\mathcal{J} = \mathbb{1}_3 + tD^2\psi = \mathbb{1}_3 + t \text{ diag}(\lambda_1, \lambda_2, \lambda_3)$, and we diagonalized the deformation tensor.

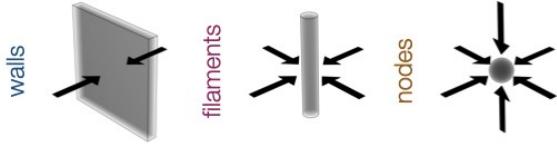


Figure 1.3: Compression of cosmic matter along different directions.

For Gaussian initial conditions, it is possible to work out the eigenvalues probability distribution [26]

$$\varrho(\lambda_1, \lambda_2, \lambda_3) = \mathcal{N} e^{-\frac{3}{2\sigma^2}(2\sum_{i=1}^3 \lambda_i^2 - \sum_{i < j} \lambda_i \lambda_j)} \prod_{j < k} |\lambda_j - \lambda_k|. \quad (1.37)$$

This statement is entirely proven in Appendix D. Here we would like to argue that knowing the eigenvalues distribution of a random field, which could be for example the random velocity potential field, could give many useful information. Indeed, the Zel'dovich tensor is the Hessian of the velocity potential random field ψ , hence a random matrix, and its eigenvalues λ_i represent the converging (diverging) rate of matter flow in each of the three spatial directions. It is clear from Eq. (1.37) that the probability distribution is proportional to the differences between all the possible eigenvalues pairs, therefore equal eigenvalues appear with zero probability. This mathematical property physically means that the Zel'dovich approximation prevents the cosmic flow from converging (diverging) with the same rates in the three spatial dimensions. Namely, exactly spherical and cylindrical collapse is forbidden, leading to non-spherical and non-cylindrical over-(under-)densities. In this way, the cosmic matter is compressed first along one direction forming two-dimensional structures, called pancakes or walls. Then, it is compressed in another direction, leading to the formation of filaments. This formation of one- and two- dimensional structures is a necessary consequence of our initial assumption of a Gaussian random field, see Fig. 1.3. This is a great result as we can observe the presence of these large-scale structures in the Universe [27].

There are two more things in Eq. (1.37) that are worth noticing. Firstly, the exponential factor kills any configuration of eigenvalues that are too far from the origin in absolute value. Secondly,

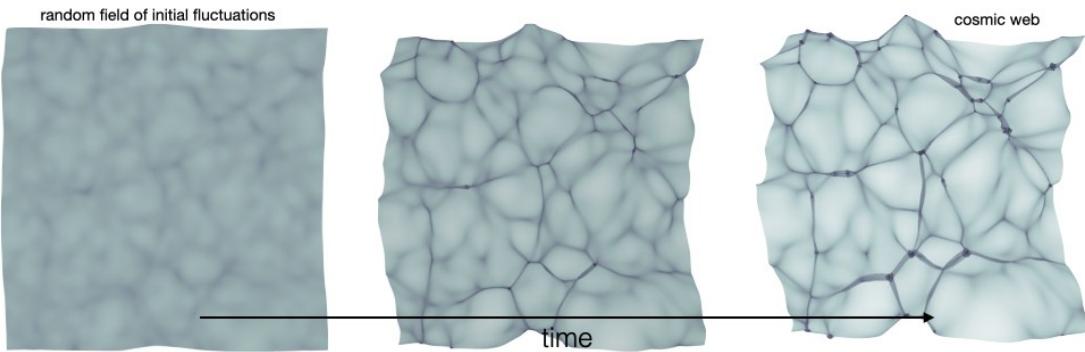


Figure 1.4: Evolution of cosmic webs as large-scale structures starting from a random field.

the presence of the element

$$\prod_{i < j} |\lambda_i - \lambda_j|, \quad (1.38)$$

also called Vandermonde determinant, is a very general property owned by a large variety of random matrices that will be studied in more detail in chapter 4. Its presence is the consequence of variable transformation since it is the determinant of the transformation matrix, see also Appendix E. In general, it is related to the "repulsive" behavior of the eigenvalues of a random matrix, which will be explained in way more detail. Furthermore, it is the signature of the strong dependence among the eigenvalues. Each eigenvalue is related to each other and thus is not possible to factorize the spectrum as $\varrho(\lambda_1, \lambda_2, \lambda_3) = \varrho_1(\lambda_1)\varrho_2(\lambda_2)\varrho_3(\lambda_3)$.

Chapter 2

Probability and random fields

In many fields of physics, the right mathematical formalism is required to properly describe the framework where the physical laws take place. That's why it is necessary to understand first the main results of probability theory and random fields, largely used by Kinetic Field Theory (KFT) and Random Matrix Theory (RMT). This introduction aims to provide the background of probability theory and the statistics of random fields that will be used in the next chapters.

2.1 Probability theory

Given a continuous random variable X and its probability distribution $\varrho_X(x)$ we can define the moments μ_k of the distribution. They are the expectation value of the power of the random variable

$$\mu_k = \langle x^k \rangle = \int_{\mathbb{R}} \varrho_X(x) x^k dx. \quad (2.1)$$

Related to this concept is the characteristic function of the probability distribution, or generator function of the moments

$$\phi(t) := \langle e^{-itx} \rangle = \int_x \varrho_X(x) e^{-itx}. \quad (2.2)$$

namely the Fourier transform of the probability distribution and therefore the probability distribution is the inverse Fourier transform of the characteristic function. The characteristic function assumes then a significant role in determining the k -th moment μ_k of the probability distribution, indeed

$$\mu_k = \langle x^k \rangle = (i\partial_t)^k \phi(t) \Big|_{t=0}. \quad (2.3)$$

We can directly see it from the power series expansion of the characteristic function

$$\phi(t) = \left\langle \sum_{k=0}^{+\infty} \frac{(-it)^k}{k!} x^k \right\rangle = \sum_{k=0}^{+\infty} \frac{(-it)^k}{k!} \mu_k. \quad (2.4)$$

Analogously, we can define a cumulants generating function, that is given by the logarithm of the characteristic function

$$c_k = (i\partial_t)^k \ln \phi(t) \Big|_{t=0}. \quad (2.5)$$

Also in this case is useful to see the power series expansion of the generating function

$$\ln \phi(t) = \sum_{k=1}^{+\infty} \frac{t^k}{k!} c_k. \quad (2.6)$$

The moments and the cumulants are related by the following relation

$$\mu_k = \sum_{|p_n|} k! \prod_n \frac{1}{p_n!(n!)^{p_n} c_n^{p_n}}, \quad (2.7)$$

with the constraint that the sum is limited over $\sum np_n = k$. For example

$$\begin{aligned} \mu_1 &= c_1, \\ \mu_2 &= c_2 + c_1^2, \\ \mu_3 &= c_3 + 3c_2c_1 + c_1^3, \end{aligned} \quad (2.8)$$

or alternatively

$$\begin{aligned} c_1 &= \mu_1, \\ c_2 &= \mu_2 - \mu_1^2, \\ c_3 &= \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3, \end{aligned} \quad (2.9)$$

and so on and so forth.

2.1.1 Multiple random variables

In the case where we have a set of random variables that define a d -dimensional space $S_X = X_1, \dots, X_N$. For example, a N particle state could be described by a set of N positions $\vec{x} = (x_1, \dots, x_N)$. The probability that the system will be found in a precise configuration is given by the joint probability distribution function (jpdf)

$$\varrho(\vec{x}) = \varrho(x_1, \dots, x_N), \quad (2.10)$$

with the normalization condition

$$\int_x \varrho(\vec{x}) = 1. \quad (2.11)$$

Then the quantity

$$\int_a^b dx_1 \dots \int_c^d dx_N \varrho(x_1, \dots, x_N), \quad (2.12)$$

gives the probability that the first variable X_1 assumes a value in the interval (a, b) and the last variable X_N is at the same moment, in the interval (c, d) . The marginal probability distribution function (pdf) is the integral over the $N - 1$ variables of the jpdf

$$\varrho(x) = \int dx_2 \dots \int dx_N \varrho(x, x_2, \dots, x_N) \quad (2.13)$$

When the jpdf is factorized, i.e. is the product of N density functions,

$$\varrho(x_1, \dots, x_N) = \varrho_1(x_1) \dots \varrho_N(x_N), \quad (2.14)$$

the variables are said to be independent, otherwise, they are dependent. When, in addition, we also have $\varrho_1(x) = \dots = \varrho_N(x)$, the random variables are called independent and identically distributed (i.i.d.). In any case, $\varrho(x_1) = \int dx_2 \dots dx_N \varrho(x_1, \dots, x_N)$ is the marginal pdf of X_1 when considered independently of the others $X_i, i = 2, \dots, N$.

Variable transformation

Let us say that we are now interested in going from the pdf of a set of random variable X_i to a pdf of another set Y_i that are functions of X_i through $x_i = x_i(y)$, namely we would like to obtain $\varrho_Y(y)$ as a function of $\varrho_X(x)$. As usual, we can use the Jacobian of the transformation and exploit it for our purpose

$$\varrho_X(x_1, \dots, x_N) dx_1 \cdots dx_N = \varrho_Y(x_1(y), \dots, x_N(y)) \left| \det \left(\frac{dx_i}{dy_i} \right) \right| dy_1 \cdots dy_N. \quad (2.15)$$

Conditional probability

Let X, Y be two continuous random variables, let us define a conditional probability density function of Y given the occurrence of the value x of X

$$\varrho_{Y|X}(y|x) = \frac{\varrho_{X,Y}(x,y)}{\varrho_X(x)}. \quad (2.16)$$

where $\varrho_{X,Y}(x,y)$ is the jpdf of X and Y and $\varrho_X(x)$ is the marginal density of X . Generalizing the concept to N variables we have

$$\varrho(x_1, \dots, x_m | x_{m+1}, \dots, x_N) = \frac{\varrho(x_1, \dots, x_N)}{\varrho(x_{m+1}, \dots, x_N)}, \quad (2.17)$$

that allows us to find the conditional probability for a fixed set of events with a set of known values $\varrho(x_1, \dots, x_m | x_{m+1}, \dots, x_N)$. The numerator is again the jpdf.

Bayes Theorem

Given two continuous random variables X and Y we are now interested in describing the probability of a variable, based on our prior knowledge of conditions related to the variable. Using the definition of conditional density (2.16)

$$\begin{aligned} \varrho_{X|Y=y}(x) &= \frac{\varrho_{X,Y}(x,y)}{\varrho_Y(y)}, \\ \varrho_{Y|X=x}(y) &= \frac{\varrho_{X,Y}(x,y)}{\varrho_X(x)}, \end{aligned} \quad (2.18)$$

therefore

$$\varrho_{X|Y=y}(x) = \frac{\varrho_{Y|X=x}(y)\varrho_X(x)}{\varrho_Y(y)}. \quad (2.19)$$

2.1.2 Central limit theorem

Let us come to one of the most important theorems in the world of probability theory, the Central Limit Theorem (CLT). It gives a convergence rule for probability distributions in probability spaces. If we consider N statistically independent and identically distributed (i.i.d.) random variables X_1, \dots, X_N , with

$$\begin{aligned} \langle x_1 \rangle &= \dots = \langle x_N \rangle = \mu, \\ \langle x_1^2 \rangle - \langle x_1 \rangle^2 &= \dots \langle x_N^2 \rangle - \langle x_N \rangle^2 = \sigma^2 < \infty. \end{aligned} \quad (2.20)$$

The CLT wants to investigate the pdf of

$$A_N = \frac{1}{N} \sum_{i=1}^N x_i. \quad (2.21)$$

It states that quantities such as the mean, the variance and the limit distribution of A_N are

$$\begin{aligned} \langle A_N \rangle &= \mu, \\ \sigma_{A_N}^2 &= \frac{\sigma^2}{N}, \\ \varrho_N(A_N) &\rightarrow \frac{1}{\sqrt{2\pi\sigma_{A_N}^2}} e^{-\frac{(A_N - \mu)^2}{2\sigma_{A_N}^2}}. \end{aligned} \quad (2.22)$$

So it does not matter what is the distribution of the individual random variables, as long as the first two moments exist, the average variable is asymptotically Gaussian distributed. The result is given by the ubiquitous appearance of the Gaussian distribution in statistical phenomena. The spirit of this theorem is resumed by the Random Matrix Theory, which gives different limit eigenvalue distributions for different classes of random matrices.

2.2 Gaussian random fields

A d -dimensional Gaussian random field $\psi(\mathbf{x}) : \mathbb{R}^d \rightarrow \mathbb{R}$ is defined via a d -dimensional multivariate normal distribution:

$$\varrho(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^d \det \mathbf{C}}} \exp\left[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^\top \mathbf{C}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right], \quad (2.23)$$

with

$$\boldsymbol{\mu} = \langle \mathbf{x} \rangle, \quad \mathbf{C} = \langle \mathbf{x} \otimes \mathbf{x} \rangle; \quad (2.24)$$

where $\mathbf{x} \in \mathbb{R}^d$ and $\mathbf{C} \in \mathbb{R}^{d \times d}$. They are the expectation value of \mathbf{x} and the covariance matrix of the distribution. One can always transform $\mathbf{x} \rightarrow \mathbf{x} - \boldsymbol{\mu}$, thus assuming a centered distribution, $\boldsymbol{\mu} = 0$, without losing generality. The covariance matrix is a $d \times d$ symmetric and positive semi-definite matrix, ensuring its eigenvalues to be real and positive, therefore $\det \mathbf{C} > 0$ and \mathbf{C}^{-1} exist. Using the characteristic function (2.2) of \mathbf{x} :

$$\phi(\mathbf{t}) = \langle e^{-i\mathbf{t}\mathbf{x}} \rangle = e^{-\frac{1}{2}\mathbf{t}^\top \mathbf{C}\mathbf{t}}, \quad (2.25)$$

as proven in Eq. B.7. Using the tools introduced in Eq. (2.3) and extending them to d -dimensions, one can take derivatives of Eq. (2.25) evaluated at $\mathbf{t} = 0$ obtaining the moments of the distribution:

$$\mu_0 = 1, \quad \mu_1 = \mathbf{0}, \quad \mu_2 = \mathbf{C}, \quad (2.26)$$

meaning that the distribution is correctly normalized, centered, and with \mathbf{C} as a covariance matrix. Note that μ_k is now a k -way symmetric tensor: μ_0 is a scalar, μ_1 is a vector, μ_2 is a matrix and so on.

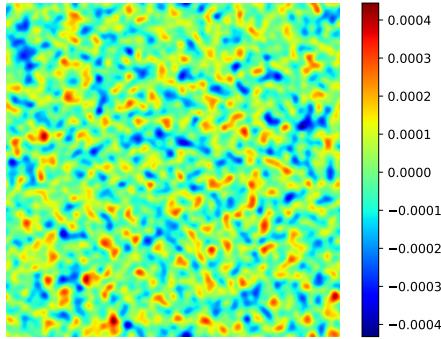


Figure 2.1: Gaussian random field, it can be taken as a model describing the CMB.

Correlation function and power spectrum

Particularly useful is the definition of correlation function ξ of a (random) field, as it quantitatively tells us what is the degree of correlation between the variables in our field. For example, in the case of uncorrelated variables, the correlation function vanishes.

Let us specialize from now on, on a 3-dimensional random field ψ , and let us define its correlation function as its covariance as

$$\xi(\vec{x}, \vec{y}) = \langle \psi(\vec{x})\psi(\vec{y}) \rangle. \quad (2.27)$$

Since our cosmological considerations imply spatial homogeneity and isotropy, the correlation function depends only on the absolute value of the separation $r = |\vec{r}| = |\vec{x} - \vec{y}|$:

$$\xi(r) = \langle \psi(\vec{x})\psi(\vec{x} + \vec{r}) \rangle. \quad (2.28)$$

In particular, consider the density contrast

$$\delta(\vec{x}) = \frac{\rho(\vec{x}) - \bar{\rho}}{\bar{\rho}}, \quad (2.29)$$

depending on the matter density $\rho(\vec{x})$ in the Universe and its average $\bar{\rho}$. More important than the autocorrelation function its Fourier transform, namely the power spectrum

$$P(k) = \int_r \xi(r) e^{-i\vec{k} \cdot \vec{r}}. \quad (2.30)$$

So, while the autocorrelation function gives the probability to find the same quantity at distance r , the power spectrum decomposes this probability into characteristic lengths, $k \approx \frac{2\pi}{L}$, and its amplitude describes the degree to which each characteristic length contributes to the total probability. It quantifies the size of the structures. With this tool we can easily describe, for example, scale invariant structures, such as fractals. Indeed, a power-law power spectrum $P(k) \sim k^{-\alpha}$ gives rise to structures at all scales, giving birth to the so-called fractal structures, as shown in Fig 2.2. It will be largely used throughout this thesis, as it is one of the main quantities used in large-scale structures physics. For example, $P_\delta(k)$ is the matter power spectrum and describes the distribution of the density contrast δ . Taking into account the power spectrum and the autocorre-

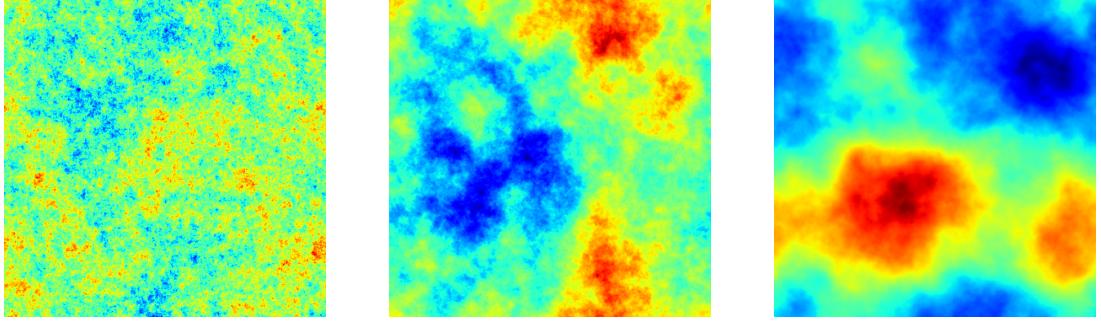


Figure 2.2: Gaussian random field with power spectrum $P(k) = k^{-\alpha}$, with $\alpha = 2, 3, 4$. Note the presence of the fractal structures.

lation function of the random field ψ , and considering spatial isotropy, one can integrate out the angle between \vec{k} and \vec{r}

$$\begin{aligned} P_\psi(k) &= \int_r \xi_\psi(r) e^{-i\vec{k}\cdot\vec{r}} = 2\pi \int_0^{+\infty} dr r^2 \xi_\psi(r) \int_{-1}^{+1} d(\cos\theta) e^{-ikr \cos\theta} \\ &= 4\pi \int_0^{+\infty} dr r^2 \xi_\psi(r) j_0(kr), \end{aligned} \quad (2.31)$$

where we introduced the spherical Bessel function j_0 of the first kind and zeroth order, see Appendix A

$$j_0(x) = \frac{\sin(x)}{x}. \quad (2.32)$$

We are left with the two equations

$$\xi_\psi(k) = \int_0^{+\infty} \frac{dk}{2\pi^2} k^2 P_\psi(k) j_0(kr), \quad P_\psi(r) = 4\pi \int_0^{+\infty} dr r^2 \xi_\psi(r) j_0(kr). \quad (2.33)$$

At this point, it is useful to consider the variance of the Fourier transform of the random field:

$$\begin{aligned} \langle \tilde{\psi}(\vec{k}) \tilde{\psi}(\vec{k}') \rangle &= \left\langle \int_x \psi(\vec{x}) e^{-i\vec{k}\cdot\vec{x}} \int_y \psi(\vec{y}) e^{-i\vec{k}'\cdot\vec{y}} \right\rangle \\ &= \int_x \int_y e^{-i(\vec{k}\cdot\vec{x} + \vec{k}'\cdot\vec{y})} \langle \psi(\vec{x}) \psi(\vec{y}) \rangle \\ &= \int_x \int_r e^{-i(\vec{k}\cdot\vec{x} + \vec{k}'\cdot(\vec{x} - \vec{r}))} \xi_\psi(r) \\ &= (2\pi)^3 \delta_D(\vec{k} + \vec{k}') \int_r \xi_\psi(r) e^{-i\vec{k}\cdot\vec{r}} \\ &= (2\pi)^3 \delta_D(\vec{k} + \vec{k}') P_\psi(k), \end{aligned} \quad (2.34)$$

where we used the Dirac delta because the field ψ is homogeneous. It indicates that Fourier modes with different wave vectors are statistically independent.

Considering the second of Eqs. (1.16) and computing its Fourier transform

$$\tilde{\rho}(\vec{k}) = \bar{\rho} \int_q (1 + \delta(\vec{q})) e^{-i\vec{k}\cdot\vec{q}} = \bar{\rho} \left[(2\pi)^3 \delta_D(k) + \tilde{\delta}(\vec{k}) \right], \quad (2.35)$$

we can then express the two-point function of the density contrast in Fourier space

$$\langle \tilde{\rho}(\vec{k}) \tilde{\rho}(\vec{k}') \rangle = \bar{\rho} \left[(2\pi)^6 \delta_D(\vec{k}) \delta_D(\vec{k}') + \langle \tilde{\delta}(\vec{k}) \tilde{\delta}(\vec{k}') \rangle \right], \quad (2.36)$$

where the last term, analogously to Eq. (2.34) reads

$$\langle \tilde{\delta}(\vec{k})\tilde{\delta}(\vec{k}') \rangle = (2\pi)^3 \delta_D(\vec{k} + \vec{k}') P_\delta(\vec{k}), \quad (2.37)$$

allowing us to write Eq. (2.36) as

$$\frac{1}{\bar{\rho}^2} \langle \tilde{\rho}(\vec{k})\tilde{\rho}(\vec{k}') \rangle = (2\pi)^3 \delta_D(\vec{k} + \vec{k}') \left[(2\pi)^3 \delta_D(\vec{k}) + P_\delta(\vec{k}) \right]. \quad (2.38)$$

Chapter 3

Kinetic Field Theory

The framework where this thesis takes place is Kinetic Field Theory, a statistical field theory describing classical ensembles of particles. It has been developed by Bartelmann *et al.* [14, 15] and aims to describe large-scale structures as ensembles of correlated classical particles out of equilibrium. Here we want to review the main steps that lead to the generating functional, and the correlation matrices, which will then be studied with the RMT formalism.

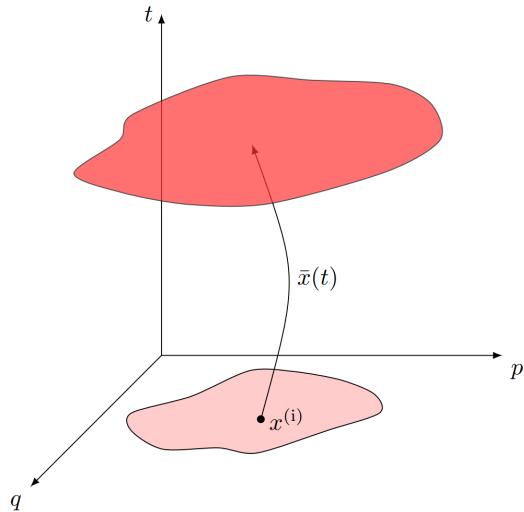


Figure 3.1: Phase space is shown schematically as a function of time. Particle trajectories originating at an initial position transport the initial probability forward in time.

Let us introduce the following notation for convenience

$$\int_x := \int_{\mathbb{R}^N} d^N \vec{x}; \quad \int_k := \int_{\mathbb{R}^N} \frac{d^N \vec{k}}{(2\pi)^N}. \quad (3.1)$$

Furthermore, let us define the Fourier transform of a function $f(\vec{x})$ and its inverse as

$$\mathcal{F}[f](\vec{k}) = \tilde{f}(\vec{k}) = \int_x f(\vec{x}) e^{-i\vec{k} \cdot \vec{x}}; \quad \mathcal{F}^{-1}[\tilde{f}](\vec{x}) = f(\vec{x}) = \int_k \tilde{f}(\vec{k}) e^{i\vec{k} \cdot \vec{x}}. \quad (3.2)$$

Furthermore, let us introduce the tensor notation used throughout this thesis, which is also the

standard notation of KFT. Since KFT considers ensembles of N particles in phase-space, it is convenient to bundle all the phase-space trajectories into a trajectory tensor

$$\mathbf{x} = x_i \otimes e_i = \begin{pmatrix} \vec{q}_i \\ \vec{p}_i \end{pmatrix} \otimes e_i, \quad (3.3)$$

where e_i is the unit vector in N dimensions. It becomes natural to define the scalar product between two tensors \mathbf{x} and \mathbf{y}

$$\mathbf{x}\mathbf{y} = (\mathbf{x}, \mathbf{y}) = (x_i \cdot y_j) (e_i \cdot e_j) = (x_i \cdot y_j) \delta_{ij} = x_i \cdot y_i. \quad (3.4)$$

3.1 Initial phase-space distribution

The initial phase-space distribution of the particles is Gaussian distribution with covariance matrix \mathbf{C}_{pp} ,

$$P(\mathbf{x}^{(i)}) = \frac{V^{-N}}{\sqrt{(2\pi)^{3N} \det \mathbf{C}_{pp}}} \exp\left(-\frac{1}{2} \mathbf{p}^{(i)\top} \mathbf{C}_{pp}^{-1} \mathbf{p}^{(i)}\right). \quad (3.5)$$

Let us see how this important result can be reached from statistical considerations. We need to start from the set of the initial positions and momenta of the N particles

$$\mathbf{x}^{(i)} = x_i^{(i)} \otimes e_i = \begin{pmatrix} \vec{q}_i^{(i)} \\ \vec{p}_i^{(i)} \end{pmatrix} \otimes e_i, \quad (3.6)$$

a remark about the notation needs to be made, namely, we will use the convention that the superscript (i) indicates an initial quantity, while the subscript i is an index, usually labeling the number of the particle, thus going from 1 to N , if it is not the case it will be clarified.

To find the probability distribution for $\mathbf{x}^{(i)}$, we need to start from the initial probability distribution for the velocity potential $\psi^{(i)}$ that we assume to be Gaussian. Therefore also the probability distribution of the data tensor of the density contrast $\delta_i^{(i)}$ and the momentum $\vec{\nabla}\psi_i^{(i)}$ is Gaussian, since they depend on the gradient and the Laplacian of the velocity field $\psi_i^{(i)}$, indeed $\delta_i^{(i)} = -\nabla^2\psi_i^{(i)}$.

$$\mathbf{d}^{(i)} = d_i^{(i)} \otimes e_i = \begin{pmatrix} \delta_i^{(i)} \\ \vec{\nabla}\psi_i^{(i)} \end{pmatrix} \otimes e_i, \quad (3.7)$$

we can write the joint probability distribution for the $4N$ values $\mathbf{d}^{(i)}$. It is indeed a multi-variate Gaussian

$$P(\mathbf{d}^{(i)}) = \frac{1}{\sqrt{(2\pi)^{4N} \det \mathbf{C}}} \exp\left(-\frac{1}{2} \mathbf{d}^{(i)\top} \mathbf{C}^{-1} \mathbf{d}^{(i)}\right), \quad (3.8)$$

with the covariance matrix

$$\mathbf{C} = \langle \mathbf{d}^{(i)} \otimes \mathbf{d}^{(i)} \rangle. \quad (3.9)$$

At this point, we just need to transform the probability distribution from $\mathbf{d}^{(i)}$ to $\mathbf{x}^{(i)}$. Let us get started with the conditional probability to find a particle in $\vec{q}_i^{(i)}$ where the density contrast is $\delta_i^{(i)}$ and with momentum $\vec{p}_i^{(i)}$ where the velocity potential has the gradient $\vec{\nabla}\psi_i^{(i)}$.

$$\begin{aligned} P(\vec{q}_i^{(i)} | \delta_i^{(i)}) &= \frac{1 + \delta_i^{(i)}}{V}, \\ P(\vec{p}_i^{(i)} | \vec{\nabla}\psi_i^{(i)}) &= \delta_D(\vec{p}_i^{(i)} - \vec{\nabla}\psi_i^{(i)}). \end{aligned} \quad (3.10)$$

since $\vec{p}_i^{(i)} = \vec{\nabla}\psi_i^{(i)}$. We can now transform $P(\mathbf{d}^{(i)})$ in $P(\mathbf{x}^{(i)})$ employing Bayes theorem

$$\begin{aligned} P(\mathbf{x}^{(i)}) &= \prod_{i=1}^N \int d\delta_i^{(i)} \int d\vec{\nabla}\psi_i^{(i)} P(\vec{q}_i^{(i)}|\delta_i^{(i)}) P(\vec{p}_i^{(i)}|\vec{\nabla}\psi_i^{(i)}) P(\mathbf{d}^{(i)}) \\ &= V^{-N} \prod_{i=1}^N \int d\delta_i^{(i)} (1 + \delta_i^{(i)}) P(\mathbf{d}^{(i)}), \end{aligned} \quad (3.11)$$

where we replaced $\vec{\nabla}\psi_i^{(i)}$ with $\vec{p}_i^{(i)}$ in $\mathbf{d}^{(i)}$ because of the Dirac delta. Furthermore we introduce the characteristic function $\phi(\mathbf{t}) = e^{-1/2\mathbf{t}^\top C\mathbf{t}}$, as a function of \mathbf{t}

$$\mathbf{t} = \begin{pmatrix} \mathbf{t}_\delta \\ \mathbf{t}_p \end{pmatrix} = \begin{pmatrix} t_{\delta_i} \\ t_{p_i} \end{pmatrix} \otimes e_i, \quad (3.12)$$

allowing us to write

$$\begin{aligned} P(\mathbf{x}^{(i)}) &= V^{-N} \prod_{i=1}^N \int d\delta_i^{(i)} (1 + \delta_i^{(i)}) \int_{\mathbf{t}} \phi(\mathbf{t}) e^{i\mathbf{t} \cdot \mathbf{d}^{(i)}} \\ &= V^{-N} \prod_{i=1}^N (1 + i\partial_{s_i}) \int d\delta_i^{(i)} \int_{\mathbf{t}} \phi(\mathbf{t}) e^{i(\mathbf{t}_\delta - \mathbf{s}) \cdot \delta^{(i)} + i\mathbf{t}_p \cdot \mathbf{p}^{(i)}} \Big|_{\mathbf{s}=0}, \end{aligned} \quad (3.13)$$

having introduced a source \mathbf{s} for the density contrast. The integral of the density contrast

$$\int d\delta_i^{(i)} e^{i(\mathbf{t}_\delta - \mathbf{s}) \cdot \delta^{(i)}} = \delta_D(\mathbf{t}_\delta - \mathbf{s}), \quad (3.14)$$

yielding

$$\begin{aligned} P(\mathbf{x}^{(i)}) &= V^{-N} \hat{D} \int_{\mathbf{t}_\delta} \int_{\mathbf{t}_p} \delta_D(\mathbf{t}_\delta - \mathbf{s}) \exp \left(-\frac{1}{2} \begin{pmatrix} \mathbf{t}_\delta \\ \mathbf{t}_p \end{pmatrix}^\top \begin{pmatrix} \mathbf{C}_{\delta\delta} & \mathbf{C}_{\delta p} \\ \mathbf{C}_{p\delta} & \mathbf{C}_{pp} \end{pmatrix} \begin{pmatrix} \mathbf{t}_\delta \\ \mathbf{t}_p \end{pmatrix} \right) e^{i\mathbf{t}_p \cdot \mathbf{p}^{(i)}} \\ &= V^{-N} \hat{D} \int_{\mathbf{t}_p} \phi(\mathbf{s}, \mathbf{t}_p) e^{i\mathbf{t}_p \cdot \mathbf{p}^{(i)}}, \end{aligned} \quad (3.15)$$

where we defined the operator

$$\hat{D} := \prod_{i=1}^N (1 + i\partial_{s_i}) \Big|_{\mathbf{s}=0} \quad (3.16)$$

and the characteristic function

$$\phi(\mathbf{s}, \mathbf{t}_p) := \exp \left(-\frac{1}{2} (\mathbf{s}^\top \mathbf{C}_{\delta\delta} \mathbf{s} + \mathbf{t}_p^\top \mathbf{C}_{pp} \mathbf{t}_p + 2\mathbf{s}^\top \mathbf{C}_{\delta p} \mathbf{t}_p) \right). \quad (3.17)$$

As last step, we just use Gaussian integral \mathbf{B} , in order to obtain

$$P(\mathbf{x}^{(i)}) = \frac{V^{-N}}{\sqrt{(2\pi)^{3N} \det \mathbf{C}_{pp}}} \hat{D} \exp \left(-\frac{1}{2} (\mathbf{p}^{(i)\top} \mathbf{C}_{pp}^{-1} \mathbf{p}^{(i)} + 2i\mathbf{p}^{(i)\top} \mathbf{A}\mathbf{s} + \mathbf{s}^\top \mathbf{B}\mathbf{s}) \right). \quad (3.18)$$

with the matrices

$$\mathbf{A} = \mathbf{C}_{pp}^{-1} \mathbf{C}_{\delta p}^\top, \quad \mathbf{B} = \mathbf{C}_{\delta\delta} - \mathbf{C}_{\delta p} \mathbf{C}_{pp}^{-1} \mathbf{C}_{\delta p}^\top. \quad (3.19)$$

At this point, it is possible to approximate $\hat{D} = 1$ and $\mathbf{s} = 0$ allowing us to consider only the momentum-momentum correlations, which is a valid assumption, especially at late cosmic times. In this way, we obtain Eq. (3.5).

3.2 Covariance matrices

We are now interested in considering the two points correlation matrices $C_{\delta_i \delta_j}$, $C_{\delta_i p_j}$, $C_{p_i p_j}$ because they compose the complete covariance matrix for the complete data set $\mathbf{d} = (\overset{\delta_i}{\vec{p}_i}) \otimes e_i$.

$$\mathbf{C} = \langle \mathbf{d} \otimes \mathbf{d} \rangle = C_{ij} \otimes E_{ij} = (C_{\delta_i \delta_j} + 2C_{\delta_i p_j} + C_{p_i p_j}) \otimes E_{ij}, \quad (3.20)$$

with the matrix $E_{ij} = e_i \otimes e_j$. The density-density correlation $C_{\delta_i \delta_j}$ is then a 1×1 matrix, the density-momentum correlation $C_{\delta_i p_j}$ is a 3×1 matrix and the momentum-momentum correlation $C_{p_i p_j}$ is a 3×3 matrix. Our focus will be put on the momentum-momentum correlation, since, as we will see, one can write the generating functional depending on the momentum-momentum correlation matrix. The i, j -th component of C is a 4×4 matrix that reads:

$$\mathbf{C} = \begin{pmatrix} C_{11} & \dots & C_{1N} \\ \vdots & \ddots & \vdots \\ C_{N1} & \dots & C_{NN} \end{pmatrix}, \quad C_{ij} = \begin{pmatrix} C_{\delta_i \delta_j} & & & \\ & \cdots & \cdots & \\ & & C_{\delta_i p_j} & \\ & & & C_{p_i p_j} \end{pmatrix}. \quad (3.21)$$

Let us work out the analytical expression of the two-point covariance matrices. Recall that

$$\delta_i = -\nabla^2 \psi_i, \quad \vec{p}_i = \vec{\nabla} \psi_i, \quad \tilde{\delta}(k) = k^2 \tilde{\psi}(k), \quad P_\delta(k) = k^4 P_\psi(k), \quad (3.22)$$

have a look at Appendix F for further details. Therefore

$$\begin{aligned} C_{\delta_i \delta_j}(r) &= \langle \delta_i \otimes \delta_j \rangle = \langle \nabla^2 \psi_i \cdot \nabla^2 \psi_j \rangle \\ &= \int_k \int_{k'} k^2 k'^2 \langle \tilde{\psi}_i(k) \tilde{\psi}_j(k') \rangle e^{i(\vec{k} \cdot \vec{x}_i + \vec{k}' \cdot \vec{x}_j)} \\ &= \int_k \int_{k'} k^2 k'^2 (2\pi)^3 \delta_D(k+k') P_\psi(k) e^{i(\vec{k} \cdot \vec{x}_i + \vec{k}' \cdot \vec{x}_j)} \\ &= \int_k k^4 P_\psi(k) e^{i(\vec{k} \cdot \vec{x}_i - \vec{k} \cdot \vec{x}_j)} \\ &= \int_k P_\delta(k) e^{i\vec{k} \cdot \vec{r}} \\ &= \frac{1}{2\pi^2} \int_0^{+\infty} dk k^2 P_\delta(k) j_0(kr) \\ &= \xi(r). \end{aligned} \quad (3.23)$$

Where we made use of Eqs. (2.30), (2.34) and of the first of Eqs. (2.33). With the same logic, the density-momentum covariance matrix is

$$\begin{aligned} C_{\delta_i p_j}(r) &= \langle \delta_i \otimes \vec{p}_j \rangle = \langle \nabla^2 \psi_i \cdot \vec{\nabla} \psi_j \rangle \\ &= \int_k \int_{k'} k^2 \left(i \vec{k}' \right) \langle \tilde{\psi}_i(k) \tilde{\psi}_j(k') \rangle e^{i(\vec{k} \cdot \vec{x}_i + \vec{k}' \cdot \vec{x}_j)} \\ &= -i \int_k dk k^2 \vec{k}' P_\psi(k) e^{i\vec{k} \cdot \vec{r}} \\ &= \frac{-i\hat{r}}{(2\pi)^2} \int_0^{+\infty} dk k^5 P_\psi(k) \int_{-1}^{+1} d\mu \mu e^{ikr\mu} \\ &= \frac{\hat{r}}{2\pi^2} \int_0^{+\infty} dk k P_\delta(k) j_1(kr) \\ &= \hat{r}\zeta(r). \end{aligned} \quad (3.24)$$

where $C_{\delta_i p_j}(r)$ is rotationally symmetric about the \vec{r} axis and we introduced the spherical Bessel function of the first kind and of the first order see App. A

$$j_1(x) = \frac{\sin(x)}{x^2} - \frac{\cos(x)}{x}. \quad (3.25)$$

For what concern the momentum-momentum correlation one needs to carry out

$$\begin{aligned} C_{p_i p_j}(r) &= \langle \vec{p}_i \otimes \vec{p}_j \rangle = \langle \vec{\nabla} \psi_i \otimes \vec{\nabla} \psi_j \rangle \\ &= \int_k \int_{k'} \left(i \vec{k} \right) \otimes \left(i \vec{k}' \right) \langle \tilde{\psi}_i(k) \tilde{\psi}_j(k') \rangle e^{i(\vec{k} \cdot \vec{x}_i + \vec{k}' \cdot \vec{x}_j)} \\ &= \int_k \left(\vec{k} \otimes \vec{k} \right) P_\psi(k) e^{i \vec{k} \cdot \vec{r}}. \end{aligned} \quad (3.26)$$

At this point, we need to introduce the two projectors

$$\pi_{\parallel} = \hat{r} \otimes \hat{r}, \quad \pi_{\perp} = \mathbb{1}_3 - \pi_{\parallel}, \quad (3.27)$$

parallel and perpendicular to the direction \vec{r} . In this way we can expand $C_{p_i p_j}(r)$

$$C_{p_i p_j}(r) = \langle \pi_{\parallel}, C_{p_i p_j} \rangle \pi_{\parallel} + \frac{1}{2} \langle \pi_{\perp}, C_{p_i p_j} \rangle \pi_{\perp}, \quad (3.28)$$

where $\langle A, B \rangle = \text{Tr}(AB)$ indicates the scalar product between symmetric matrices. The two scalar products read

$$\begin{aligned} \langle \pi_{\parallel}, C_{p_i p_j} \rangle &= \int_k k^2 \mu^2 P_\psi(k) e^{i \vec{k} \cdot \vec{r}} = \frac{1}{(2\pi)^2} \int_0^{+\infty} dk k^4 P_\psi(k) \int_{-1}^{+1} d\mu \mu^2 e^{ikr\mu}, \\ \langle \pi_{\perp}, C_{p_i p_j} \rangle &= \int_k k^2 (1 - \mu^2) P_\psi(k) e^{i \vec{k} \cdot \vec{r}} = \frac{1}{(2\pi)^2} \int_0^{+\infty} dk k^4 P_\psi(k) \int_{-1}^{+1} d\mu (1 - \mu^2) e^{ikr\mu}, \end{aligned} \quad (3.29)$$

where $\mu = \cos \theta$ is the cosine of the angle between \vec{k} and \vec{r} . Using the properties of the spherical Bessel functions in appendix A

$$\begin{aligned} \int_k k^2 P_\psi(k) e^{i \vec{k} \cdot \vec{r}} &= \frac{1}{2\pi^2} \int_0^{+\infty} dk P_\delta(k) \left[\frac{3j_1(kr)}{kr} - j_2(kr) \right], \\ \int_k k^2 \mu^2 P_\psi(k) e^{i \vec{k} \cdot \vec{r}} &= \frac{1}{2\pi^2} \int_0^{+\infty} dk P_\delta(k) \left[\frac{j_1(kr)}{kr} - j_2(kr) \right], \end{aligned} \quad (3.30)$$

that, plugged back in Eq. (3.29), yield

$$\begin{aligned} \langle \pi_{\parallel}, C_{p_i p_j} \rangle &= \frac{1}{2\pi^2} \int_0^{+\infty} dk P_\delta(k) \left[\frac{j_1(kr)}{kr} - j_2(kr) \right] = -[a_1(r) + a_2(r)], \\ \langle \pi_{\perp}, C_{p_i p_j} \rangle &= \frac{1}{\pi^2} \int_0^{+\infty} dk P_\delta(k) \frac{j_1(kr)}{kr} = -2a_1(r), \end{aligned} \quad (3.31)$$

having defined the two correlation functions

$$\begin{aligned} a_1(r) &:= -\frac{1}{2\pi^2} \int_0^{+\infty} dk P_\delta(k) \frac{j_1(kr)}{kr}, \\ a_2(r) &:= \frac{1}{2\pi^2} \int_0^{+\infty} dk P_\delta(k) j_2(kr), \end{aligned} \quad (3.32)$$

allowing us to write Eq. (3.28) as

$$\begin{aligned} C_{p_i p_j}(r) &= -[a_1(r) + a_2(r)] \pi_{\parallel} - a_1(r) \pi_{\perp} \\ &= -a_1(r) \mathbb{1}_3 - a_2(r) \pi_{\parallel}. \end{aligned} \quad (3.33)$$

At this point we should calculate the limit for small r of the correlation functions $a_1(r)$ and $a_2(r)$, namely

$$\lim_{r \rightarrow 0} a_1(r) = \lim_{x \rightarrow 0} -\frac{1}{2\pi^2} \int_0^{+\infty} dk P_\delta(k) \frac{j_1(kr)}{kr} = -\frac{1}{3} \frac{1}{2\pi^2} \int_0^{+\infty} dk P_\delta(k) = -\frac{\sigma_1^2}{3}, \quad (3.34)$$

$$\lim_{r \rightarrow 0} a_2(r) = 0.$$

since

$$\lim_{x \rightarrow 0} \frac{j_1(x)}{x} = \frac{1}{3}, \quad \lim_{x \rightarrow 0} j_2(x) = 0, \quad (3.35)$$

for completeness we also have

$$\lim_{r \rightarrow 0} \xi(r) = \lim_{r \rightarrow 0} \frac{1}{2\pi^2} \int_0^{+\infty} dk k^2 P_\delta(k) j_0(kr) = \sigma_2^2, \quad (3.36)$$

$$\lim_{r \rightarrow 0} \zeta(r) = 0.$$

and we defined the moments of the initial density-fluctuation power spectrum

$$\sigma_n^2 = \frac{1}{2\pi^2} \int_0^{+\infty} dk k^{2n-2} P_\delta(k). \quad (3.37)$$

Using recursion relations between Bessel spherical function (A.8), we can derive useful relations between our statistical quantities $\xi(r), \zeta(r), a_1(r)$ and $a_2(r)$. Namely, since

$$\begin{aligned} \zeta'(r) &= \xi(r) - \frac{2}{r} \zeta(r), \\ a'_2(r) &= \zeta(r) - \frac{3}{r} a_2(r) \end{aligned} \quad (3.38)$$

we obtain

$$\begin{aligned} \zeta(r) &= a'_2(r) + \frac{3}{r} a_2(r), \\ \xi(r) &= \zeta'(r) + \frac{2}{r} \zeta(r) = a''_2(r) + \frac{5}{r} a'_2(r) + \frac{3}{r^2} a_2(r). \end{aligned} \quad (3.39)$$

Momentum-momentum covariance matrix

Let us now consider the \mathbf{C}_{pp} matrix defined as the following

$$\mathbf{C}_{pp} = C_{p_i p_j} \otimes E_{ij}, \quad (3.40)$$

$$\mathbf{C}_{pp} = \begin{pmatrix} C_{p_1 p_1} & \dots & C_{p_1 p_N} \\ \vdots & \ddots & \vdots \\ C_{p_N p_1} & \dots & C_{p_N p_N} \end{pmatrix}, \quad C_{p_i p_i} = -\frac{\sigma_1^2}{3} \mathbb{1}_3, \quad C_{p_i p_j} = -a_1(r) \mathbb{1}_3 - a_2(r) \pi_{\parallel}. \quad (3.41)$$

Let us visualize, for a better understanding, the \mathbf{C}_{pp} matrix for two particles,

$$\mathbf{C}_{pp} = \begin{pmatrix} C_{p_1 p_2} & C_{p_1 p_2} \\ C_{p_2 p_1} & C_{p_2 p_2} \end{pmatrix}, \quad C_{p_1 p_1} = C_{p_2 p_2} = \begin{pmatrix} \frac{\sigma_1^2}{3} & 0 & 0 \\ 0 & \frac{\sigma_1^2}{3} & 0 \\ 0 & 0 & \frac{\sigma_1^2}{3} \end{pmatrix}, \quad (3.42)$$

$$C_{p_1 p_2} = C_{p_2 p_1} = \begin{pmatrix} a_1(r) + a_2(r) r_x^2 & a_2(r) r_x r_y & a_2(r) r_x r_z \\ a_2(r) r_x r_y & a_1(r) + a_2(r) r_y^2 & a_2(r) r_y r_z \\ a_2(r) r_x r_z & a_2(r) r_y r_z & a_1(r) + a_2(r) r_z^2 \end{pmatrix}, \quad (3.43)$$

where $\vec{r} = r(r_x, r_y, r_z)^\top$ with module r and components r_j , $j = x, y, z$. This matrix and its generalization for N particles will be central in our discussion in the next chapter with the random matrices approach. More specifically we will study its eigenvalues distribution and connect it to the generating functional.

3.3 Generating functional

We would need to start with the probability $P(\mathbf{x}, t)$ to find a particle ensemble at time t at the phase space position \mathbf{x} . To do that, we must make use of the conditional probability machinery, namely

$$P(\mathbf{x}(t)) = \int d\mathbf{x}^{(i)} P(\mathbf{x}(t)|\mathbf{x}^{(i)}) P(\mathbf{x}^{(i)}), \quad (3.44)$$

as our probability is given by the multiplication between the probability to find the initial ensemble at an initial position $\mathbf{x}^{(i)}$ and the conditional probability for the ensemble to be in the position $\mathbf{x}(t)$ at time t , given their initial position. Of course, we need to consider all the possible initial positions, that's why we integrate over $\mathbf{x}^{(i)}$. Since we are dealing with a classical theory, and therefore with a classical ensemble, the trajectories $\bar{\mathbf{x}}$ are described by deterministic equations of motion and so we can express the transition probability by a Dirac delta that vanishes for all those trajectories that are different from the ones that solve the equations of motion starting at $\mathbf{x}^{(i)}$, namely

$$P(\mathbf{x}(t)|\mathbf{x}^{(i)}) = \delta_D[\mathbf{x} - \bar{\mathbf{x}}]. \quad (3.45)$$

The generating function is thus the path integral over all the trajectories \mathbf{x} , augmented by a source field \mathbf{J} ,

$$\begin{aligned} Z[\mathbf{J}] &= \int \mathcal{D}\mathbf{x} P(\mathbf{x}(t)) e^{i(\mathbf{J}, \mathbf{x})} \\ &= \int \mathcal{D}\mathbf{x} \int d\mathbf{x}^{(i)} P(\mathbf{x}(t)|\mathbf{x}^{(i)}) P(\mathbf{x}^{(i)}) e^{i(\mathbf{J}, \mathbf{x})} \\ &= \int \mathcal{D}\mathbf{x} \int d\mathbf{x}^{(i)} \delta_D[\mathbf{x} - \bar{\mathbf{x}}] P(\mathbf{x}^{(i)}) e^{i(\mathbf{J}, \mathbf{x})} \\ &= \int d\mathbf{x}^{(i)} P(\mathbf{x}^{(i)}) e^{i(\mathbf{J}, \bar{\mathbf{x}})} \\ &:= \int d\Gamma e^{i(\mathbf{J}, \bar{\mathbf{x}})}, \end{aligned} \quad (3.46)$$

where we defined the phase space measure

$$d\Gamma = d\mathbf{x}^{(i)} P(\mathbf{x}^{(i)}). \quad (3.47)$$

The source \mathbf{J} has been chosen such that

$$\langle \mathbf{x}(t_1) \rangle = -i \frac{\delta}{\delta \mathbf{J}(t_1)} Z[\mathbf{J}]|_{\mathbf{J}=0}. \quad (3.48)$$

3.3.1 Density operator

The particle number density of our particle ensemble at time t is a sum over delta distributions

$$\rho(\vec{q}, t) = \sum_{j=1}^N \delta_D(\vec{q} - \vec{q}_j(t)), \quad (3.49)$$

that can be transformed in Fourier space

$$\tilde{\rho}(\vec{k}, t) = \sum_{j=1}^N e^{-i\vec{k}\cdot\vec{q}_j(t)}. \quad (3.50)$$

Making use of the functional derivative in Eq. (3.48) for each position component \vec{J}_{q_j} of the source field \mathbf{J}

$$\vec{q}_j(t) \rightarrow -i \frac{\delta}{\delta \vec{J}_{q_j}(t)}, \quad (3.51)$$

we obtain the N -particle density operator

$$\hat{\rho}(\vec{k}, t) = \sum_{j=1}^N \exp \left(-\vec{k} \cdot \frac{\delta}{\delta \vec{J}_{q_j}(t)} \right) = \sum_{j=1}^N \hat{\rho}_j(\vec{k}, t), \quad (3.52)$$

namely the sum of the one-particle density operator $\hat{\rho}_j(\vec{k}, t)$ for each particles j . The functional derivative in the exponent represents a translation, yielding

$$\begin{aligned} \hat{\rho}_j(\vec{k}_j, t_j) \mathbf{J} &= \exp \left[-\vec{k}_j \cdot \frac{\delta \mathbf{J}}{\delta \vec{J}_{q_j}(t_j)} \right] \\ &= \exp \left[-\vec{k}_j \cdot \frac{\delta}{\delta \vec{J}_{q_j}(t_j)} \left(\frac{\vec{J}_{q_k}(t)}{\vec{J}_{p_k}} \right) \otimes e_k \right] \\ &= \exp \left[-\vec{k}_j \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \delta_D(t - t_j) \delta_{jk} \otimes e_k \right], \end{aligned} \quad (3.53)$$

where the right-hand-side of the last equation is exactly the shift generated by the functional derivative $\mathbf{J} \rightarrow \mathbf{J} + \mathbf{L}$. The application of n -one particle density operators to the generating functional (3.46) acts in the same way as the evaluation of the generating functional at the *shift tensor* \mathbf{L}

$$\hat{\rho}_1(\vec{k}_1, t_1) \cdots \hat{\rho}_n(\vec{k}_n, t_n) Z[\mathbf{J}] \Big|_{\mathbf{J}=0} = Z[\mathbf{L}], \quad (3.54)$$

where the shift tensor reads

$$\mathbf{L} = - \sum_{j=1}^n \vec{k}_j \cdot \frac{\delta \mathbf{J}(t)}{\delta \vec{J}_{q_j}(t_j)} = - \sum_{j=1}^n \vec{k}_j \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \delta_D(t - t_j) \otimes e_j. \quad (3.55)$$

In the end, applying n density operators to the generating functional

$$\begin{aligned} \langle \hat{\rho}_1(\vec{k}_1, t_1) \cdots \hat{\rho}_n(\vec{k}_n, t_n) \rangle &= \sum_{j_1, \dots, j_n=1}^N \hat{\rho}_{j_1}(\vec{k}_1, t_1) \cdots \hat{\rho}_{j_n}(\vec{k}_n, t_n) Z[\mathbf{J}] \Big|_{\mathbf{J}=0} \\ &= \prod_{r=0}^{n-1} (N-r) Z[\mathbf{L}] \\ &\approx N^n Z[\mathbf{L}]. \end{aligned} \quad (3.56)$$

The first equality in the second line is valid due to the indistinguishability of particles within the ensemble. This implies that each tuple of particles (j_1, \dots, j_n) among the $N(N-1)\cdots(N-n+1)$ possible combinations must yield identical statistical outcomes. The ultimate approximation holds when N is significantly greater than n , as is often the case. In cosmology, it is reasonable to assume that any sufficiently large volume contains an exceedingly vast number of particles. So from now on we will consider the last approximation as an equality.

3.3.2 Free generating functional

The form of the free generating functional we are interested in is

$$\begin{aligned} Z_0[\mathbf{L}] &= \frac{V^{-N}}{\sqrt{(2\pi)^{3N}}} \int \frac{d\mathbf{q} d\mathbf{p}}{\sqrt{\det \mathbf{C}_{pp}}} \exp \left(-\frac{1}{2} \mathbf{p}^\top \mathbf{C}_{pp}^{-1} \mathbf{p} + i\mathbf{L}_q \cdot \mathbf{q} + i\mathbf{L}_p \cdot \mathbf{p} \right) \\ &= V^{-N} \int d\mathbf{q} \exp \left(-\frac{1}{2} \mathbf{L}_p^\top \mathbf{C}_{pp} \mathbf{L}_p + i\mathbf{L}_q \cdot \mathbf{q} \right), \end{aligned} \quad (3.57)$$

with the two shift tensors $\mathbf{L}_q, \mathbf{L}_p$ defined as

$$\mathbf{L}_q = - \sum_{i=1}^l \vec{k}_i \otimes e_i, \quad \mathbf{L}_p = - \sum_{i=1}^l \vec{k}_i t_i \otimes e_i. \quad (3.58)$$

In order to arrive at this expression we need to start considering the free generating functional (3.46) with the initial phase-space distribution (3.5)

$$\begin{aligned} Z_0[\mathbf{J}] &= \int d\Gamma e^{i(\mathbf{J}, \mathbf{x}_0)} \\ &= \int d\mathbf{x}^{(i)} P(\mathbf{x}^{(i)}) e^{i(\mathbf{J}, \mathbf{x}_0)} \\ &= \frac{V^{-N}}{\sqrt{(2\pi)^{3N}}} \int d\mathbf{q} \int d\mathbf{p} \frac{1}{\sqrt{\det \mathbf{C}_{pp}}} \exp \left(-\frac{1}{2} \mathbf{p}^\top \mathbf{C}_{pp}^{-1} \mathbf{p} \right) e^{i(\mathbf{J}, \mathbf{x}_0)}. \end{aligned} \quad (3.59)$$

The trajectories of our particles ensemble are given by

$$\begin{aligned} \bar{\mathbf{x}} &= \mathbf{x}_0 + \mathbf{x}_I \\ &= \mathbf{G}(t, 0)\mathbf{x}^{(i)} + \int_0^t dt' \mathbf{G}(t, t') \mathbf{F}(t') \\ &= \mathbf{G}(t, 0) x_i^{(i)} \otimes e_i + \int_0^t dt' \mathbf{G}(t, t') \begin{pmatrix} 0 \\ \vec{f}_i \end{pmatrix} \otimes e_i \\ &= \begin{pmatrix} g_{qq} \mathbb{1}_3 & g_{qp} \mathbb{1}_3 \\ g_{pq} \mathbb{1}_3 & g_{pp} \mathbb{1}_3 \end{pmatrix} \otimes \mathbb{1}_N \begin{pmatrix} \vec{q}_i^{(i)} \\ \vec{p}_i^{(i)} \end{pmatrix} \otimes e_i + \mathbf{x}_I. \end{aligned} \quad (3.60)$$

where we will neglect the term \mathbf{x}_I since we are dealing with non-interacting particles. Redefining our source field $\mathbf{J} \rightarrow \mathbf{J} + \mathbf{L}$, allowing us to write $\mathbf{L} = -\vec{k}_1 \frac{\delta \mathbf{J}(t)}{\delta J_{q_i}(t_1)}$ and creating the shift

$$\mathbf{L}(t) = - \sum_{i=1}^l \vec{k}_i \begin{pmatrix} 1 \\ 0 \end{pmatrix} \delta(t - t_i) \otimes e_i. \quad (3.61)$$

In the case of the power spectrum, we would only need $l = 2$

$$\begin{aligned} \mathbf{L}(t) &= - \begin{pmatrix} \vec{k}_1 \\ 0 \end{pmatrix} \delta(t - t_1) \otimes e_1 - \begin{pmatrix} \vec{k}_2 \\ 0 \end{pmatrix} \delta(t - t_2) \otimes e_2 \\ &= - \begin{pmatrix} 1 \\ 0 \end{pmatrix} (\delta(t - t_1) \vec{k}_1 \otimes e_1 + \delta(t - t_2) \vec{k}_2 \otimes e_2), \end{aligned} \quad (3.62)$$

that becomes

$$\mathbf{L}(t) = - \begin{pmatrix} 1 \\ 0 \end{pmatrix} \delta_D(t - t_1) \vec{k}_1 \otimes (e_1 - e_2), \quad (3.63)$$

for synchronous power spectra. At this point, we carry out the calculation for the element $(\mathbf{L}, \mathbf{x}_0)$, in the case $l = 1$

$$\begin{aligned}
(\mathbf{L}, \mathbf{x}_0) &:= \int_0^{t'} dt \mathbf{L}(t) \mathbf{x}_0(t) \\
&= - \int_0^{t'} dt \vec{k}_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} \delta_D(t - t_1) \otimes e_1 \cdot \mathbf{G}(t_1, 0) \begin{pmatrix} \vec{q}_1^{(i)} \\ \vec{p}_1^{(i)} \end{pmatrix} \otimes e_i \\
&= - \begin{pmatrix} \vec{k}_1 \\ \vec{0} \end{pmatrix} \mathbf{G}(t_1, 0) \begin{pmatrix} \vec{q}_1^{(i)} \\ \vec{p}_1^{(i)} \end{pmatrix} \\
&= - \begin{pmatrix} \vec{k}_1 \\ \vec{0} \end{pmatrix} \begin{pmatrix} \mathbb{1}_3 & g_H(t_1, 0) \\ 0 & \dot{g}_H(t_1, 0) \end{pmatrix} \begin{pmatrix} \vec{q}_1^{(i)} \\ \vec{p}_1^{(i)} \end{pmatrix} \\
&= - \begin{pmatrix} \vec{k}_1 \\ \vec{0} \end{pmatrix} \begin{pmatrix} \vec{q}_1^{(i)} + g_H(t_1, 0) \vec{p}_1^{(i)} \\ \dot{g}_H(t_1, 0) \vec{p}_1^{(i)} \end{pmatrix} \\
&= - \vec{k}_1 \left(\vec{q}_1^{(i)} + g_H(t_1, 0) \vec{p}_1^{(i)} \right) \\
&:= L_q \vec{q}_1^{(i)} + L_p \vec{p}_1^{(i)},
\end{aligned} \tag{3.64}$$

immediately generalizable for the case $l = 2$, namely for the power spectrum:

$$\begin{aligned}
(\mathbf{L}, \mathbf{x}_0) &:= \int_0^{t'} dt \mathbf{L}(t) \mathbf{x}_0(t) \\
&= - \int_0^{t'} dt \left(\delta(t - t_1) \vec{k}_1 \otimes e_1 + \delta(t - t_2) \vec{k}_2 \otimes e_2 \right) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot \mathbf{G}(t, 0) \begin{pmatrix} \vec{q}_i^{(i)} \\ \vec{p}_i^{(i)} \end{pmatrix} \otimes e_i \\
&= - \begin{pmatrix} \vec{q}_1^{(i)} + g_H(t_1, 0) \vec{p}_1^{(i)} \\ \dot{g}_H(t_1, 0) \vec{p}_1^{(i)} \end{pmatrix} \begin{pmatrix} \vec{k}_1 \\ \vec{0} \end{pmatrix} - \begin{pmatrix} \vec{q}_2^{(i)} + g_H(t_2, 0) \vec{p}_2^{(i)} \\ \dot{g}_H(t_2, 0) \vec{p}_2^{(i)} \end{pmatrix} \begin{pmatrix} \vec{k}_2 \\ \vec{0} \end{pmatrix} \\
&= - \left(\vec{q}_1^{(i)} + g_H(t_1, 0) \vec{p}_1^{(i)} \right) \vec{k}_1 - \left(\vec{q}_2^{(i)} + g_H(t_2, 0) \vec{p}_2^{(i)} \right) \vec{k}_2 \\
&:= \mathbf{L}_q \mathbf{q}^{(i)} + \mathbf{L}_p \mathbf{p}^{(i)},
\end{aligned} \tag{3.65}$$

with the two shift tensor

$$\mathbf{L}_q = - \sum_{i=1}^l g_{qq_i} \vec{k}_i \otimes e_i, \quad \mathbf{L}_p = - \sum_{i=1}^l g_{qp_i} \vec{k}_i \otimes e_i. \tag{3.66}$$

At the end of the day, what we remain with is:

$$\begin{aligned}
Z_0[\mathbf{L}] &= \frac{V^{-N}}{\sqrt{(2\pi)^{3N}}} \int d\mathbf{q} \int d\mathbf{p} \frac{1}{\sqrt{\det \mathbf{C}_{pp}}} \exp \left(-\frac{1}{2} \mathbf{p}^\top \mathbf{C}_{pp}^{-1} \mathbf{p} \right) e^{-i \vec{k}_1 \cdot (\vec{q}^{(i)} + g_H(t_1, 0) \vec{p}^{(i)})} \\
&= \frac{V^{-N}}{\sqrt{(2\pi)^{3N}}} \int d\mathbf{q} \frac{e^{-i \mathbf{L}_q \mathbf{q}^{(i)}}}{\sqrt{\det \mathbf{C}_{pp}}} \int d\mathbf{p} \exp \left(-\frac{1}{2} \mathbf{p}^\top \mathbf{C}_{pp}^{-1} \mathbf{p} \right) e^{-i \mathbf{L}_p \mathbf{p}^{(i)}}.
\end{aligned} \tag{3.67}$$

Now, we would like to carry out the momentum integral. Making use of the Gaussian integral, and dropping the (i) superscript, we now have:

$$Z_0[\mathbf{L}] = V^{-N} \int d\mathbf{q} \exp \left(-\frac{1}{2} \mathbf{L}_p^\top \mathbf{C}_{pp} \mathbf{L}_p - i \mathbf{L}_q \mathbf{q} \right). \tag{3.68}$$

For the case of the power spectrum, $l = 2$, the last expression simplifies to

$$\begin{aligned}
Z_0[\mathbf{L}] &= \frac{(2\pi)^3 \delta_D(k_1 + k_2)}{V^N} \int d\mathbf{q}_1 \exp \left(-\frac{1}{2} \mathbf{L}_p^\top \mathbf{C}_{pp} \mathbf{L}_p + ik_1 q_2 \right) \\
&= \frac{(2\pi)^3 \delta_D(k_1 + k_2)}{V^2} \int_{q_2} \exp \left(-\frac{1}{2} \mathbf{L}_p^\top \mathbf{C}_{pp} \mathbf{L}_p + ik_1 q_2 \right).
\end{aligned} \tag{3.69}$$

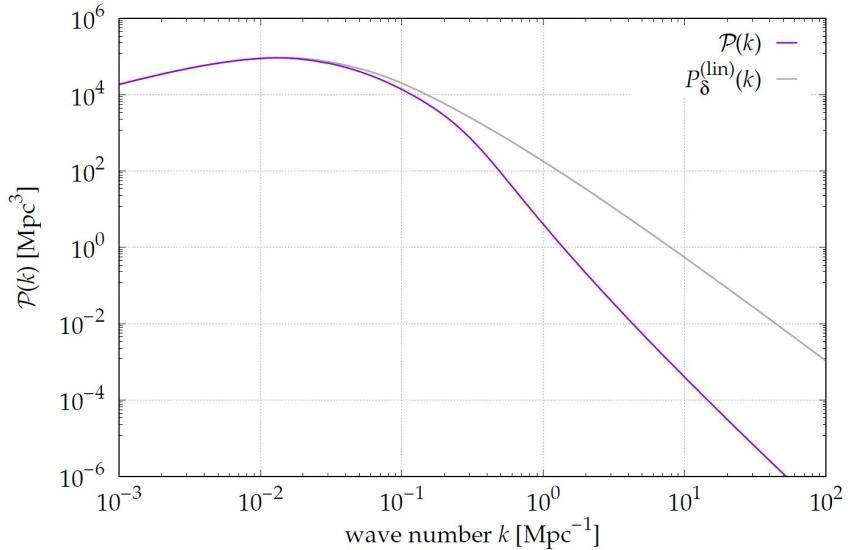


Figure 3.2: The non-linear power spectrum $\mathcal{P}(k)$ (3.72) at redshift $z = 0$ (purple line), with the linearly evolved power spectrum (gray line).

where $d\mathbf{q}_1 = \prod_{i=2}^N dq_i$ and we integrated over all the particles $j > 2$. Using the free two-point correlator function (3.56)

$$\left\langle \rho(\vec{k}_1, t_1) \rho(\vec{k}_2, t_2) \right\rangle = N^2 Z_0[\mathbf{L}], \quad (3.70)$$

and the relation with the free power spectrum $\mathcal{P}(k)$ (2.38)

$$(2\pi)^3 \delta_D(k_1 + k_2) [(2\pi)^3 \delta_D(k) + \mathcal{P}(k)] = \frac{V^2}{N^2} \left\langle \rho(\vec{k}_1, t_1) \rho(\vec{k}_2, t_2) \right\rangle, \quad (3.71)$$

we obtain, for $k \neq 0$,

$$\begin{aligned} \mathcal{P}(k) &= \int_q \exp \left(-\frac{1}{2} \mathbf{L}_p^\top \mathbf{C}_{pp} \mathbf{L}_p + i \vec{k} \cdot \vec{q} \right) \\ &= \int_q \exp \left(-t^2 \vec{k}^\top (C_{p_1 p_1} - C_{p_1 p_2}) \vec{k} \right) e^{i \vec{k} \cdot \vec{q}} \\ &= e^{-t^2 k^2 \sigma_1^2 / 3} \int_q \exp \left(-t^2 \vec{k}^\top a_{\parallel}(\vec{q}_2, \mu) \vec{k} \right) e^{i \vec{k} \cdot \vec{q}}, \end{aligned} \quad (3.72)$$

where $a_{\parallel}(\vec{q}_2, \mu) = a_1(\vec{q}_2) + \mu^2 a_2(\vec{q}_2)$ with μ the cosine of the angle between \vec{k} and \vec{q}_2 .

Chapter 4

Random Matrix Theory

Random matrix theory (RMT) is a powerful mathematical-physical tool that has been developed lately in the last century thanks to the work of Wishart [20] and then by Wigner [21, 22], who used it to describe the statistics of excited levels in complex nuclei. Random matrices methods have been developed by illustrious theorists such as Tao, Parisi, and Metha [19, 28, 29] spreading RMT into new areas of physics and mathematics. Nowadays random matrices are largely used in all branches of physics [30], and many applications can be found in [31, 32].

A special class of random matrices we will later focus on, are the Euclidean random matrices (ERMs) [18]. They play a significant role in the description of vibrations in topologically disordered systems [33–37], in relaxation in glasses [38–40] and in Anderson localization [41–43]. We will then analyze and consider some of the main results obtained for Hermitian ERMs [44], such as the high and low-density expansion, and apply them to the covariance matrices of Kinetic Field Theory.

4.1 Matrix basis

In this chapter, we will give the basis of RMT to understand the main aspects one should focus on to comprehend ERMs. In particular, we will analyze the relation between a large random matrix H and the statistical properties of its eigenvalues. As one could expect, the eigenvalue distribution strongly depends on the choice of the randomicity of the matrix. We will consider different kinds of random matrices, such as Gaussian random matrices, Wishart matrices, and Euclidean Random Matrix (ERM), where the last ensemble describes the covariance matrices of Kinetic Field theory. For further details, one could always look at [44, 45].

As we will deal with a lot of matrices from now on, is safe to recall some important relations between some properties of the matrix and its eigenvalues. Let us start by considering a real symmetric matrix H ($H = H^\top$), and let us state without proof that

Theorem 1. *Every symmetric matrix H has real eigenvalues.*

Hence

Theorem 2. *Every symmetric matrix H is similar to a diagonal matrix of its eigenvalues. In other words,*

$$H = H^\top \rightarrow H = \mathcal{O} D \mathcal{O}^\top, \quad (4.1)$$

where \mathcal{O} is an orthogonal matrix ($\mathcal{O}^{-1} = \mathcal{O}^\top$), and D is a diagonal matrix whose entries are the eigenvalues of H .

Since all the matrices that we will consider are going to be symmetric, this is a great simplification for our system. Furthermore, considering that a covariance matrix is obey to

Definition 3. *An $N \times N$ symmetric real matrix H is said to be **positive-semidefinite** if and only if*

$$\mathbf{x}^\top M \mathbf{x} \geq 0, \quad \forall \mathbf{x} \in \mathbb{R}^n. \quad (4.2)$$

we can finally state that

Theorem 4. *A positive-semidefinite matrix H has real positive eigenvalues*

constraining the eigenvalues on the real positive axis.

4.2 Gaussian random matrices

Let us consider by now the easiest example of a random matrix, namely a Gaussian random matrix. It is a $N \times N$ matrix H , whose entries are independently sampled from a Gaussian probability distribution function $\mathcal{N}(\mu, \sigma)$. Without loss of generality, we can assume that $\mu = 0$ and $\sigma = 1$. To exploit Theorem 1 we must symmetrize our matrix H using $H_s = \frac{H+H^\top}{2}$. After that H_s will be a symmetric matrix, hence with real eigenvalues. In this way, the matrix H_s belongs to the so-called Gaussian Orthogonal Ensemble (GOE). This is not of course the unique ensemble for a RMT, for example, one could in principle use complex or quaternionic entries, giving rise to the Gaussian Unitary Ensemble (GUE) and Gaussian Symplectic Ensemble (GSE). In that case one would require hermitianity and self-duality respectively, to have real eigenvalues. For practical reasons, we will analyze only the GOE, for any generalization, one could always consult [45]. In this section, we will focus on the GOE, namely the ensemble of $N \times N$ Hermitian (symmetric) matrices that have independent and identically distributed (i.i.d) zero-mean Gaussian entries.

Jpdf of the GOE

Let us get started with the jpdf of the entries for the GOE. Since our matrix is symmetric, the free entries of a $N \times N$ matrix are $\frac{N(N-1)}{2}$. The diagonal terms then read $(H_s)_{ii} = H_{ii}$ and the off-diagonal entries $(H_s)_{ij} = \frac{H_{ij} + H_{ji}}{2}$. For this reason, the variances of the off-diagonal entries are 1/2 of the one of diagonal entries. The pdf reads

$$\begin{aligned} \varrho[H_s] := \varrho[H_{11}, \dots, H_{NN}] &= \prod_{i=1}^N \frac{e^{-H_{ii}^2/2}}{\sqrt{2\pi}} \prod_{i < j}^N \frac{e^{-H_{ij}^2}}{\sqrt{\pi}} \\ &= \frac{1}{2^{N/2} \pi^{N(N+1)/4}} e^{-\frac{1}{2} \text{Tr}(H_s^2)}, \end{aligned} \quad (4.3)$$

where the normalization constant has been provided by [46]. Here is the critical point, knowing the pdf of the entries, in principle, we are allowed to derive the analytical expression of the eigenvalue distribution $\varrho(\lambda)$. That's why is possible to do it for the ensembles like GOE or Wishart matrices, but not for ERMs, for which we will see that the pdf of the entries is not known.

Wigner's surmise

Let us consider one of the first results from GOE, as we will see, this leads to a very general result, that is going to be valid for any ensemble, namely that the eigenvalues are strongly correlated among them. In particular, considering a 2×2 matrix belonging to the GOE we can study the spacing $s = \lambda_2 - \lambda_1$, with $\lambda_1 < \lambda_2$, between the two eigenvalues. This law is known under the name of Wigner's surmise and the behavior of s is given by

$$\varrho(s) = \frac{s}{2} e^{-\frac{s^2}{4}}. \quad (4.4)$$

This means that the probability of sampling two eigenvalues close to each other ($s \rightarrow 0$) is very small. It is as if each eigenvalue feels the presence of the other and tries to avoid it. This feature describes well the fact that the eigenvalues of random matrices are strongly not independent, making the jpdf not factorizable, meaning that Eq. (2.14) is not valid. This is a general property owned by Random Matrices known with the name of *level repulsion*: the eigenvalues repel each other, contrary to what random variables do. In this case the distribution of gaps between adjacent i.i.d. random variables follows the scaling [45]

$$\varrho(s) \sim e^{-s}. \quad (4.5)$$

The probability of vanishing gaps, $s \rightarrow 0$ does not vanish, leading to a "attractive" behaviour among the eigenvalues, rather than a "repulsive" as in the case of RMT.

Jpdf of Eigenvalues

The jpdf for the eigenvalues of a $N \times N$ Gaussian matrix is [45] can be found using Theorem 5 in D and reads

$$\varrho(\lambda_1, \dots, \lambda_N) = \frac{1}{Z_N} e^{-\frac{1}{2} \sum_{i=1}^N \lambda_i^2} \prod_{j < k} |\lambda_j - \lambda_k|, \quad (4.6)$$

with the normalization constant

$$Z_N = \frac{(2\pi)^{N/2}}{\Gamma(3/2)} \prod_{j=1}^N \Gamma(1 + j/2), \quad (4.7)$$

enforcing $\int_{\lambda} \varrho(\lambda_1, \dots, \lambda_N) = 1$. The Gamma function is defined as the usual $\Gamma(x) = \int_0^{+\infty} dt t^{x-1} e^{-t}$. The eigenvalues are meant to be unordered here.

Let us notice two things, that can be useful at this point. First, if we rewrite Eq. (4.6) in a slightly different way, namely exponentiating the product term $\prod_{j < k} |\lambda_j - \lambda_k| = \exp \left\{ \sum_{j < k} \ln |\lambda_j - \lambda_k| \right\}$,

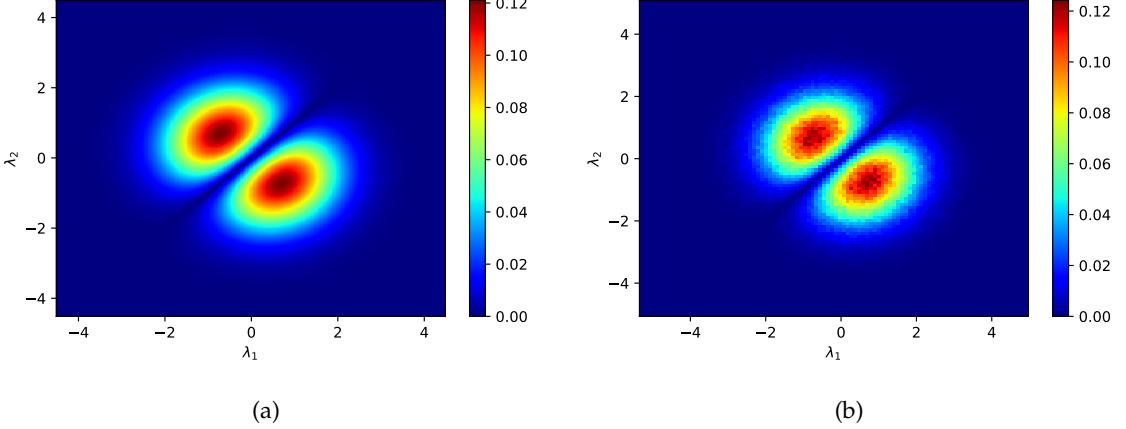


Figure 4.1: Eigenvalues for a 2×2 GOE matrix. Comparison between analytical expression (4.6) in (a) and numerical diagonalization in (b).

then the jpdf becomes of the form

$$\varrho(\lambda_1, \dots, \lambda_N) = \frac{e^{-\mathcal{V}(\lambda_1, \dots, \lambda_N)}}{\mathcal{Z}_N}, \quad (4.8)$$

of some function \mathcal{V} of the eigenvalues that represent the energy function of an ensemble. This is exactly the statistical physics definition for a system to be found in the state $\{\lambda^*\} = \lambda_1, \dots, \lambda_N$, as known as the *Boltzmann distribution* for the *canonical ensemble*. This feature will be used in the next section 4.2.1 to calculate the marginalized pdf $\varrho(\lambda)$ for the GOE, as the \mathcal{V} can represent the potential energy of a gas of charged particles as a function of their position λ .

Secondly, if we stare at Eq. (4.6) intensely, we will realize that it is in the same form as Eq. (1.37), and this is not a coincidence at all, this is a very general result from RMT, the Gaussian factor kills all the configurations with large eigenvalues, while the term $\prod_{j < k} |\lambda_j - \lambda_k|$ kills configurations where two eigenvalues get too close to each other. This repulsion factor also makes the eigenvalues strongly non-independent. Every eigenvalue feels the presence of all the others, and the jpdf 4.6 does not factorize. Hence, the classical tools for independent random variables are useless here.

Semicircle law

So how would it be possible to recover the marginalized pdf $\varrho(\lambda) = \int_{\lambda} \varrho(\lambda_1, \dots, \lambda)$ if the jpdf is not factorizable? The answer to this question is given in the next section, for now, let us just analyze the result, and let us see what we can learn from this. Rescaling the eigenvalues $\lambda_i \rightarrow \frac{\lambda_i}{\sqrt{N}}$ we can find the marginal probability distribution function in the limit of $N \rightarrow \infty$, obtaining the so-called *Wigner's semicircle law*

$$\varrho_{\text{SC}}(\lambda) = \frac{1}{\pi} \sqrt{2 - \lambda^2}, \quad \lambda \in [-\sqrt{2}, +\sqrt{2}]. \quad (4.9)$$

As we can see in Fig 4.2, the match becomes better as the dimension N of the matrix increases. A result for finite N will be discussed in section 4.4.3 and can be found in [45].

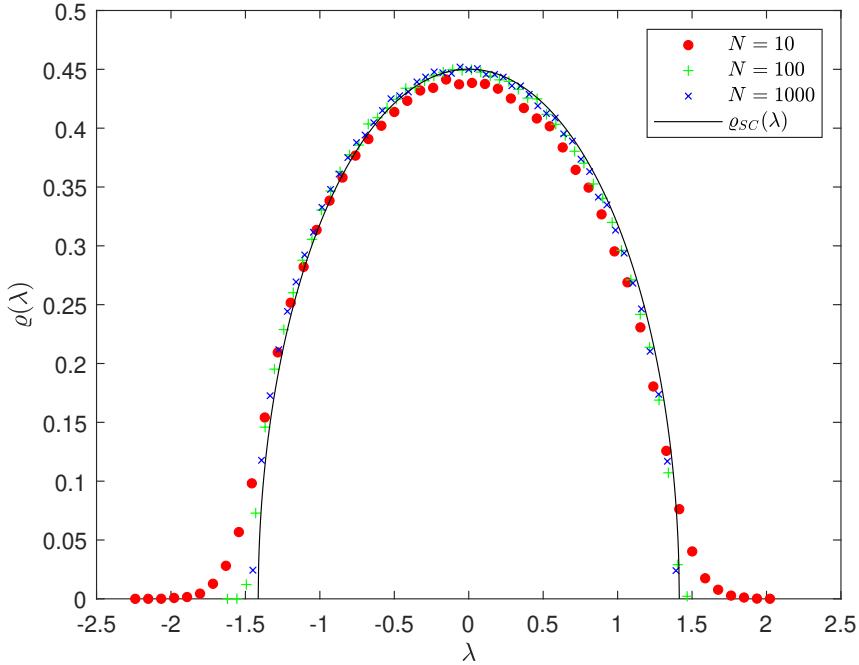


Figure 4.2: Numerical validation of the semicircle law for different sizes N of GOE matrix.

There are several ways to arrive at Eq. (4.9), using the most diverse technique of mathematical physics such as the Coulomb gas technique or making use of the complex resolvent. We will analyze both techniques as they are general, and not only valid for the GOE but could be used for a large class of Random matrices.

4.2.1 Coulomb gas

The Coulomb gas (Dyson gas) technique has been used to evaluate the spectrum of GOE Random matrices by Dyson and Wigner [47, 48]. The reason for the first name will be clear soon. Let us consider the jpdf for the GOE, rescaling the eigenvalues such as $\lambda_i \rightarrow \lambda_i \sqrt{N}$, the normalization constant reads

$$\mathcal{Z}_N = \mathcal{C}_N \int_{\lambda} e^{-\beta N^2 \mathcal{V}[\lambda]}, \quad (4.10)$$

with $\mathcal{C}_N = (\sqrt{N})^{N+N(N-1)/2}$, and where we called the potential

$$\mathcal{V}[\lambda] = \frac{1}{2N} \sum_i \lambda_i^2 - \frac{1}{2N^2} \sum_{i \neq j} \ln |\lambda_i - \lambda_j|, \quad (4.11)$$

obtaining the canonical partition function, integrating the Gibbs-Boltzmann weight $e^{-\beta N^2 \mathcal{V}[\lambda]}$, with the fictitious inverse temperature $\beta = 1$, over all possible positions λ of the particles. It describes a set of static particles at thermodynamical equilibrium on a line, under the effect of a quadratic potential and a repulsive logarithmic term. The term βN^2 ensures that taking the limit of $N \rightarrow \infty$ is both a thermodynamic and zero-temperature limit. At this point, the problem has

been brought in a classical statistical physics problem, namely finding the equilibrium positions λ_i at zero temperature of the particles minimizing the free energy

$$F = -\frac{1}{\beta} \ln \mathcal{Z}_{\mathcal{N}}, \quad (4.12)$$

of the system.

At this point, it is clear the reason for the name, indeed, the logarithmic repulsive potential is characteristic of a two-dimensional charge. The 2-d electric field and the potential, generated by a point charge on a plane, can be derived by Gauss's law assuming rotational symmetry

$$\vec{E}(r) \sim \frac{1}{r} \hat{r}, \quad V(r) = \int E(r) dr \sim \ln r. \quad (4.13)$$

In this way, our set of eigenvalues can be described by charged particles in 2 dimensions, constrained on a line, repelling each other by Coulomb interaction, and submitted to an external potential, determined by the form of the probability distribution $\varrho[H]$. This also explains why this method is general, it can be applied as far as the distribution of the entries of the random matrix is known. For a complete discussion of the calculations, [45] is a good reference.

4.2.2 Resolvent

A completely different approach could be the one involving the use of the complex function known as *resolvent*

$$G_N(z) = \frac{1}{N} \text{Tr}(z \mathbb{1}_N - H)^{-1} = \frac{1}{N} \sum_{i=1}^N \frac{1}{z - \lambda_i}, \quad z \in \mathbb{C} \quad (4.14)$$

for a random matrix H . The resolvent has N poles at the locations λ_i of each eigenvalue. Since we are interested in the limit of $N \rightarrow \infty$, the sum gets converted into an integral, and the poles will form a cut on the real line, so that, averaging

$$\langle G_N(z) \rangle \rightarrow G_{\infty}^{(\text{av})}(z) = \int dx \frac{\varrho(x)}{z - x}, \quad (4.15)$$

where we weighted the integrand with the average density of eigenvalues $\varrho(\lambda)$ at point λ . Now, by taking the limit of small ε of

$$G_{\infty}^{(\text{av})}(z) = \lim_{\varepsilon \rightarrow 0} G_{\infty}^{(\text{av})}(x - i\varepsilon), \quad (4.16)$$

and with

$$G_{\infty}^{(\text{av})}(x - i\varepsilon) = \int dx' \frac{\rho(x')}{x - i\varepsilon - x'} = \int dx' \frac{\rho(x')(x - x')}{(x - x')^2 + \varepsilon^2} + i \int dx' \rho(x') \frac{\varepsilon}{(x - x')^2 + \varepsilon^2}. \quad (4.17)$$

Considering both the real and the imaginary part, the two following relations can be derived:

$$\begin{aligned} \varrho(\lambda) &= -\frac{1}{\pi} \lim_{\varepsilon \rightarrow 0} \text{Im} \left\{ G_{\infty}^{(\text{av})}(x - i\varepsilon) \right\}, \\ \text{Pr} \int_{-\infty}^{+\infty} d\lambda \frac{\varrho(\lambda)}{\lambda - \lambda'} &= \text{Re} \left\{ G_{\infty}^{(\text{av})}(x - i\varepsilon) \right\}, \end{aligned} \quad (4.18)$$

using the following representation of the delta function

$$\delta(x) = \frac{1}{\pi} \lim_{\varepsilon \rightarrow 0} \frac{\varepsilon}{x^2 + \varepsilon^2}. \quad (4.19)$$

Using the first of Eqs. (4.18), it is possible to recover in a few steps the semicircle law [45].

4.3 Wishart-Laguerre matrices

Let us consider, for completeness, another class of matrices, namely Wishart Laguerre (WL) matrices, $N \times N$ square matrices W with correlated entries. They are constructed as $W = HH^\dagger$, where H is a $N \times M$ matrix ($M \geq N$) filled with i.i.d. Gaussian entries. In the case of real entries, the \dagger is replaced by $^\top$. Differently from Gaussian matrices, Wishart matrices have N non-negative eigenvalues. Indeed, Wishart matrices W are positive semidefinite.

Jpdf of the Wishart-Laguerre ensemble

Also in this case is possible to write the jpdf of the entries of a matrix W belonging to WL ensemble

$$\varrho[W] = \frac{1}{2^{MN/2} \Gamma_N(M/2)} e^{-\frac{1}{2} \text{Tr } W} (\det W)^{\frac{1}{2}(M-N-1)}, \quad (4.20)$$

where the multivariate Gamma function has the following expression

$$\Gamma_n(x) = \pi^{n(n-1)/4} \prod_{i=1}^n \Gamma\left(x - \frac{i-1}{2}\right). \quad (4.21)$$

From Eq. (4.20) we can realize that the entries are correlated, as the determinant kills any chance of factorizing the jpdf, unless for specific combinations of M and N for which the determinant disappears.

Jpdf of Eigenvalues

The jpdf of eigenvalues can be written down immediately, expressing everything in terms of the eigenvalues and adding the Vandermonde determinant, see Theorem 5 in D

$$\varrho(\lambda_1, \dots, \lambda_N) = \frac{1}{Z_N^{(L)}} e^{-\frac{1}{2} \sum_{i=1}^N \lambda_i} \prod_i^N \lambda_i^{\alpha/2} \prod_{j < k} |\lambda_j - \lambda_k|, \quad (4.22)$$

where $\alpha = M - N - 1$ and $Z_N^{(L)} = \frac{\Gamma_M(N/2)\Gamma_N(M/2)}{\pi^{N^2}}$, and the eigenvalues are meant to be unordered.

Marčenko-Pastur density

The average density of eigenvalues has the following scaling form for $N, M \rightarrow \infty$ (such that $c = \frac{N}{M} \leq 1$ is kept fixed)

$$\varrho(\lambda) \rightarrow \frac{1}{N} \varrho_{\text{MP}}\left(\frac{\lambda}{N}\right), \quad (4.23)$$

where the Marčenko-Pastur scaling function (the analogue of the semicircle $\varrho_{SC}(\lambda)$ for the Gaussian ensemble) is given by [49]

$$\varrho_{\text{MP}}(\lambda) = \frac{1}{2\pi\lambda} \sqrt{(\lambda - \zeta_-)(\zeta_+ - \lambda)}, \quad \lambda \in (\zeta_-, \zeta_+], \quad (4.24)$$

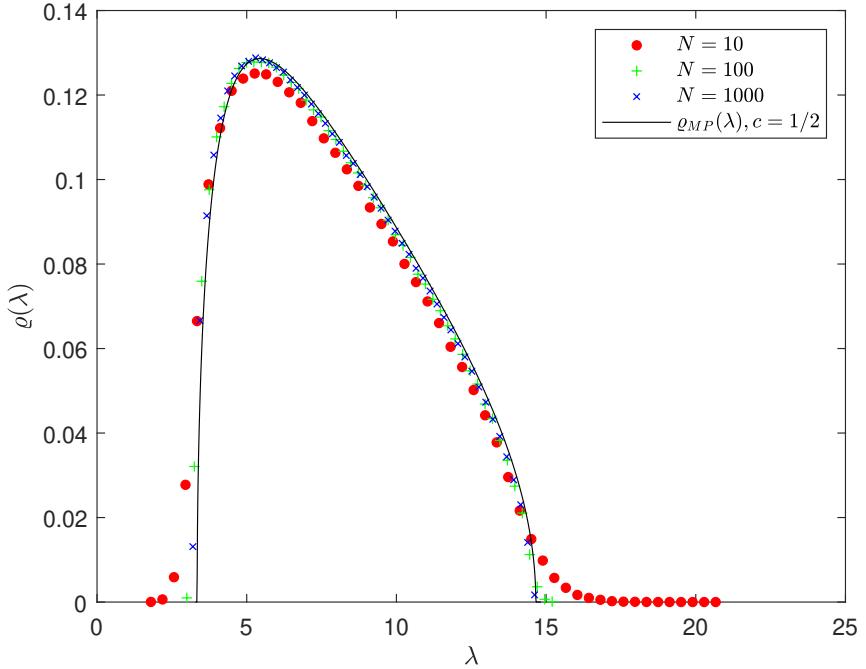


Figure 4.3: Comparison between the Marcenko-Pastur density, for $c = 1/2$, and the corresponding histograms obtained from the numerical diagonalization of different sizes N of random Wishart matrices. The histograms are obtained from respectively 10^5 , 10^4 , and 10^3 Wishart matrices.

with $\zeta_{\pm} = (1 \pm c^{-1/2})^2$. Also in this case, as for the GOE, it is possible to use for example the Coulomb gas technique or the resolvent method.

4.4 Euclidean random matrices

After an introduction to the previous two important classes of random matrices, we will now focus on a third special class of random matrices, namely the so-called Euclidean random matrices. This class of random matrices is important for us as it can describe a covariance matrix. Indeed, we can use ERM for the density fluctuations correlation matrix $C_{\delta\delta}$, for the momentum component correlation matrix $C_{p_i p_j}$ or its absolute value $C_{|p||p|}$. Furthermore, also the complete momentum correlation matrix, under certain assumptions, can be identified as an ERM. But what is exactly an ERM, and why is it called like that? Each entry of a $N \times N$ Euclidean random matrix is identified by a deterministic function $f(\mathbf{r}_i - \mathbf{r}_j)$ of the positions of pairs of points randomly distributed in a region of finite volume V of the Euclidean space, and hence the name. The single entries then read $A_{ij} = f(\mathbf{r}_i, \mathbf{r}_j)$, $i = 1, \dots, N$. When we treat the covariance matrices of KFT as ERMs, then, the function f is always only a function of the absolute value of the distance $f(r_{ij}) = f(|\mathbf{r}_i - \mathbf{r}_j|)$ and this role is played by the correlation function $\xi(r)$ or a combination of $a_1(r)$ and $a_2(r)$. The points are uniformly distributed in a finite region of the three-dimensional space, and represent the particles of our KFT framework. Differently from the previous two classes of random matrices, for which $\varrho[A]$ is known, for the ERM ensemble it is

not known, so is impossible to derive $\varrho(\lambda_1, \dots, \lambda_N)$ immediately. There are nonetheless simple heuristic approaches that give good results under certain conditions, which are going to be studied in the next section. In particular, if happens that the function f decays fast enough for large r that, in the limit of $V \rightarrow \infty$ the eigenvalue density $\varrho(\lambda)$ depends only on the density $\rho = N/V$, then it is possible to derive a simple expression for the eigenvalue distribution in the high or low-density regime.

It is important to notice that all the diagonal terms of every autocorrelation matrix are equal to each other since they compute the correlation of any particle with itself. Therefore all the matrices have been rescaled by the diagonal terms. More details about numerical diagonalization can be found in Appendix G

4.4.1 High density expansion

For any ERM A , we can always write

$$\sum_{j=1}^N A_{ij} \Phi_j(\mathbf{k}) = \lambda_i(\mathbf{k}) \Phi_i(\mathbf{k}), \quad (4.25)$$

with the eigenvectors components $\Phi_i(\mathbf{k}) = e^{i\mathbf{k} \cdot \mathbf{r}_i}$ and

$$\lambda_i(\mathbf{k}) = \sum_{j=1}^N e^{-i\mathbf{k} \cdot (\mathbf{r}_i - \mathbf{r}_j)} f(\mathbf{r}_i - \mathbf{r}_j). \quad (4.26)$$

We assume now that the density $\rho = \frac{N}{V}$ is large enough for the phase $-i\mathbf{k} \cdot (\mathbf{r}_i - \mathbf{r}_j)$ to vary only weakly between neighboring points r_i and r_j . This condition occurs when $\rho^{1/3} \gg k$. In this way, the sum (4.26) can be approximated by an integral, and $\lambda_i(\mathbf{k})$ does not depend on i anymore, becoming

$$\lambda(\mathbf{k}) = \rho \int_r f(\mathbf{r}) e^{-i\mathbf{k} \cdot \mathbf{r}} := \rho f_0(\mathbf{k}), \quad (4.27)$$

with $f_0(\mathbf{k})$ being the Fourier transform of $f(\mathbf{r})$. Summing over the different eigenvalues labeled by \mathbf{k} we obtain

$$\varrho(\lambda) = \frac{1}{\rho} \int_k \delta[\lambda - \rho f_0(\mathbf{k})]. \quad (4.28)$$

From a more physical point of view, the high-density limit is an interesting case, since the number particle density ρ is high at late cosmic time, when large-scale structures have already formed. In this sense simulating an ensemble of particles at high density would describe CDM particles at late cosmic time.

4.4.2 Low density expansion

In the low-density limit $\rho \rightarrow 0$, for a rapidly monotonically decaying function $f(\mathbf{r}_i - \mathbf{r}_j)$, the matrix element A_{ij} are not negligible only if the points \mathbf{r}_i and \mathbf{r}_j are nearest neighbors. In this case, the matrix A can be approximated by a block diagonal matrix with $N/2$ 2×2 blocks. The

eigenvalues of each block will be then $\lambda_{1,2} = f(\mathbf{0}) \pm f(\mathbf{r}_i - \mathbf{r}_j)$. Then the eigenvalue distribution is

$$\varrho(\lambda) = \frac{1}{2} \int_{\Delta r} p_{nn}(\Delta r) \{ \delta[\lambda - f(0) - f(\Delta r)] + \delta[\lambda - f(0) + f(\Delta r)] \}, \quad (4.29)$$

where $p_{nn}(\Delta r)$ is the probability (C.4) of finding two nearest neighbors at a distance Δr :

$$p_{nn}(\Delta r) = d\mathcal{V}\rho\Delta r^{d-1}e^{-\mathcal{V}\rho\Delta r^d} = 4\pi\rho\Delta r^2e^{-\frac{4}{3}\pi\rho\Delta r^3}, \quad (4.30)$$

with the volume of the d -dimensional unit sphere $\mathcal{V} = \frac{\pi^{d/2}}{\Gamma(d/2+1)}$. Therefore the cumulative probability distribution function (cpdf) $C(\lambda) = \int_{\lambda}^{\infty} d\lambda' \varrho(\lambda')$ takes the form

$$C(\lambda) = \begin{cases} 1 & \lambda < 0; \\ \frac{1}{2} \operatorname{sgn}[f(0) - \lambda] e^{-\mathcal{V}\rho\{f^{-1}(|f(0)-\Lambda|)\}^d} + \frac{1}{2} & 0 < \lambda < 2f(0); \\ 0 & \lambda > 2f(0). \end{cases} \quad (4.31)$$

Since

$$\varrho(\lambda) = -\frac{d}{d\lambda}C(\lambda), \quad (4.32)$$

fixing $f(0) = 1$ in some units system, the previous equation yields

$$\varrho(\lambda) = -\frac{\mathcal{V}\rho d}{2} e^{-\mathcal{V}\rho\{f^{-1}(|1-\lambda|)\}^d} f^{-1}(|1-\lambda|)^{d-1} (f^{-1}(|1-\lambda|))', \quad 0 < \lambda < 2. \quad (4.33)$$

4.4.3 Considerations

When we consider ERM is important to spend some time thinking about how to set the parameters, namely how large can the matrix be, which densities are worth considering, and so on. The limiting factor here is the number of particles since a $N \times N$ ERM describes a N particle set. In the numerical simulations done in this thesis, the maximum that has been considered is a 1000×1000 matrix. This leads to lots of limitations as one may expect. For example, in principle, one could expect that the eigenvalue spectrum of a random matrix (not Euclidean) changes with the size of the matrix, and this is indeed true. One can just think at the different eigenvalue distribution for a GOE matrix: recall that the semicircle law (4.9) is only valid for $N \rightarrow \infty$. For any even finite N , for example, the spectrum changes according to [30, 50]

$$\varrho_{\text{GOE},N}(\lambda) = \frac{1}{2N} \sum_{k=0}^{N/2-1} e^{-\frac{\lambda^2}{2}} [R_{2k}(\lambda)\Phi_{2k+1}(\lambda) - R_{2k+1}(\lambda)\Phi_{2k}(\lambda)], \quad (4.34)$$

where

$$\Phi_k(x) = \int_{-\infty}^{+\infty} dy R_k(x)e^{-\frac{y^2}{2}} \operatorname{sign}(x-y), \quad (4.35)$$

and

$$R_k(x) = \begin{cases} \frac{\sqrt{2}}{\pi^{\frac{1}{4}} 2^{\frac{k}{2}} (k)!!} H_k(x) & k \text{ even}; \\ \frac{\sqrt{2}}{\pi^{\frac{1}{4}} 2^{\frac{k+3}{2}} (k-2)!!} [-H_k(x) + 2(k-1)H_{k-2}(x)] & k \text{ odd}; \end{cases} \quad (4.36)$$

and the $H_k(x)$ are the Hermite polynomials. For an ERM the situation is a bit more delicate since the matrix depends on the distances between the particles. Indeed, *a priori* the spectrum is

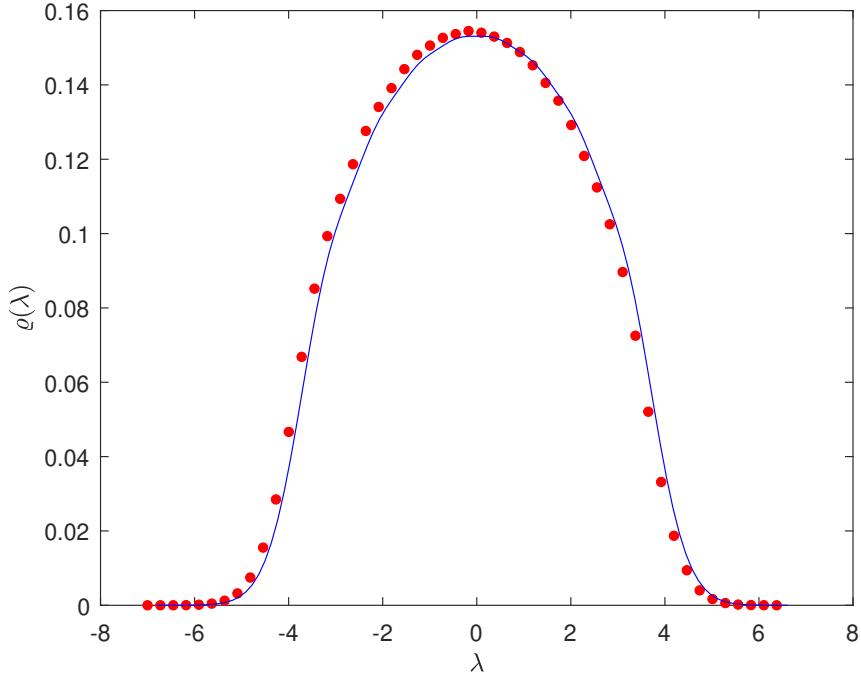


Figure 4.4: Comparison between numerically generated eigenvalue histograms of 10^6 GOE matrices of size $N = 8$ and the corresponding theoretical densities (4.34).

a function of the number density ρ and the number of points N . So, differently from the GOE, any ERM depends on two parameters. All this is true in principle, but actually, under some conditions, the eigenvalue spectrum could only depend on the density ρ and not anymore on N in the limit of $N \rightarrow \infty$. This becomes true if the random matrix $A_{ij}(\vec{r}) = f(\vec{r}_i - \vec{r}_j)$ is defined by a function $f(\vec{r})$ that decays fast enough for large \vec{r} . In particular, for all our applications will be always the case that $A_{ij}(r) = f(|\vec{r}_i - \vec{r}_j|)$ so from now on the \vec{r} will be replaced with the simpler r . If that is true, in the limit of $V \rightarrow \infty$ the translational invariance is restored at fixed density $\rho = \frac{N}{V}$. So if this would be our case we could ignore that our simulations have a limited number of matrix entries. In the first section of the next chapter, we will discuss all the possible choices for the function f and their relations to cosmological meaning. In particular, we can state that if the function is a Gaussian or a decreasing exponential

$$f(r) \sim e^{-r^2/2a^2}, \quad f(r) \sim e^{-r/R}, \quad (4.37)$$

then the spectrum does not depend on the matrix's size for large N . In contrast, if it is considered a function that does not decay fast enough, and even more, oscillates, such as

$$f(r) = \frac{\sin(k_0 r)}{k_0 r}, \quad (4.38)$$

whose spectrum is shown in fig 4.6 , its eigenvalues distribution depends on N , even for large N . In this case, we wouldn't be allowed to use the high and low-density expansions 4.4.1 and 4.4.2. Last but not least, let us consider a function that decays as a power law for large r

$$f(r) \sim r^{-n}. \quad (4.39)$$

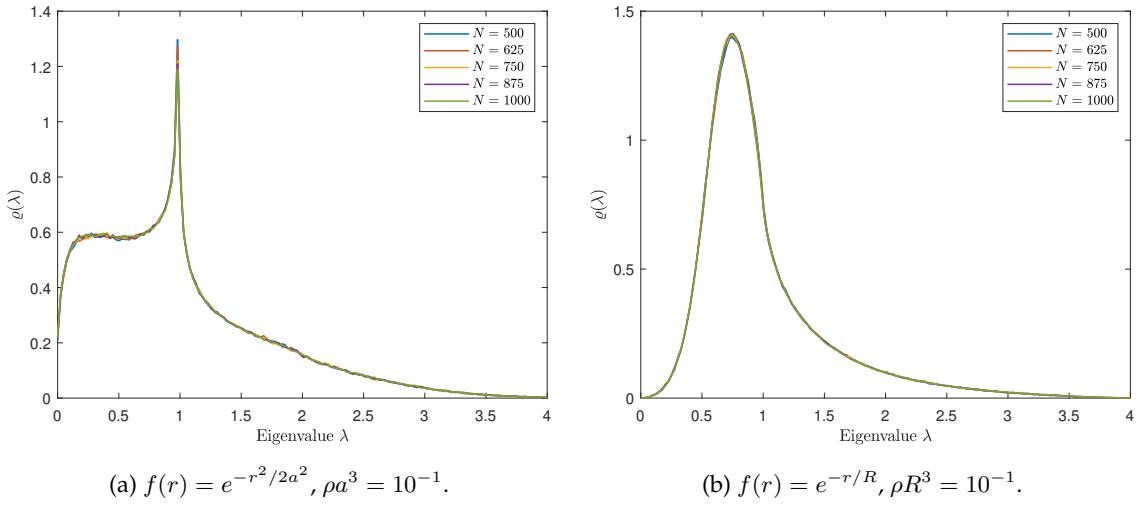


Figure 4.5: Eigenvalues distribution for $N \times N$ Gaussian and exponential ERM, obtained through diagonalization of 10^3 matrices.

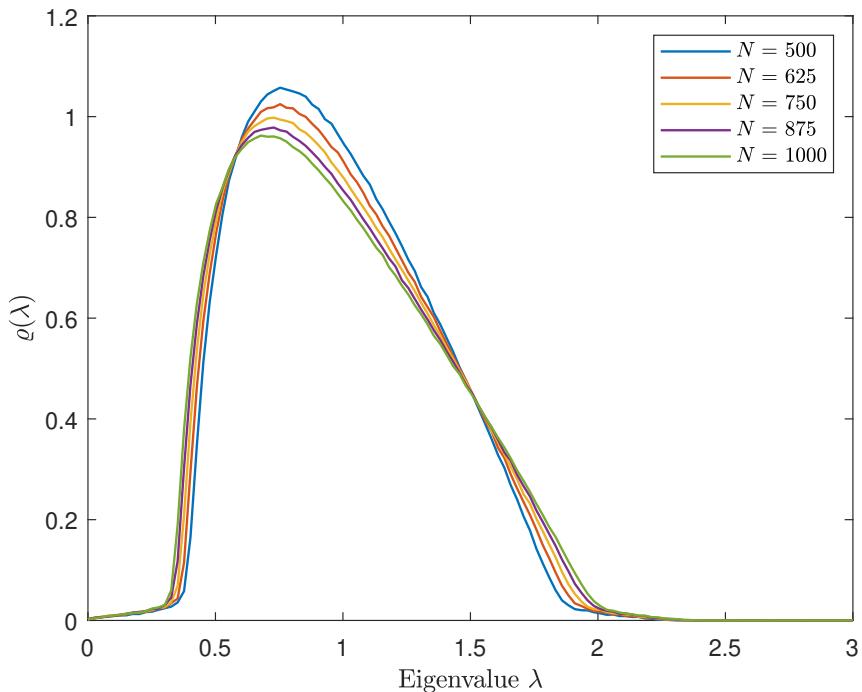


Figure 4.6: Eigenvalues distribution for a $f(r) = \frac{\sin(k_0 r)}{k_0 r}$ ERM, at number density $\rho \lambda_0^3 = 10^{-1}$, where $\lambda_0 = \frac{2\pi}{k_0}$. Obtained through diagonalization of 10^3 matrices.

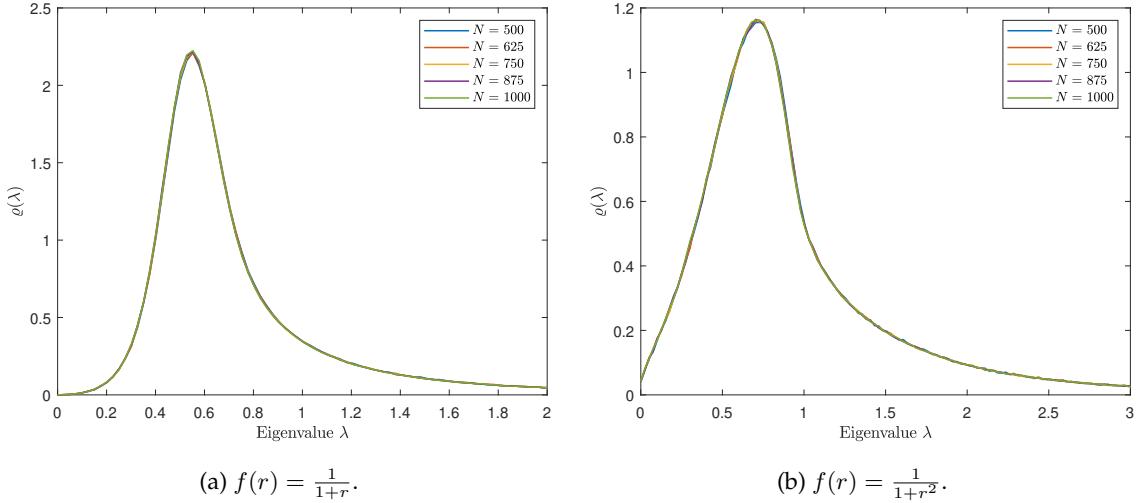


Figure 4.7: Eigenvalues distribution for power-law decaying ERM, at number density $\rho = 10^{-1}$ and different sizes. Obtained through diagonalization of 10^3 matrices.

It is thus interesting to understand what exponent n enables this "size" invariance. Let us observe how the distribution of the eigenvalues depends on the size N of the matrix for large $|\vec{r}|$. We are interested in this case because both the correlation functions $a_1(r)$ and $a_2(r)$ take the form $f(r) \sim r^{-2}$, in the limit of large r . As we can see, in this case, we can safely consider our function to decay fast enough, Fig. 4.7a and 4.7b.

After these considerations, we can be free to use the high and the low-density expansions presented in the previous sections to the autocorrelation ERM of KFT such as

$$C_{\delta\delta} = \xi_\delta(r), \quad C_{|p||p|} = -3a_1(r) - a_2(r). \quad (4.40)$$

The results are shown in the next chapter.

Chapter 5

Applications and results

This chapter is structured as follows: first, we will need to discuss the different choices of the possible power spectrum, and hence correlation functions, that can be done. Then we can analyze the eigenvalue distribution of the covariance matrices $C_{\delta\delta}$, $C_{|p||p|}$ and \mathbf{C}_{pp} , where the last one does not belong to the ERM ensemble. We will also investigate the high and low-density limits previously introduced. Finally, we discuss what is the relation between these spectra and the generating functional of KFT.

5.1 Choice of correlations

Before discussing the eigenvalues distribution of the correlation matrices, let us first discuss the form of the possible and interesting correlation functions and the power spectrum that we will use in this section. In particular, we will consider three kinds of density contrast power spectrum $P_\delta(k)$, one derived by an exponential correlation function, a Gaussian power spectrum, and the one that describes the Cold Dark Matter (CDM).

Gaussian power spectrum

For a Gaussian correlation function, and therefore for a Gaussian power spectrum.

$$\xi_\delta(r) = Ae^{-\frac{r^2}{2a^2}}, \quad P_\delta(k) = A(2\pi)^{\frac{3}{2}}a^3e^{-\frac{k^2a^2}{2}}, \quad (5.1)$$

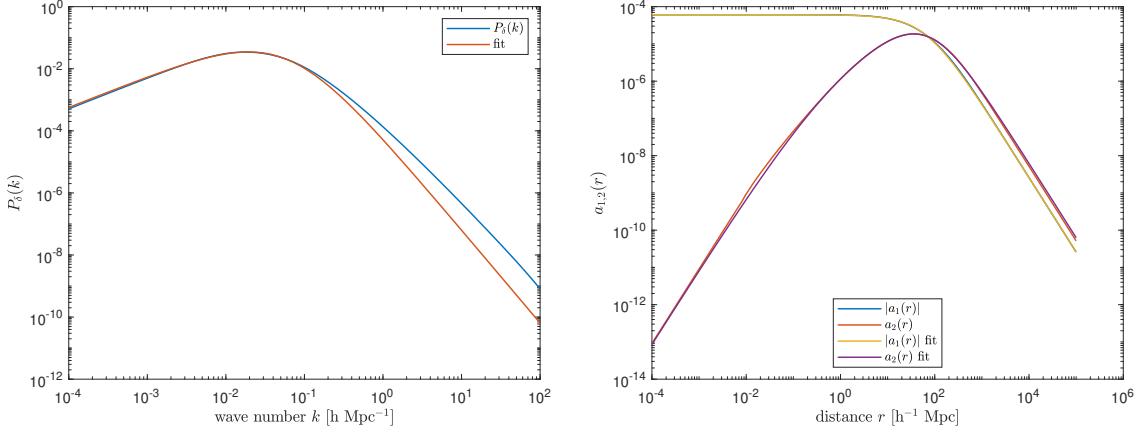
using Eq. (B.3). Furthermore

$$\sigma_1^2 = \frac{A}{16\pi^2}, \quad \sigma_2^2 = \frac{A}{16\pi^2a^3}. \quad (5.2)$$

Unfortunately, there are no analytic forms for the correlation functions $\zeta(r)$, $a_1(r)$ and $a_2(r)$. Nonetheless, they can be derived numerically.

Exponential correlation function

The first choice that is investigated, is the one of an exponentially decaying correlation function, we can derive the Power spectrum analytically using the definition (2.30), and the calculation can



(a) Initial density power spectrum, with the fitting function (5.6)

(b) Momentum correlation function $a_1(r)$ and $a_2(r)$ with their fit (5.8)

Figure 5.1: Statistical functions of CDM model.

be found in Appendix F

$$\xi_\delta(r) = Ae^{-r/R}, \quad P_\delta(k) = \frac{8\pi A}{R \left(\frac{1}{R^2} + k^2 \right)^2}, \quad (5.3)$$

where A is the correlation amplitude and R is the correlation length. We can compute the entire set of variables

$$\sigma_1^2 = AR^2, \quad \sigma_2^2 = A, \quad (5.4)$$

The full calculations can be found in F. With some symbolic effort

$$\begin{aligned} \zeta(r) &= \sigma_2^2 \frac{R^3}{r^2} \left\{ 2 - e^{-r/R} \left[2 + \frac{r}{R} \left(2 + \frac{r}{R} \right) \right] \right\}, \\ a_1(r) &= -\sigma_1^2 \frac{R^3}{r^3} e^{-r/R} \left(2 + \frac{r}{R} \right) \left[2 + \frac{r}{R} - e^{r/R} \left(2 - \frac{r}{R} \right) \right], \\ a_2(r) &= \sigma_1^2 \frac{R^3}{r^3} \left\{ \frac{r^2}{R^2} - 12 + e^{-r/R} \left(2 + \frac{r}{R} \right) \left[6 + \frac{r}{R} \left(3 + \frac{r}{R} \right) \right] \right\}, \end{aligned} \quad (5.5)$$

Cold Dark Matter

The power spectrum that better fits the data for Cold Dark Matter (CDM) has been numerically determined by Bardeen *et al.* and it is of the form of

$$P_\delta(k) = \frac{Ak}{\left(1 + \frac{k}{k_0} \right)^4}, \quad (5.6)$$

with the numerical value of the amplitude $A = 5.69 \text{ Mpc}^4 \text{h}^{-4}$ and the scale $k_0 = 0.0580 \text{ h Mpc}^{-1}$.

Furthermore

$$\sigma_1^2 = 1.8179 \cdot 10^{-4} \frac{\text{Mpc}^2}{\text{h}^2}, \quad \sigma_2^2 = 1.3268 \cdot 10^{-4}, \quad (5.7)$$

and with the fitted momentum correlation functions

$$\begin{aligned} a_1(r) &= -\frac{\sigma_1^2}{3} \left(1 + \frac{r}{\alpha_1} \left(1 + \frac{r}{\beta_1} \right) \right)^{-1}, \\ a_2(r) &= \frac{\sigma_2^2}{15} r^2 \left(1 + \sqrt{\frac{r}{\alpha_2}} \left(1 + \frac{r}{\beta_2} \right) \right)^{-8/3}, \end{aligned} \quad (5.8)$$

with the fit coefficients $\alpha_1 = 42.50 \text{ Mpc h}^{-1}$, $\beta_1 = 101.02 \text{ Mpc h}^{-1}$, and $\alpha_2 = 0.77 \text{ Mpc h}^{-1}$, $\beta_2 = 75.30 \text{ Mpc h}^{-1}$, they are shown in Fig. 5.11a. In the end, we can use Eq. (3.39) to obtain $\xi(r)$ and $\zeta(r)$, writing for simplicity

$$a_2(r) = \frac{\sigma_2^2 r^2}{15} f^{-n}(r), \quad f(r) = 1 + \sqrt{\frac{r}{\alpha_2}} \left(1 + \frac{r}{\beta_2}\right), \quad (5.9)$$

with $n = \frac{8}{3}$, the derivatives are

$$\begin{aligned} a'_2(r) &= a_2(r) \left[\frac{2}{r} - n \frac{f'(r)}{f(r)} \right], \\ a''_2(r) &= a_2(r) \left[\frac{2}{r^2} - 4 \frac{n}{r} \frac{f'(r)}{f(r)} + n(n+1) \frac{f'^2(r)}{f^2(r)} - n \frac{f''(r)}{f(r)} \right], \end{aligned} \quad (5.10)$$

yielding

$$\begin{aligned} \xi_\delta(r) &= a_2(r) \left[\frac{15}{r^2} - 9 \frac{n}{r} \frac{f'(r)}{f(r)} + n(n+1) \frac{f'^2(r)}{f^2(r)} - n \frac{f''(r)}{f(r)} \right], \\ \zeta(r) &= a_2(r) \left[\frac{5}{r} - n \frac{f'(r)}{f(r)} \right]. \end{aligned} \quad (5.11)$$

5.2 Eigenvalues distribution of $C_{\delta\delta}$

In this section, we will consider the density fluctuation correlation matrix $C_{\delta\delta}(r) = \xi_\delta(r)$, in this way, we define our correlation matrix as an ERM with entries defined by the function $f(|\mathbf{r}_i - \mathbf{r}_j|) = f(r_{i,j})$:

$$C_{\delta_i \delta_j}(r) = f(|\mathbf{r}_i - \mathbf{r}_j|) = \xi_\delta(r_{i,j}). \quad (5.12)$$

This case can be easily treated with the ERM machinery. In particular, will be considered the high and the low-density limit. The choice of $\xi_\delta(r)$ is essential, as its form defines the eigenvalues distribution. After the consideration made in 4.4.3 we can safely assume that our eigenvalues spectra do not depend on the dimension of the matrix in the limit of large N , but only on the particle density ρ . We will consider all three cases introduced in the previous section, and in particular, for the CDM case, we use the numerical integrated version for

$$\xi_\delta(r) = \frac{1}{2\pi^2} \int_0^{+\infty} dk k^2 P_\delta(k) j_0(kr), \quad (5.13)$$

assuming for $P_\delta(k)$ the fit expression (5.6).

As we can notice from the plots 5.8 and 5.3, in any considered case, the behavior of the distribution is similar, gets a strong peak in $\lambda = 0$ for the high-density limit, and a peak in $\lambda = 1$ for the low-density limit. This pushes us to consider the high and the low-density limits using the main results of ERM theory.

5.2.1 High density limit

In the high-density limit, using the formula (4.28)

$$\varrho(\lambda) = \frac{1}{\rho} \int_k \delta[\lambda - \rho f_0(k)] = \frac{1}{\rho} \int_k \delta[\lambda - \rho P_\delta(k)], \quad (5.14)$$

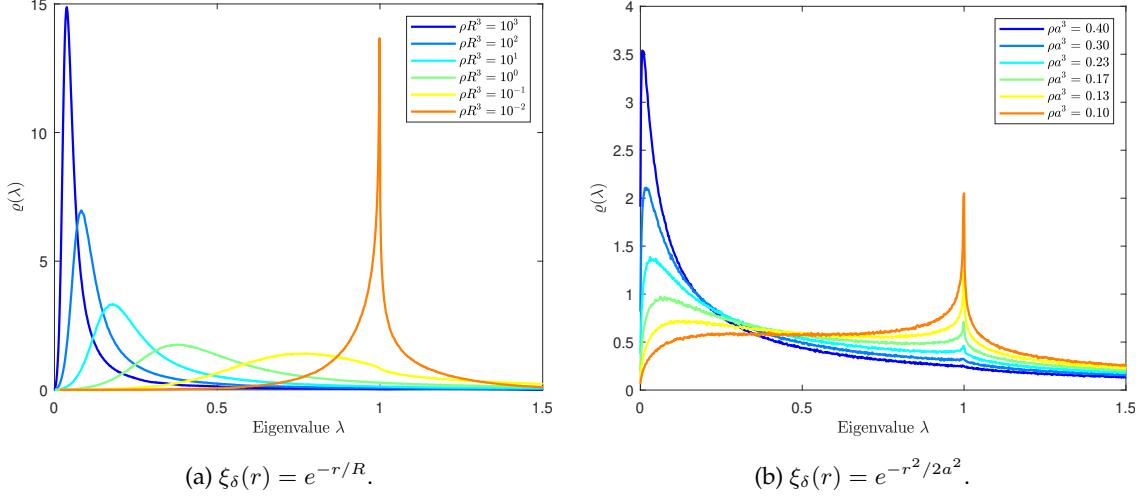


Figure 5.2: Eigenvalue distribution of 500×500 $C_{\delta\delta}$ ERM at different densities, obtained with 10^4 samples.

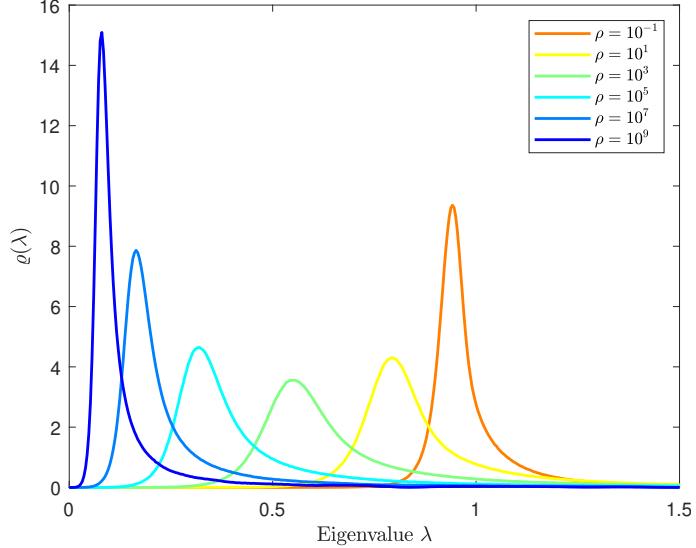


Figure 5.3: Numerical evaluation of the eigenvalue distribution obtained with 10^4 samples of 500×500 density contrast correlation matrix $C_{\delta\delta} = \xi_\delta(r)$ using the Bardeen fit.

where

$$f_0(k) = \int_r f(\vec{r}) e^{-i\vec{k}\cdot\vec{r}} = \int_r \xi_\delta(r) e^{-i\vec{k}\cdot\vec{r}} = P_\delta(k), \quad (5.15)$$

one can derive the form of the eigenvalue spectrum. Given the presence of the δ distribution in the integral (5.14), the form of the power spectrum determines the analyticity of the eigenvalue distribution. Indeed, one can express it with a famous property of the δ distribution

$$\delta(f(x)) = \sum_i \frac{\delta(x - x_i)}{|f'(x_i)|}, \quad (5.16)$$

summing over all the x_i that are zeros of the function $f(x)$. In particular, we will show that for the Gaussian and the exponential correlation function, an analytic expression for $\varrho(\lambda)$ can be found.

Gaussian power spectrum

We show, for this case, that we can derive the analytical expression of the eigenvalue distribution under the assumption of a high-density limit for an ERM. We assume a Gaussian power spectrum as in Eq. (5.1).

$$f_0(k) = (2\pi)^{\frac{3}{2}} a^3 e^{-\frac{k^2 a^2}{2}}, \quad (5.17)$$

one obtains [18]

$$\varrho(\lambda) = \frac{1}{\sqrt{2\pi\rho a^3} \lambda} \sqrt{\ln \frac{(2\pi)^{\frac{3}{2}} \rho a^3}{\lambda}}. \quad (5.18)$$

Let us see how can easily recover the previous result with a simple calculation. Starting from plugging the expression of $P_\delta(k)$ in Eq. (5.14) one obtains

$$\varrho(\lambda) = \frac{1}{\rho} \int \frac{d^3 \vec{k}}{(2\pi)^3} \delta \left[\lambda - (2\pi)^{3/2} \rho a^3 e^{-\frac{k^2 a^2}{2}} \right]. \quad (5.19)$$

Since the power spectrum depends only on the absolute value of k , we can go in spherical coordinates, using the fact that $d^3 \vec{k} = 4\pi k^2 dk$

$$\varrho(\lambda) = \frac{4\pi}{\rho(2\pi)^3} \int_0^{+\infty} dk k^2 \delta \left[\lambda - (2\pi)^{\frac{3}{2}} \rho a^3 e^{-\frac{k^2 a^2}{2}} \right]. \quad (5.20)$$

At this point, we can make use of the property (5.16) where in our case, the function f can be identified by

$$\begin{aligned} f(k) &= \lambda - (2\pi)^{\frac{3}{2}} \rho a^3 e^{-\frac{k^2 a^2}{2}}, \\ f'(k) &= (2\pi)^{\frac{3}{2}} \rho a^5 k e^{-\frac{k^2 a^2}{2}}, \end{aligned} \quad (5.21)$$

with zeros of $f(k)$ in

$$k_i^* = \pm \frac{\sqrt{2}}{a} \sqrt{\ln \frac{(2\pi)^{\frac{3}{2}} \rho a^3}{\lambda}}. \quad (5.22)$$

We also need to evaluate the derivative at the zeros

$$f'(k_i^*) = \pm \frac{\sqrt{2}}{a} \sqrt{\ln \frac{(2\pi)^{\frac{3}{2}} \rho a^3}{\lambda}} a^2 (2\pi)^{\frac{3}{2}} \rho a^3 \frac{\lambda}{(2\pi)^{\frac{3}{2}} \rho a^3} = \pm \sqrt{2} a \lambda \sqrt{\ln \frac{(2\pi)^{\frac{3}{2}} \rho a^3}{\lambda}}. \quad (5.23)$$

but since we only care about the square of k_i^* for the numerator and its absolute value for the denominator, we only consider it to have one zero. Plugging everything in Eq. (5.20)

$$\begin{aligned} \varrho(\lambda) &= \frac{1}{2\pi^2 \rho} \int_0^{+\infty} dk k^2 \frac{\delta \left[k - \frac{\sqrt{2}}{a} \sqrt{\ln \frac{(2\pi)^{\frac{3}{2}} \rho a^3}{\lambda}} \right]}{\left| \sqrt{2} a \lambda \sqrt{\ln \frac{(2\pi)^{\frac{3}{2}} \rho a^3}{\lambda}} \right|} \\ &= \frac{1}{2\pi^2 \rho} \frac{\left(\frac{\sqrt{2}}{a} \sqrt{\ln \frac{(2\pi)^{\frac{3}{2}} \rho a^3}{\lambda}} \right)^2}{\sqrt{2} a \lambda \sqrt{\ln \frac{(2\pi)^{\frac{3}{2}} \rho a^3}{\lambda}}} \\ &= \frac{1}{\sqrt{2} \pi^2 \rho a^3 \lambda} \sqrt{\ln \frac{(2\pi)^{\frac{3}{2}} \rho a^3}{\lambda}}, \end{aligned} \quad (5.24)$$

with $\lambda \in (0, (2\pi)^{3/2} \rho a^3]$. We compare this equation with the results of numerical simulations in the left panel of Fig. 5.4.

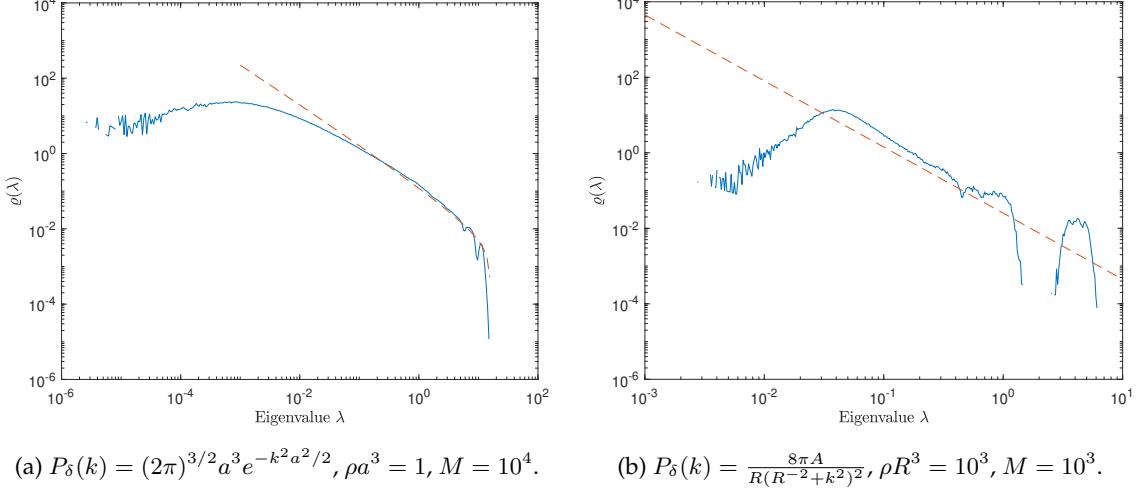


Figure 5.4: Eigenvalue density of 100×100 ERM with Gaussian and exponential elements at high density. Analytical approximation (dashed line) (5.24) in Fig. (a) and (5.27) in Fig. (b), is compared to numerical diagonalization (solid line).

Exponential power spectrum

Making use of the power spectrum and the correlation function defined in Eq. (5.3), we can follow the same steps as in the last section to recover the high-density limit of the eigenvalue distribution. Furthermore, the form of the power spectrum can be inverted analytically, allowing us to obtain an analytical expression for the distribution $\varrho(\lambda)$. Equipped with

$$P_\delta(k) = \frac{8\pi A}{R(R^{-2} + k^2)^2}, \quad P'_\delta(k) = -\frac{32\pi Ak}{R(R^{-2} + k^2)^3}, \quad (5.25)$$

one can solve the equation $\lambda = \rho P_\delta(k)$ with solutions

$$k^* = \pm \sqrt{-R^{-2} + \sqrt{\frac{8\pi A \rho}{R \lambda}}}, \quad (5.26)$$

where we have noticed that we can only consider real solutions and that we still care only about the absolute value of k^* , so that the eigenvalues distribution becomes

$$\begin{aligned} \varrho(\lambda) &= \frac{1}{2\pi^2 \rho} \frac{k^{*2}}{|\rho P'_\delta(k^*)|} \\ &= \frac{R}{64\pi^3 A \rho^2} \sqrt{-R^{-2} + \sqrt{\frac{8\pi A \rho}{R \lambda}}} \left(\frac{8\pi A \rho}{R \lambda} \right)^{\frac{3}{2}} \\ &= \frac{1}{8\pi^2 \rho \lambda} \sqrt{\frac{8\pi A \rho}{R \lambda}} \sqrt{-R^{-2} + \sqrt{\frac{8\pi A \rho}{R \lambda}}}, \end{aligned} \quad (5.27)$$

with $\lambda \in (0, 8\pi A \rho R^3]$. The numerical comparison is presented in the right panel of Fig. 5.4.

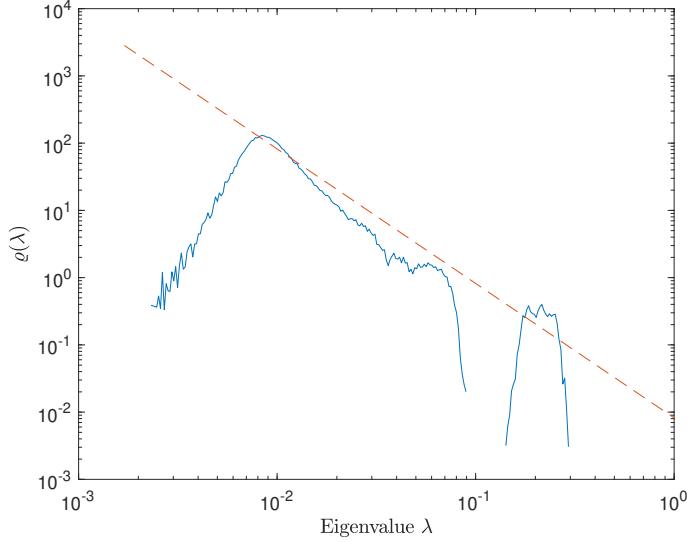


Figure 5.5: Eigenvalue density of 100×100 ERM $C_{\delta\delta}$ with power spectrum from the CDM fit $P_{\delta}(k) = \frac{Ak}{(1+k/k_0)^4}$. Theoretical approximation (dashed line) from the numerical resolution of Eq. (5.29) is compared to numerical diagonalization (solid line) of $M = 10^3$ matrices at $\rho = 10^{15}$ number density.

CDM model

Let us consider now the power spectrum for a Cold Dark Matter model. Using the fit provided by Bardeen et al. in [7]

$$P_{\delta}(k) = \frac{Ak}{\left(1 + \frac{k}{k_0}\right)^4}, \quad P'_{\delta}(k) = A \frac{\left(1 - 3\frac{k}{k_0}\right)}{\left(1 + \frac{k}{k_0}\right)^5}, \quad (5.28)$$

with the numerical value of the amplitude $A = 5.69 \text{ Mpc}^4 \text{h}^{-4}$ and the scale $k_0 = 0.0580 \text{ h Mpc}^{-1}$. In this case, it is not possible to obtain any analytical expression for the eigenvalue distribution as the power spectrum is not analytically invertible in simple terms of basic functions, but we can still numerically evaluate

$$\varrho(\lambda) = \frac{1}{\rho} \int_k \delta [\lambda - \rho P_{\delta}(k)], \quad (5.29)$$

yielding

$$\varrho(\lambda) = \frac{1}{2\pi^2\rho} \int_0^{+\infty} dk k^2 \sum_{k^*} \frac{\delta(k - k^*)}{|\rho P'_{\delta}(k^*)|} = \frac{1}{2\pi^2\rho^2} \sum_{k^*} \frac{k^{*2}}{|P'_{\delta}(k^*)|}, \quad (5.30)$$

with $\lambda \in \left(0, \rho A k_0 \frac{3^3}{4^4}\right)$. Where k^* are the zeros of the equation $P_{\delta}(k) = \frac{\lambda}{\rho}$, and must be evaluated numerically. Since $P_{\delta}(k)$ is not monotonic, there could be two zeros for some values of λ . Furthermore, since $P_{\delta,\max}(k) = P_{\delta}\left(\frac{k_0}{3}\right) = Ak_0 \frac{3^3}{4^4}$, we can read the maximum value λ_{\max} for which $P_{\delta,\max}(k) = \frac{\lambda_{\max}}{\rho}$ has solutions, giving the range in which λ can exist. The results are shown in Fig. 5.5.

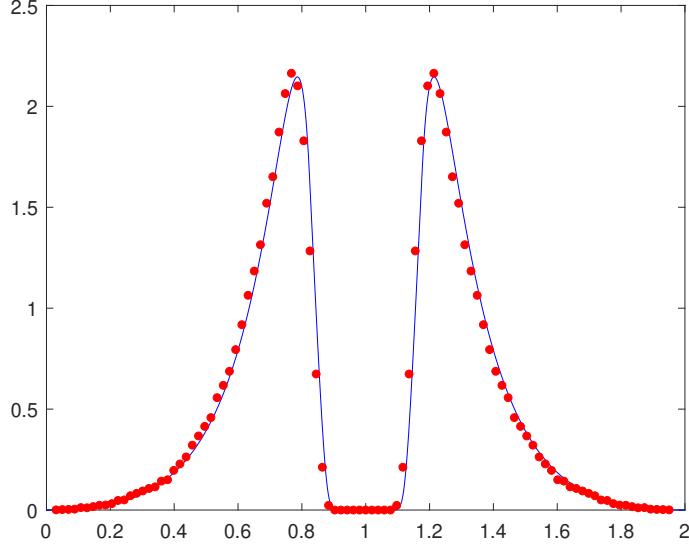


Figure 5.6: Numerical diagonalization (dots) of 10^5 samples of the 2×2 matrix (5.33), with number density $\rho = 2 \cdot 10^6$, compared with the analytical expression (5.37) (solid blue line).

5.3 Eigenvalues distribution of \mathbf{C}_{pp}

Let us come to the momentum-momentum correlation matrix. In its full glory, it is a $3N \times 3N$ matrix, containing the correlation information for N particles.

$$\mathbf{C}_{pp} = C_{p_i p_j} \otimes E_{ij}, \quad C_{p_i p_j} = -a_1(r)\mathbb{1}_3 - a_2(r)\pi_{\parallel}. \quad (5.31)$$

It has the structure as in Eq. (3.41). As in there, it is not possible to use the ERM machinery, but it is nonetheless possible to make some considerations. At first, we will get the analytic expression of $\varrho(\lambda)$ for two particles, and then we will see under which condition we can use ERM theory for \mathbf{C}_{pp} . In the end, we will consider the \mathbf{C}_{pp} in its extended form. For all the considerations made in this section, is assumed the form of $a_1(r)$ and $a_2(r)$ of the CDM model to be as in Eq. (5.8)

5.3.1 Two particles

At first, we consider the simplest case possible, namely with only two particles, neglecting the $a_2(r)$ contribution and rescaling the matrix. The limit in which $a_2(r)$ can be safely neglected corresponds to the limit of high density, as we can notice from Fig. 5.11a, under a certain distance, the a_2 contribution is negligible. Furthermore a block matrix of the form

$$\mathbf{C}_{pp,2} = \begin{pmatrix} \mathbb{1}_3 & -a_1(r)\mathbb{1}_3 \\ -a_1(r)\mathbb{1}_3 & \mathbb{1}_3 \end{pmatrix}, \quad \lambda(r) = (1 \pm a_1(r)), \quad (5.32)$$

has three times the same eigenvalue of the rescaled version (5.33). In this case, the resulting eigenvalue distribution can simply be derived by transforming the probability distribution of a distance between two points in a volume, using Eq. (2.15). This is because the resulting matrix

and relative eigenvalues are of the form

$$C_{pp,2} = \begin{pmatrix} 1 & -a_1(r) \\ -a_1(r) & 1 \end{pmatrix}, \quad \lambda_{1,2}(r) = 1 \pm a_1(r), \quad (5.33)$$

The probability distribution of the distance between two random points in a box of length $L = 1$ has been computed in [51], it reads

$$\varrho_{\text{CLP}}(l) = \begin{cases} -l^2 [(l-8)l^2 + \pi(6l-4)] & 0 \leq l \leq 1; \\ 2l [(l^2 - 8\sqrt{l^2-1} + 3)l^2 - 4\sqrt{l^2-1} + 12l^2 \sec^{-1} l + \pi(3-4l) - \frac{1}{2}] & 1 < l \leq \sqrt{2}; \\ l [(1+l^2)(6\pi + 8\sqrt{l^2-2} - 5 - l^2) - 16l \csc^{-1}(\sqrt{2-2l^2}) + \\ + 16l \tan^{-1}(l\sqrt{l^2-2}) - 24(l^2+1) \tan^{-1}(\sqrt{l^2-2})] & \sqrt{2} < l \leq \sqrt{3}. \end{cases} \quad (5.34)$$

Transforming this probability distribution using Eq. (2.15), through the correlation function $a_1(r)$, one obtains the analytical distribution of the eigenvalues for a $2 \times 2 C_{pp}$ matrix. Considering the first correlation function $a_1(r)$ defined in Eq. (5.8), for example, we can write down the explicit analytic formula of $\varrho(\lambda)$ in the simple case $N = 2$, using Eq. (2.15). First of all we need to invert $\lambda_{1,2}(r)$

$$r_{1,2}(y = \lambda_{1,2}) = -\frac{\beta_1}{2} + \frac{\sqrt{\beta_1 [4\alpha_1 (\pm \frac{\sigma_1^2}{3} - y + 1) + \beta_1(y-1)]}}{2\sqrt{y-1}}. \quad (5.35)$$

Then considering its derivative

$$\begin{aligned} r'_{1,2}(y = \lambda_{1,2}) &= \mp \frac{\sigma_1^2}{3} \frac{\alpha_1 \sqrt{\beta_1}}{(y-1)^{3/2} \sqrt{4\alpha_1 (\pm \frac{\sigma_1^2}{3} - y + 1) + \beta_1(y-1)}} \\ &= \mp \frac{\sigma_1^2}{3} \frac{\alpha_1 \beta_1}{2(y-1)^2 (r(\lambda_{1,2}) + \frac{\beta_1}{2})}, \end{aligned} \quad (5.36)$$

we have everything to compute $\varrho(\lambda_{1,2})$

$$\varrho(y = \lambda_{1,2}) = -\varrho_{\text{CLP},i} \left(\frac{r_{1,2}(y)}{R} \right) \frac{r'_{1,2}(y)}{R}, \quad (5.37)$$

in each of the three intervals i of Eq. (5.34). In this way, we transformed $\varrho_{\text{CLP}}(r) \rightarrow \varrho(\lambda)$, using basic probability knowledge, and the results are shown in Fig. 5.6.

Considering again the $a_2(r)$ contribution, the eigenvalues of $C_{pp,2}$ of the form of Eq. (3.42) are

$$\lambda_{1,2,5,6} = 1 \pm a_1(r), \quad \lambda_{3,4} = 1 \pm a_1(r) \pm a_2(r), \quad (5.38)$$

where $\lambda_i < \lambda_j$ if $i < j$. The distribution for different densities is shown in Fig. 5.7.

It consists of two symmetric peaks, one is relative to the contribution of $a_1(r)$, namely the one that is further away from the symmetry axis $\lambda = 1$. The other peak is given by the combined contribution of $a_1(r)$ and $a_2(r)$ and is the most internal one. As anticipated, as long as we increase the density, the internal peak relative to the eigenvalues $\lambda_{3,4}$ gets closer and closer to the first and

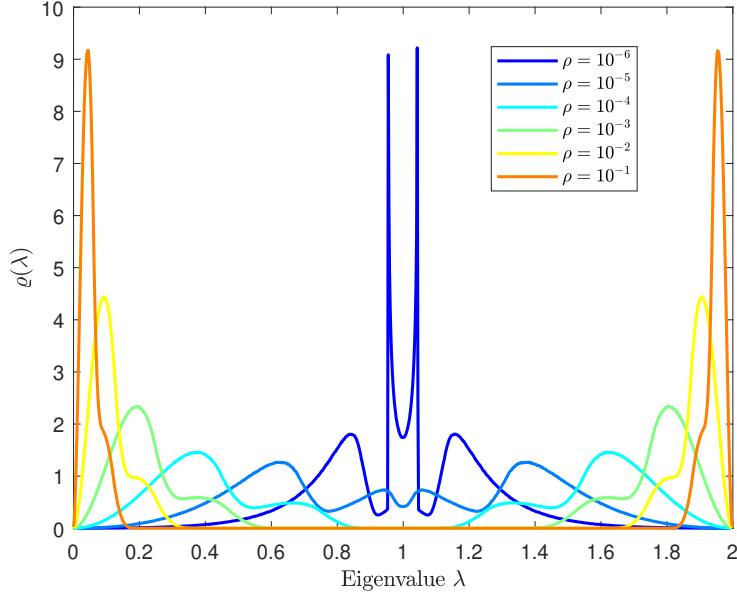


Figure 5.7: Eigenvalue spectrum for the momentum-momentum covariance matrix \mathbf{C}_{pp} with two particles, obtained with 10^7 matrices.

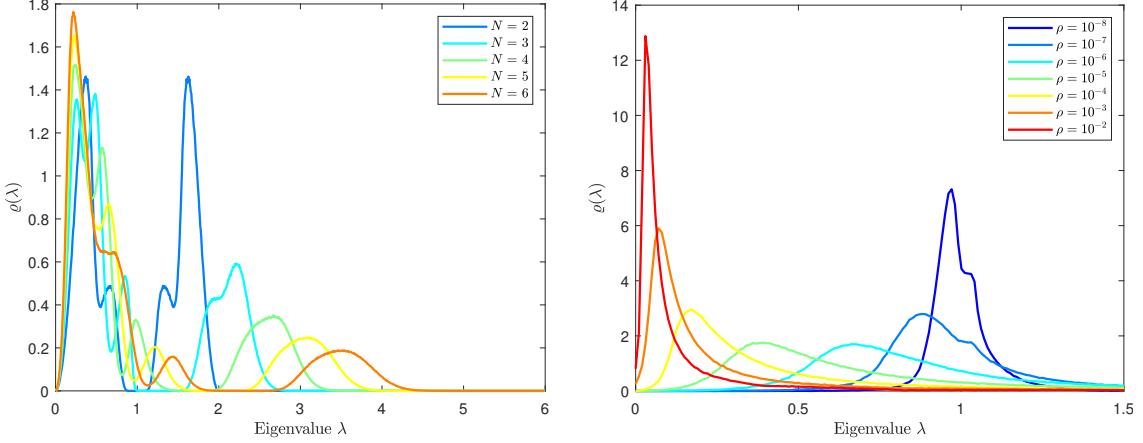
so the six eigenvalues can be considered to be $\lambda_{1,\dots,6} = 1 \pm a_1(r)$. As expected in the high-density limit, the eigenvalues distribution tends to a Dirac distribution in $\lambda = 0$ and $\lambda = 2$, as $\lim_{r \rightarrow 0} a_1(r) = 1$ if we set $\sigma_1^2/3$ to unity. While in the low density limit, since $\lim_{r \rightarrow +\infty} a_1(r) = 0$ the peak is at $\lambda = 1$. This behavior will be respected also when we will consider N particles.

5.3.2 N particles

When we consider N particles, hence with a $3N \times 3N$ correlation matrix, what we can observe studying the behavior of the eigenvalue distribution, is that of the four peaks in the case of $N = 2$ particles, only one survives, see Fig. (5.8a). In contrast, the others get suppressed more and more as N increases. We can also observe that if N is large enough, then the spectrum does not depend on N anymore, as we discussed in the last section of the previous chapter, and the distribution converges to a limit function. In the limit of large N , the spectrum consists only of one peak with a small fluctuation given by the presence of $a_2(r)$ if the density is low enough. While for high densities, we can not notice any difference between the case where we consider $a_2(r)$ and the one where we do not, see Fig. 5.8b.

High-density limit

In this section, we will use ERM theory to obtain an analytical form of the eigenvalue distribution in the high-density limit. In order to do that, we first need to neglect $a_2(r)$ in the expression of \mathbf{C}_{pp} , and we just showed that this approximation can be safely carried on in the limit of high-density. Before proceeding we thus need to show that the difference of the eigenvalue distributions of the



(a) Behavior of small matrices at constant number density $\rho = 10^{-4}$. Obtained with numerical diagonalization of $M = 10^6$ matrices. Is noticeable the rise of the left peak.

(b) Plots of 10^4 samples of $300 \times 300 \mathbf{C}_{pp}$ matrices at different densities. The limit regime pushes the distribution to $\lambda = 0$ for the high-density limit and to $\lambda = 1$ for the low-density.

Figure 5.8: Numerical evaluation of eigenvalue distribution of \mathbf{C}_{pp} matrices.

matrices with and without $a_2(r)$ are negligible, see Fig. 5.9. As we can see from the plot, the contribution of $a_2(r)$ in the high-density regime is negligible in the main body of the distribution, while the approximation does not hold for the tails. We can thus rescale our \mathbf{C}_{pp} matrix into an ERM in order to find an expression that could hold for the body. At this point is convenient to use the high-density expansion introduced in Sec. 4.4.1 to derive an expression for the eigenvalue distribution in this limit. Differently from the density correlation matrix case studied in Sec. 5.2.1, in this case, is not possible to obtain any analytical expression for the eigenvalues distribution, but we can still numerically evaluate it. In this case, to compute the Fourier transform of the defining function of \mathbf{C}_{pp} , $f(r) = -a_1(r)$, we need a little trick. Namely, we first need to consider the absolute momentum correlation matrix $C_{|p_i||p_j|} = -3a_1(r) - a_2(r)$

$$\begin{aligned}
C_{|p_i||p_j|}(r) &= \langle |\vec{p}_i| \otimes |\vec{p}_j| \rangle = \left\langle \left| \vec{\nabla} \psi_i \right| \left| \vec{\nabla} \psi_j \right| \right\rangle \\
&= \int_k \int_{k'} \left| i\vec{k} \right| \left| i\vec{k}' \right| \langle \tilde{\psi}_i(k) \tilde{\psi}_j(k') \rangle e^{i(\vec{k} \cdot \vec{x}_i + \vec{k}' \cdot \vec{x}_j)} \\
&= \int_k k^2 P_\psi(k) e^{i\vec{k} \cdot \vec{r}} \\
&= \int_k \frac{P_\delta(k)}{k^2} e^{i\vec{k} \cdot \vec{r}} \\
&= \frac{1}{2\pi^2} \int_0^{+\infty} P_\delta(k) j_0(kr) \\
&= -3a_1(r) - a_2(r).
\end{aligned} \tag{5.39}$$

But we already noticed that the contribution given by $a_2(r)$ is negligible in the limit of high density, therefore we have the following relation

$$C_{|p_i||p_j|}^\infty(r) = -3a_1(r) = \int_k \frac{P_\delta(k)}{k^2} e^{i\vec{k} \cdot \vec{r}}, \tag{5.40}$$

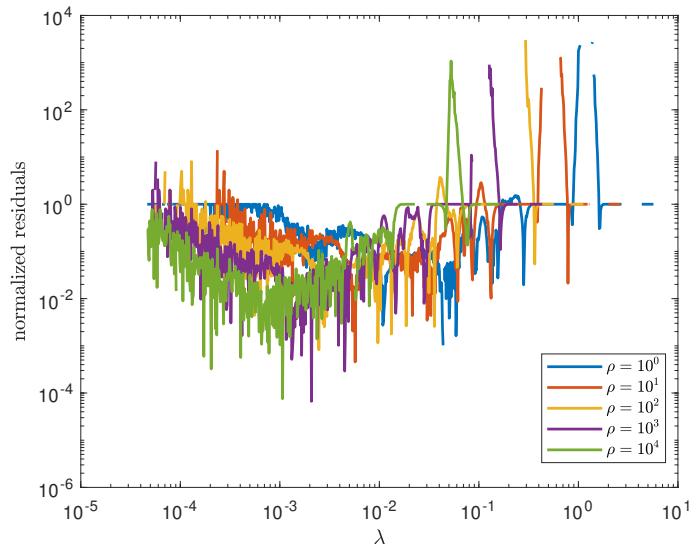


Figure 5.9: Comparison between the eigenvalue distribution of \mathbf{C}_{pp} with $a_2(r)$, and without.

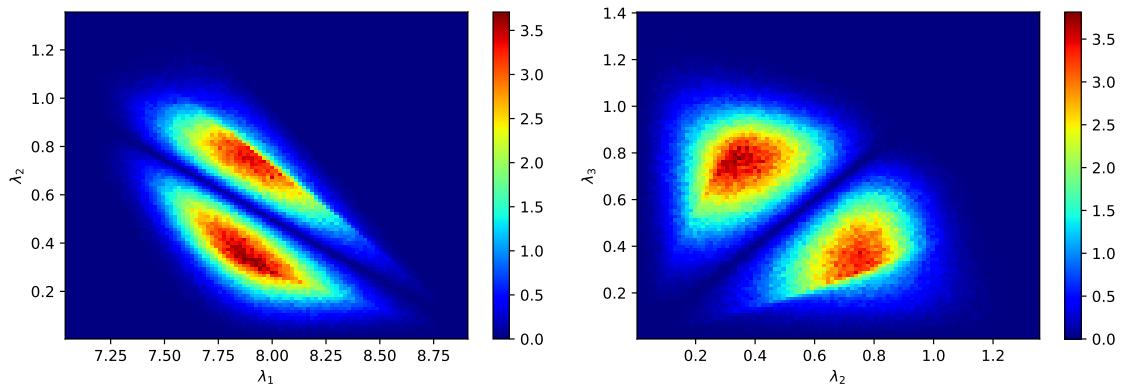


Figure 5.10: Sample counts of a $3 \times 3 C_{|p||p|}$ matrix, with 10^6 samples and $a_1(r), a_2(r)$ of the form of Eq. (5.8).

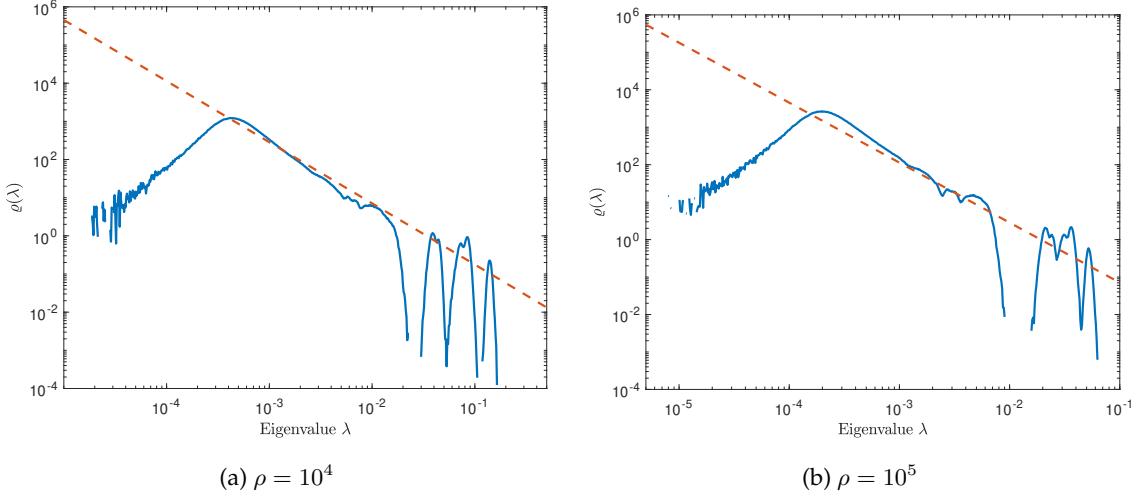


Figure 5.11: Eigenvalue density of 10^4 samples of the $3N \times 3N$ C_{pp} matrix at high density. The approximated Eq.(5.42) (dashed line) is compared to numerical diagonalization (solid line).

anti-Fourier transforming this, and dividing by 3, one obtains

$$f_0(k) = \int_r f(\vec{r}) e^{-i\vec{k} \cdot \vec{r}} = \int_r -a_1(r) e^{-i\vec{k} \cdot \vec{r}} = \frac{P_\delta(k)}{3k^2}. \quad (5.41)$$

In this way the high-density equation for the eigenvalue distribution becomes

$$\begin{aligned} \varrho(\lambda) &= \frac{1}{\rho} \int_k \delta \left[\lambda - \rho \frac{P_\delta(k)}{3k^2} \right] \\ &= \frac{1}{2\pi^2 \rho^2} \int dk k^2 \frac{\delta(k - k^*)}{\left| \left(\frac{P_\delta(k)}{3k^2} \right)' \right|_{k^*}} \\ &= \frac{3}{2\pi^2 \rho^2} \frac{k^{*5}}{|k^* P'_\delta(k^*) - 2P_\delta(k^*)|}, \end{aligned} \quad (5.42)$$

where now k^* are the zeros of the equations $\frac{P_\delta(k)}{k^2} = \frac{3\lambda}{\rho}$, and must be evaluated numerically. Since the left-hand side of the last equation is a monotonous decreasing function of k , we know that only one solution satisfies the equation. The results in Fig. 5.11 tell us that the approximation of high-density limit manages to describe the eigenvalue distribution quite well in a range of $\rho = 10^2 - 10^6$. Unfortunately, it is a good approximation only in the decreasing area of the plot. Furthermore, this approximation fails to describe the islands that appear in correspondence of the higher eigenvalues.

Low-density limit

Another interesting application is the low-density limit, as RMT provides for this case a closed analytical formula, depending only on the defining function of the ERM and the number density. Specializing Eq. (4.33) in $d = 3$ dimensions, we have

$$\varrho(\lambda) = 2\pi\rho e^{-4\pi/3\rho \{f^{-1}(|1-\lambda|)\}^3} f^{-1}(|1-\lambda|)^2 (f^{-1}(|1-\lambda|))', \quad 0 < \lambda < 2. \quad (5.43)$$

We apply this approximation to the C_{pp} correlation matrix, neglecting $a_2(r)$.

$$C_{p_i p_j}(r) = f(r) = -a_1(|\vec{r}_i - \vec{r}_j|). \quad (5.44)$$

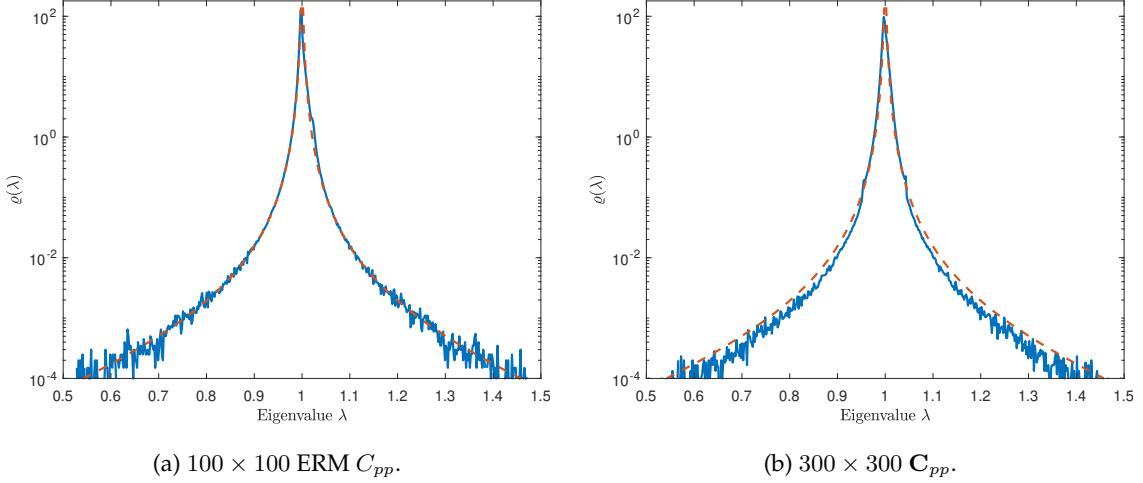


Figure 5.12: Low-density eigenvalues distribution with CDM power spectrum and momentum correlation functions. Analytical approximation (dashed line) (5.43) is compared to numerical diagonalization (solid line) of $M = 10^5$ matrices at $\rho = 10^{-10}$ number density.

since we have an analytical fit for $a_1(r)$ and hence an analytical expression for $a_1^{-1}(r)$

$$\begin{aligned} f^{-1}(y) &= -\frac{\beta_1}{2} + \frac{\sqrt{\beta_1 y (\beta_1 y - 4\alpha_1(y-1))}}{2y}, \\ (f^{-1})'(y) &= -\frac{\alpha_1 \beta_1}{y \sqrt{\beta_1 y (\beta_1 y - 4\alpha_1(y-1))}}, \end{aligned} \quad (5.45)$$

The results are shown in Fig. 5.12.

5.3.3 Individual probability distribution

Another different property that is interesting to calculate and study is the pdf of the individual eigenvalues $\varrho(\lambda_i)$. In particular, could be possible to fit them with a Gaussian distribution. Let us consider for this case both the original momentum covariance matrix \mathbf{C}_{pp} and the ERM $C_{|p||p|}$. A good match is observed for \mathbf{C}_{pp} as the eigenvalues are not chosen from the extrema of the spectrum, see Fig. 5.13, while for $C_{|p||p|}$ a good match is obtained for all the spectrum, see Fig. 5.14.

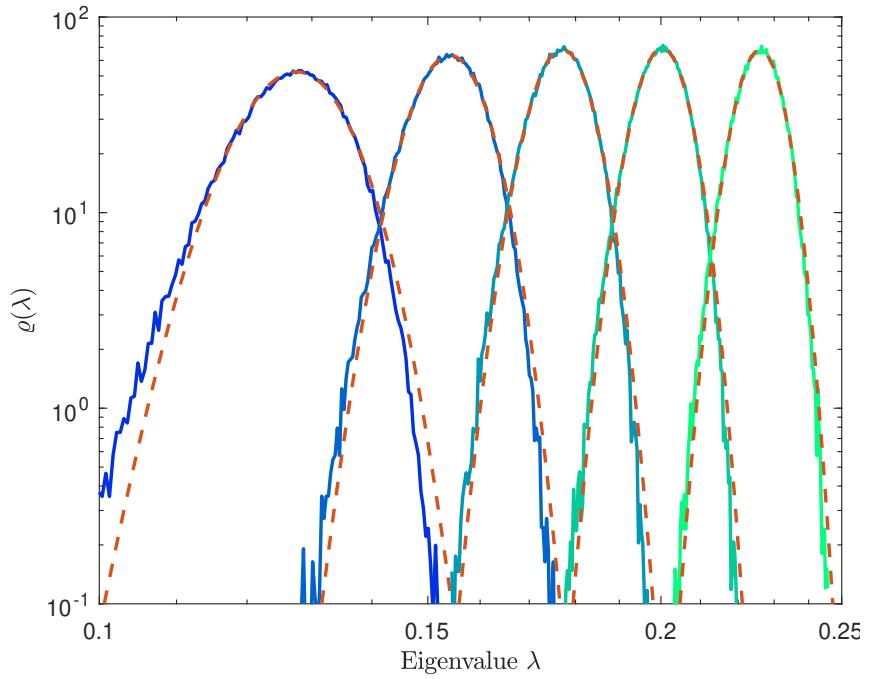


Figure 5.13: Some of the single eigenvalues distribution for the C_{pp} matrix, obtained from 10^5 samples of the 300×300 C_{pp} of $\rho = 10^{-4}$ number density, compared with a Gaussian fit.

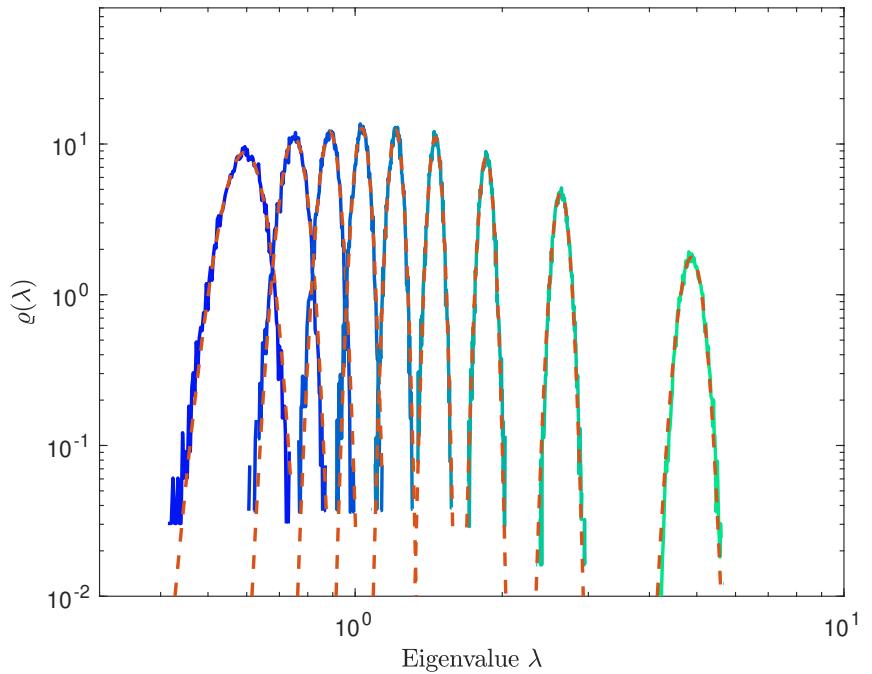


Figure 5.14: Some of the single eigenvalues distribution for the $C_{|p||p|}$ ERM, obtained from 10^4 samples of the 100×100 matrix of $\rho = 10^{-4}$ number density, compared with a Gaussian fit.

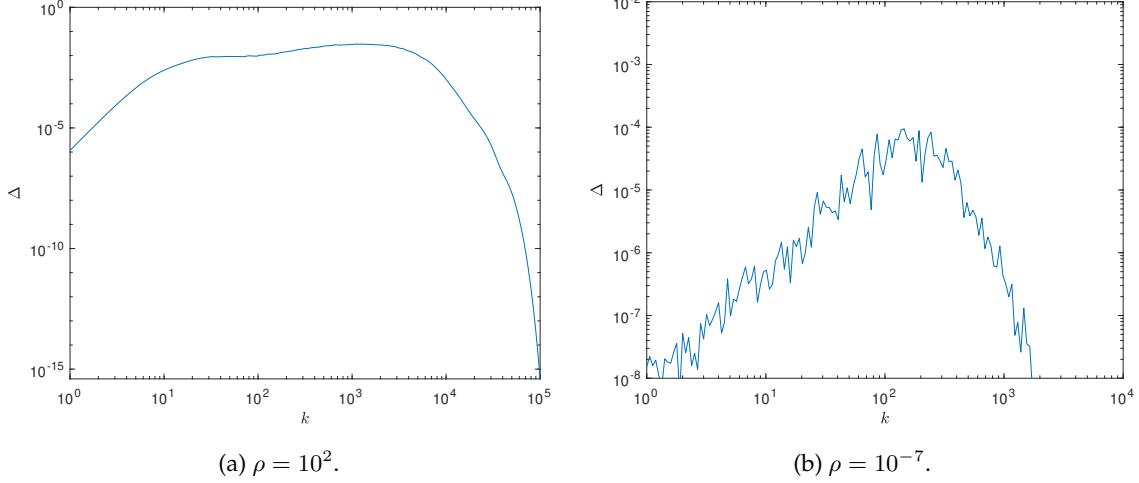


Figure 5.15: Comparison of the quantity $\Delta = \langle e^{-1/2\mathbf{k}^\top \mathbf{C}_{pp}\mathbf{k} - i\mathbf{k}\mathbf{q}} \rangle - \langle e^{-1/2\mathbf{k}^\top \mathbf{C}_{pp}\mathbf{k}} \rangle \langle e^{-i\mathbf{k}\mathbf{q}} \rangle$ at different densities, with 10^3 samples of the $3N \times 3N$ \mathbf{C}_{pp} matrix, with $N = 200$.

5.4 Generating functional

Recall that our generating functional reads

$$Z_0[\mathbf{L}] = V^{-N} \int d\mathbf{q} \exp \left(-\frac{1}{2} \mathbf{L}_p^\top \mathbf{C}_{pp} \mathbf{L}_p + i \mathbf{L}_q \mathbf{q} \right). \quad (5.46)$$

In the easiest case, the integrand function could be written as

$$\exp \left(-\frac{t^2}{2} \mathbf{k}^\top \mathbf{C}_{pp} \mathbf{k} + i \mathbf{k} \mathbf{q} \right), \quad (5.47)$$

where \mathbf{C}_{pp} is diagonalizable through a rotation

$$\exp \left(-\frac{t^2}{2} \mathbf{k}^\top \mathcal{O}^{-1} \mathcal{O} \mathbf{C}_{pp} \mathcal{O}^{-1} \mathcal{O} \mathbf{k} + i \mathbf{k} \mathcal{O}^{-1} \mathcal{O} \mathbf{q} \right) = \exp \left(-\frac{t^2}{2} \mathbf{k}'^\top \mathbf{C}'_{pp} \mathbf{k}' + i \mathbf{k}' \mathbf{q}' \right). \quad (5.48)$$

Where the primed quantities have been rotated, the \mathbf{C}_{pp} matrix is now diagonal $\mathbf{C}_{pp} = \text{diag}(\lambda_1, \dots, \lambda_{3N})$. From now on, for simplicity, it will be written without the primes anyway. The first question we are going to investigate is whether the average of the integrand function of the generating functional is factorizable or not, namely if

$$\left\langle \exp \left(-\frac{t^2}{2} \mathbf{k}^\top \mathbf{C}_{pp} \mathbf{k} + i \mathbf{k} \mathbf{q} \right) \right\rangle = \left\langle \exp \left(-\frac{t^2}{2} \mathbf{k}^\top \mathbf{C}_{pp} \mathbf{k} \right) \right\rangle \langle \exp(i \mathbf{k} \mathbf{q}) \rangle. \quad (5.49)$$

This question has been investigated numerically and the answer is given in Fig. 5.15. We can notice that for lower density the oscillations are way more important than in the higher density cases.

5.4.1 Moments expansion

Since we can use Eq. (5.49), we are now interested in calculating

$$\left\langle \exp \left(-\frac{t^2}{2} \mathbf{k}^\top \mathbf{C}_{pp} \mathbf{k} \right) \right\rangle = \left\langle \exp \left(-\frac{t^2}{2} \sum_{i=1}^{3N} \mathbf{k}_i^2 \lambda_i \right) \right\rangle, \quad (5.50)$$

related to the eigenvalues distribution of the \mathbf{C}_{pp} matrix. We first use the definition of mean value

$$\left\langle \exp\left(-\frac{t^2}{2} \sum_{i=1}^{3N} \mathbf{k}_i^2 \lambda_i\right) \right\rangle = \int d\lambda_1 \dots d\lambda_{3N} \varrho(\lambda_1, \dots, \lambda_{3N}) \exp\left(-\frac{t^2}{2} \sum_{i=1}^{3N} \mathbf{k}_i^2 \lambda_i\right). \quad (5.51)$$

then we make use of power series expansion for the exponential

$$\exp\left(-\frac{t^2}{2} \sum_{i=1}^{3N} \mathbf{k}_i^2 \lambda_i\right) = \sum_{n=0}^{+\infty} \frac{\left(-\frac{t^2}{2} \sum_{i=1}^{3N} \mathbf{k}_i^2 \lambda_i\right)^n}{n!}. \quad (5.52)$$

As the last step, we need the n -th order expansion for a polynomial with m terms, namely the multinomial theorem

$$(x_1 + \dots + x_m)^n = \sum_{\alpha} \frac{n!}{j_1! \dots j_m!} \prod_{t=1}^m x_t^{j_t}. \quad (5.53)$$

where with α we intend all the configuration for which $\sum_{i=1}^m j_i = n$. In this way, we can easily rewrite Eq. (5.52) as

$$\sum_{n=0}^{+\infty} \left(-\frac{t^2}{2}\right)^n \frac{\left(\sum_{i=1}^{3N} \mathbf{k}_i^2 \lambda_i\right)^n}{n!} = \sum_{n=0}^{+\infty} \left(-\frac{t^2}{2}\right)^n \frac{1}{n!} \sum_{\alpha} \frac{n!}{j_1! \dots j_{3N}!} \prod_{t=1}^{3N} \mathbf{k}_i^{2j_t} \lambda_i^{j_t}. \quad (5.54)$$

Plugging it into Eq. (5.51) one obtains

$$\begin{aligned} \left\langle \exp\left(-\frac{t^2}{2} \mathbf{k}^\top \mathbf{C}_{pp} \mathbf{k}\right) \right\rangle &= \int d\lambda_1 \dots d\lambda_{3N} \varrho(\lambda_1, \dots, \lambda_{3N}) \exp\left(-\frac{t^2}{2} \sum_{i=1}^{3N} \mathbf{k}_i^2 \lambda_i\right) \\ &= \int d\lambda_1 \dots d\lambda_{3N} \varrho(\lambda_1, \dots, \lambda_{3N}) \sum_{n=0}^{+\infty} \left(-\frac{t^2}{2}\right)^n \sum_{\alpha} \frac{1}{j_1! \dots j_{3N}!} \prod_{t=1}^{3N} \mathbf{k}_i^{2j_t} \lambda_i^{j_t} \\ &= \sum_{n=0}^{+\infty} \left(-\frac{t^2}{2}\right)^n \sum_{\alpha} \frac{1}{j_1! \dots j_{3N}!} \left\langle \prod_{t=1}^{3N} \mathbf{k}_i^{2j_t} \lambda_i^{j_t} \right\rangle. \end{aligned} \quad (5.55)$$

The first three terms would be

$$\left\langle \exp\left(-\frac{t^2}{2} \mathbf{k}^\top \mathbf{C}_{pp} \mathbf{k}\right) \right\rangle \simeq 1 - \frac{t^2}{2} \sum_{i=1}^{3N} \mathbf{k}_i^2 \langle \lambda_i \rangle + \frac{t^4}{8} \sum_{i,j=1}^{3N} \mathbf{k}_i \mathbf{k}_j \langle \lambda_i \lambda_j \rangle, \quad (5.56)$$

and so on and so forth. Remember that an extra 2 factor is there, either because of the $j_i!$ coefficient that is equal to 2 if $i = j$ or because we have to divide by 2 every term with $i \neq j$ or would be counted twice. The moments are defined through

$$\begin{aligned} \langle \lambda_i \rangle &= \int d\lambda_i \varrho(\lambda_i) \lambda_i, \\ \langle \lambda_i \lambda_j \rangle &= \int d\lambda_i d\lambda_j \varrho(\lambda_i, \lambda_j) \lambda_i \lambda_j, \\ \langle \lambda^n \rangle &= \int d\lambda \varrho(\lambda) \lambda^n = \frac{1}{N} \langle \text{Tr } \mathbf{C}_{pp}^n \rangle, \end{aligned} \quad (5.57)$$

Chapter 6

Conclusions

This thesis investigates the eigenvalues distribution of the random momentum-momentum and density-density autocorrelation matrices that rise in KFT. It is possible to use ERM tools since these matrices are Euclidean Random Matrices, a special class of random matrices whose entries depend on a function of the relative distances between point particles sampled in a box. It is first checked that the eigenvalues distribution of the ERM does not depend on the size N of the matrix, in the limit of large N . Three different initial density power spectra $P_\delta(k)$ have been considered to define the ERM. This feature enabled two things, first, the possibility to use some of the RMT tools, such as the high and low-density expansion, and, second, making sure to study an ensemble of $N \gg 1$ particles even using limited sized matrices.

For the density-density autocorrelation matrix $C_{\delta\delta}$, since it is a proper ERM by definition, the whole ERM machinery has been used without any other assumptions. The threshold for the high and the low-density limits has been fixed, having investigated the eigenvalues spectrum. Those limits have been further studied resulting in analytical expressions for the eigenvalues distribution, that agree with the numerical diagonalization for the decaying part of the spectrum. This limit is particularly interesting as it describes the Universe at late cosmic time.

For the momentum-momentum autocorrelation matrix C_{pp} , which is not a proper ERM, is first given the analytic expression in the simple case $N = 2$ particles, when it is rescaled to an ERM matrix neglecting the contribution of the $a_2(r)$ momentum correlation function. Then it is justified why in the high-density limit is legitimate to neglect the $a_2(r)$ contribution, and an analytical expression of the eigenvalues distribution is given. It turns out that also the low-density limit describes quite well the eigenvalues distribution of C_{pp} even though in this case $a_2(r)$ is not negligible.

The single eigenvalues distribution of C_{pp} are considered. It turned out that, especially the central eigenvalues, could be fitted with a Gaussian.

In the end, is investigated how the known eigenvalues distribution can be used to simplify the calculation for the free generating functional. A possible way is to use the moment expansion so that the previous information about the eigenvalue distribution can now be applied. For this reason is tempting to try to find a relation between the mean of the eigenvalues and the density of the associated random matrix. Further developments of the project should include the high and low-density eigenvalue distribution in the expression of the generating functional, ...

Appendix A

Spherical Bessel functions

The spherical Bessel functions of the first kind $j_n(x)$, are one of the two linearly independent solutions, together with $y_n(x)$ of the radial part of the Helmholtz eigenvalue equation

$$\nabla^2 f = -k^2 f, \quad (\text{A.1})$$

that has expression

$$x^2 \frac{d^2 y}{dx^2} + 2x \frac{dy}{dx} + (x^2 - n(n+1))y = 0. \quad (\text{A.2})$$

They are related to the ordinary Bessel function $J_n(x)$ by

$$j_n(x) = \sqrt{\frac{\pi}{2x}} J_{n+\frac{1}{2}}(x). \quad (\text{A.3})$$

where n is an integer. While the ordinary Bessel functions of the first kind solve Bessel's differential equation

$$x^2 \frac{d^2 y}{dx^2} + x \frac{dy}{dx} + (x^2 - n^2)y = 0. \quad (\text{A.4})$$

Given the high presence of these kinds of functions throughout the thesis, it is useful to gather all the properties that have been used. The spherical Bessel functions arise when an integral in spherical coordinates of an imaginary exponential function must be carried on

$$\int_{-1}^{+1} d\mu \mu^n e^{ixn} = (-i\partial_x)^n \int_{-1}^{+1} d\mu e^{ixn} = 2(-i\partial_x)^n j_0(x). \quad (\text{A.5})$$

In general, they can be written as

$$\begin{aligned} j_n(x) &= (2x)^n \sum_{k=0}^{+\infty} \frac{(-1)^k (k+n)!}{k! (2(k+n)+1)!} z^{2k} \\ &= x^n \sum_{k=0}^{+\infty} \frac{(-1)^k}{k! (2(k+n)+1)!!} \left(\frac{z^2}{2}\right)^k \\ &= (-x)^n \left(\frac{1}{x} \frac{d}{dx}\right)^n \frac{\sin x}{x}, \end{aligned} \quad (\text{A.6})$$

such as the first Bessel functions are

$$j_0(x) = \frac{\sin x}{x}; \quad j_1(x) = \frac{\sin x}{x^2} - \frac{\cos x}{x}; \quad j_2(x) = \frac{(3-x^2)\sin x}{x^3} - 3\frac{\cos x}{x^2}. \quad (\text{A.7})$$

It is useful to define some recursion relations

$$j_{n-1}(x) = (2n+1) \frac{j_n(x)}{x} - j_{n+1}(x), \quad j'_n(z) = \frac{n}{x} j_n(x) - j_{n+1}(x). \quad (\text{A.8})$$

And even more useful is to specialize them for $j_0(x)$ and $j_1(x)$

$$j''_0(x) = -j'_1(x) = j_2(x) - \frac{j_1(x)}{x}, \quad j_0(x) = \frac{3j_1(x)}{x} - j_2(x). \quad (\text{A.9})$$

Appendix B

Gaussian integrals

One-dimensional

$$\int_{-\infty}^{+\infty} e^{-a(x+b)^2} dx = \sqrt{\frac{\pi}{a}}. \quad (\text{B.1})$$

$$\int_{-\infty}^{+\infty} e^{-ax^2+jx} dx = \sqrt{\frac{\pi}{a}} e^{\frac{j^2}{4a}}. \quad (\text{B.2})$$

$$\int_{-\infty}^{+\infty} e^{-ax^2+ijx} dx = \sqrt{\frac{\pi}{a}} e^{-\frac{j^2}{4a}}. \quad (\text{B.3})$$

Multi-dimensional

$$\int_x e^{-\frac{1}{2}\mathbf{x}^\top A\mathbf{x}} = \sqrt{\frac{(2\pi)^N}{\det A}}. \quad (\text{B.4})$$

$$\int_x e^{-\frac{1}{2}\mathbf{x}^\top A\mathbf{x} + j^\top \mathbf{x}} = \sqrt{\frac{(2\pi)^N}{\det A}} e^{\frac{1}{2}j^\top A^{-1}j}. \quad (\text{B.5})$$

$$\int_x e^{-\frac{1}{2}\mathbf{x}^\top A\mathbf{x} + ij^\top \mathbf{x}} = \sqrt{\frac{(2\pi)^N}{\det A}} e^{-\frac{1}{2}j^\top A^{-1}j}. \quad (\text{B.6})$$

Others

$$\begin{aligned} \langle e^{i\mathbf{t}\mathbf{x}} \rangle &= \int_x e^{i\mathbf{t}\mathbf{x}} \varrho(\mathbf{x}) \\ &= \frac{1}{\sqrt{(2\pi)^d \det \mathbf{C}}} \int_x e^{i\mathbf{t}\mathbf{x}} e^{-1/2\mathbf{x}^\top \mathbf{C}^{-1}\mathbf{x}} \\ &= e^{-\frac{1}{2}\mathbf{t}^\top \mathbf{C}\mathbf{t}}. \end{aligned} \quad (\text{B.7})$$

$$\begin{aligned} \int_0^{+\infty} dx x^n e^{-ax} &= (-\partial_a)^n \int_0^{+\infty} dx e^{-ax} \\ &= (-\partial_a)^n \left[\frac{e^{-ax}}{-a} \right]_0^{+\infty} \\ &= (-\partial_a)^n \left(\frac{1}{a} \right) \\ &= \frac{n!}{a^{n+1}}, \quad a > 0, n > -1. \end{aligned} \quad (\text{B.8})$$

Appendix C

Nearest neighbours

Let us consider a distribution of n points randomly distributed in a three-dimensional region of the space, with constant spatial density ρ . In order to find the probability distribution $p(\Delta r)$ for finding the nearest neighbor of a given point in a distance between Δr and $\Delta r + dr$, we need to first set the integral-differential equation that the distribution satisfies, namely

$$p(\Delta r)d\Delta r = \left(1 - \int_0^{\Delta r} d\Delta r' p(\Delta r')\right) 4\pi\Delta r^2\rho d\Delta r, \quad (\text{C.1})$$

that is the probability of not finding it closer than Δr times the probability of finding it in a shell of radius Δr and width $d\Delta r$. Doing some algebra we arrive at the differential equation

$$\begin{aligned} \frac{p(\Delta r)}{4\pi\Delta r^2\rho} &= 1 - \int_0^{\Delta r} d\Delta r' p(\Delta r'), \\ \frac{d}{d\Delta r} \left(\frac{p(\Delta r)}{4\pi\Delta r^2\rho} \right) &= -p(\Delta r), \end{aligned} \quad (\text{C.2})$$

that can be easily solved

$$\begin{aligned} \frac{p'(\Delta r)\Delta r^2 - 2p(\Delta r)\Delta r}{\Delta r^4} &= -4\pi\rho p(\Delta r), \\ \frac{p'(\Delta r)}{p(\Delta r)} &= \frac{2}{\Delta r} - 4\pi\rho\Delta r^2, \end{aligned} \quad (\text{C.3})$$

which leads to the result

$$p(\Delta r) = 4\pi\rho\Delta r^2 e^{-\frac{4}{3}\pi\rho\Delta r^3}. \quad (\text{C.4})$$

The result is immediately generalizable to d dimensions if we take into account that in Eq. (C.1) $4\pi\Delta r^2$ is replaced with $dV\Delta r^{d-1}$ where V is the d -dimensional unit sphere volume, leading to

$$p_d(\Delta r) = dV\rho\Delta r^{d-1} e^{-V\rho\Delta r^d}. \quad (\text{C.5})$$

Appendix D

Eigenvalues spectrum of the Zel'dovich tensor

The Jacobian matrix of the cosmic flow in Zel'dovich approximation is given by

$$\mathcal{J} = \frac{\partial \vec{q}(t)}{\partial \vec{q}^{(i)}} = \mathbb{1}_3 + b(t) D^2 \psi, \quad (\text{D.1})$$

where $D^2 \psi$ is the Hessian of the velocity potential. In this appendix, we derive the expression (1.37) that was given without proof. Since ψ is a random field, so is its Hessian matrix, and the components are distributed with covariance

$$C_{ijkl} = \langle \partial_{ij}^2 \psi \partial_{kl}^2 \psi \rangle, \quad (\text{D.2})$$

that can be written in terms of the power spectrum P_ψ

$$\begin{aligned} \langle \partial_{ij}^2 \psi \partial_{kl}^2 \psi \rangle &= \int_k \int_{k'} \langle k_i k_j \tilde{\psi}(\vec{k}) k'_k k'_l \tilde{\psi}(\vec{k}') \rangle e^{i(\vec{k} \cdot \vec{x} + \vec{k}' \cdot \vec{x})}, \\ &= \int_k P_\psi(k) k_i k_j k_k k_l, \end{aligned} \quad (\text{D.3})$$

which vanishes for an odd number of times occurrence of the indices (i, j, k, l) . Employing the initial density-fluctuation power spectrum

$$\sigma_n^2 = \frac{1}{2\pi^2} \int_0^{+\infty} dk k^{2n-2} P_\delta(k), \quad (\text{D.4})$$

one has

$$\begin{aligned} \langle \partial_{ii}^2 \psi \partial_{ii}^2 \psi \rangle &= \int_k k^4 \mu^4 P_\psi(k) = \frac{\sigma_2^2}{5}, \\ \langle \partial_{ii}^2 \psi \partial_{jj}^2 \psi \rangle &= \int_k k^4 \mu^2 \frac{1 - \mu^2}{2} P_\psi(k) = \frac{\sigma_2^2}{15}, \quad i \neq j. \end{aligned} \quad (\text{D.5})$$

Therefore, the covariance matrix of the data vector $\vec{d} = (\partial_{11}^2, \partial_{22}^2, \partial_{33}^2, \partial_{12}^2, \partial_{13}^2, \partial_{23}^2)^\top \psi$ is

$$C = \langle \vec{d} \otimes \vec{d} \rangle = \frac{\sigma_2^2}{15} \begin{pmatrix} M & 0 \\ 0 & \mathbb{1}_3 \end{pmatrix}, \quad M = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{pmatrix}, \quad (\text{D.6})$$

with determinant

$$\det C = \frac{20\sigma_2^{12}}{15^6}. \quad (\text{D.7})$$

The data vector \vec{d} has Gaussian distribution

$$\varrho[D^2\psi] = \varrho[d_1, d_2, \dots, d_6] = \varrho(\vec{d}) = \frac{15^3}{(2\pi)^2\sigma_2^6\sqrt{20}}e^{-Q/2}. \quad (\text{D.8})$$

with the quadratic form

$$\begin{aligned} Q &= \vec{d}^\top C^{-1} \vec{d} \\ &= \frac{3}{\sigma_2^2}[2(d_1^2 + d_2^2 + d_3^2) - (d_1d_2 + d_1d_3 + d_2d_3)] + \frac{15}{\sigma_2^2}(d_4^2 + d_5^2 + d_6^2). \end{aligned} \quad (\text{D.9})$$

Original proof

Since $D^2\psi = S$ is symmetric, hence it can be diagonalized, such as

$$S = R^\top \Lambda R, \quad \Lambda = \text{diag}(\lambda_1, \lambda_2, \lambda_3), \quad (\text{D.10})$$

by an orthogonal matrix R parametrised by three Euler angles (α, β, γ) such that $R^\top R = \mathbb{1}_3$. Our goal is to obtain the Jacobian determinant of the transformation $\vec{d} \rightarrow \vec{d}' = (\lambda_1, \lambda_2, \lambda_3, 0, 0, 0)^\top$ using Eq. (2.15)

$$\varrho(\lambda_1, \lambda_2, \lambda_3)d\lambda_1 d\lambda_2 d\lambda_3 = \varrho(\vec{d})\mathcal{J}dd_1 dd_2 dd_3 dd_4 dd_5 dd_6, \quad (\text{D.11})$$

The line element of the metric of any symmetric matrix S is given by

$$ds^2 = \text{Tr}(\delta S^2) = \text{Tr}((\delta R^\top)\Lambda R + R^\top(\delta\Lambda)R + R^\top\Lambda(\delta R))^2, \quad (\text{D.12})$$

where with δS we mean the differential of a matrix. Since the matrix element S_{ij} can be written as

$$S_{ij} = \sum_{l,m} R_{il}\Lambda_{lm}R_{jm} = \sum_l R_{il}\lambda_l R_{jl}. \quad (\text{D.13})$$

The infinitesimal matrix δS has entries $(\delta S)_{ij} = dS_{ij}$ given by the differential of S_{ij} . We defined the scalar product between symmetric matrices to be $\langle A, B \rangle = \text{Tr}(AB)$. From $R^\top R = \mathbb{1}_3$, we have

$$0 = (\delta R^\top)R + R^\top(\delta R), \quad (\text{D.14})$$

yielding

$$\begin{aligned} \delta R^\top &= -R^\top(\delta R)R^\top, \\ \delta R &= -R(\delta R^\top)R, \end{aligned} \quad (\text{D.15})$$

allowing us to write

$$\begin{aligned} ds^2 &= \text{Tr}(-R^\top(\delta R)R^\top\Lambda R + \delta\Lambda - R^\top\Lambda R(\delta R^\top)R)^2 \\ &= \text{Tr}(-R^\top(\delta R)\Lambda + \delta\Lambda - \Lambda(\delta R^\top)R)^2 \\ &= \text{Tr}((\delta\Lambda) + [R(\delta R^\top), \Lambda])^2, \end{aligned} \quad (\text{D.16})$$

where $[A, B] = AB - BA$ is the commutator operator between two matrices. Since $R^\top(\delta R) = -R(\delta R^\top)$ is asymmetric, we can write $(R(\delta R^\top))_{ij} = \varepsilon_{ijk}\omega_k$, with the use of the total anti-symmetric tensor.

$$\begin{aligned} [R(\delta R^\top), \Lambda] &= \begin{pmatrix} 0 & \omega_3 & \omega_2 \\ -\omega_3 & 0 & \omega_1 \\ -\omega_2 & -\omega_1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} - \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \begin{pmatrix} 0 & \omega_3 & \omega_2 \\ -\omega_3 & 0 & \omega_1 \\ -\omega_2 & -\omega_1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & \omega_3(\lambda_2 - \lambda_1) & \omega_2(\lambda_3 - \lambda_1) \\ \omega_3(\lambda_2 - \lambda_1) & 0 & \omega_1(\lambda_3 - \lambda_2) \\ \omega_2(\lambda_3 - \lambda_1) & \omega_1(\lambda_3 - \lambda_2) & 0 \end{pmatrix}, \end{aligned} \quad (\text{D.17})$$

and

$$(\delta\Lambda + [R(\delta R^\top), \Lambda])^2 = \begin{pmatrix} d\lambda_1 & \omega_3(\lambda_2 - \lambda_1) & \omega_2(\lambda_3 - \lambda_1) \\ \omega_3(\lambda_2 - \lambda_1) & d\lambda_2 & \omega_1(\lambda_3 - \lambda_2) \\ \omega_2(\lambda_3 - \lambda_1) & \omega_1(\lambda_3 - \lambda_2) & d\lambda_3 \end{pmatrix}^2, \quad (\text{D.18})$$

leading to

$$\text{Tr}(\delta\Lambda + [R(\delta R^\top), \Lambda])^2 = d\lambda_1^2 + d\lambda_2^2 + d\lambda_3^2 + 2[(\lambda_2 - \lambda_3)^2\omega_1^2 + (\lambda_1 - \lambda_3)^2\omega_2^2 + (\lambda_1 - \lambda_2)^2\omega_3^2]. \quad (\text{D.19})$$

The volume element in the space of symmetric 3×3 matrices is then

$$\mathcal{J} = 2\pi^2|\lambda_2 - \lambda_3||\lambda_1 - \lambda_3||\lambda_1 - \lambda_2|d\lambda_1 d\lambda_2 d\lambda_3, \quad (\text{D.20})$$

expressed by the eigenvalues only. Integrating over the Euler angles it result in a $2\pi^2$ factor from the integration of the three sphere. The eigenvalues of the Zel'dovich tensor have the probability distribution

$$\varrho(\lambda_1, \lambda_2, \lambda_3) = \frac{15^3}{8\pi\sigma_2^6\sqrt{5}} |(\lambda_2 - \lambda_3)(\lambda_1 - \lambda_3)(\lambda_1 - \lambda_2)| e^{-\frac{3}{2\sigma_2^2} [2(\lambda_1^2 + \lambda_2^2 + \lambda_3^2) - \lambda_1\lambda_2 - \lambda_1\lambda_3 - \lambda_2\lambda_3]}. \quad (\text{D.21})$$

General prof for invariant models

Invariant models

In this section we give a general proof for those models that are rotationally invariant, namely the ensemble of random matrices in which each matrix H' is related to any other matrix H via the transformation

$$H' = UHU^{-1}, \quad (\text{D.22})$$

where U is an orthogonal (unitary) matrix in case of real symmetric (complex hermitian) matrices with the same probability

$$\rho[H]dH_{11} \cdots dH_{NN} = \rho[H']dH'_{11} \cdots dH'_{NN}. \quad (\text{D.23})$$

This condition requires two more. First we want the **jpdf** of the entries to have the same form before and after the transformation

$$\varrho[H] = \varrho[UHU^{-1}]. \quad (\text{D.24})$$

This constraints $\varrho[H]$ to be a function of the traces of the first N power of H [45], [52],

$$\varrho[H] = \phi(\text{Tr } H, \text{Tr } H^2, \dots, \text{Tr } H^N). \quad (\text{D.25})$$

Furthermore we require that

$$dH_{11} \cdots dH_{NN} = dH'_{11} \cdots dH'_{NN}. \quad (\text{D.26})$$

In this class of matrices lay the Wishart-Laguerre ensemble, the Gaussian Orthogonal Ensemble (GOE), our correlation matrices defined as Euclidean Random Matrices, and the covariance matrix of the entries of the Zel'dovich tensor. After this preliminaries we are ready to state a theorem [46] about this class of matrices.

Theorem 5. *Let the real symmetric $N \times N$ matrix H be invariant under orthogonal similarity transformations (D.24), and having jpdf of entries as in Eq. (D.25). Then the jpdf of the N ordered eigenvalues of H is*

$$\rho_{\text{ord}}(\lambda_1, \dots, \lambda_N) = \frac{\pi^{N^2/2}}{\Gamma_N(N/2)} \phi\left(\sum_i \lambda_i, \dots, \sum_i \lambda_i^N\right) \prod_{i < j} (\lambda_i - \lambda_j). \quad (\text{D.27})$$

The normalizing factor corresponds to

$$\frac{\pi^{N^2/2}}{\Gamma_N(N/2)} = \frac{\int_{\mathbb{V}_N} dO}{2^N}, \quad (\text{D.28})$$

in the orthogonal case, the details of this consideration are further explained in [45]. The integral in the right-hand side of Eq. (D.28) represents the "volume" occupied by orthogonal matrices in \mathbb{R}^{N^2} [53] and it reads

$$\text{Vol}(\mathbb{V}_N) = \int_{\mathbb{V}_N} dO = \frac{2^N \pi^{N^2/2}}{\Gamma_N(N/2)}, \quad (\text{D.29})$$

where we used the multivariate Gamma

$$\Gamma_n(x) = \pi^{n(n-n)/4} \prod_{i=1}^n \Gamma\left(x - \frac{i-1}{2}\right). \quad (\text{D.30})$$

Calling

$$\mathcal{D}O = \frac{dO}{\text{Vol}(\mathbb{V}_N)}, \quad (\text{D.31})$$

this defines the so-called *Haar measure* on the orthogonal group. The Haar measure is invariant under orthogonal conjugation, and defines a probability space on orthogonal matrices.

Proof

We know that

$$\begin{aligned} \varrho(H_{11}, \dots, H_{NN}) \prod_{i \leq j} dH_{ij} &= \varrho(H_{11}(\Lambda, O), \dots, H_{NN}(\Lambda, O)) |\mathcal{J}(H \rightarrow \{\Lambda, O\})| dO \prod_{i=1}^N d\lambda_i \\ &= \hat{\varrho}(\lambda_1, \dots, \lambda_N, O) dO \prod_{i=1}^N d\lambda_i, \end{aligned} \quad (\text{D.32})$$

where the Jacobian of the transformation is

$$\mathcal{J}(H \rightarrow \{\Lambda, O\}) = \prod_{i > j} (\lambda_i - \lambda_j), \quad (\text{D.33})$$

Furthermore, we need to consider that

$$\hat{\varrho}(\lambda_1, \dots, \lambda_N) d\lambda_1 \dots d\lambda_N = d\lambda_1 \dots d\lambda_N \int_{\mathbb{V}_N} dO \hat{\varrho}(\lambda_1, \dots, \lambda_N, O), \quad (\text{D.34})$$

the integration over \mathbb{V}_N is possible when the marginal jpdf of entries is independent of eigenvectors, namely

$$\hat{\varrho}(\lambda_1, \dots, \lambda_N, O) = \varrho(H_{11}(\Lambda), \dots, H_{NN}(\Lambda)) |\mathcal{J}(H \rightarrow \{\Lambda\})|, \quad (\text{D.35})$$

so that Eq. (D.34) becomes

$$\hat{\varrho}(\lambda_1, \dots, \lambda_N) = \varrho(H_{11}(\Lambda), \dots, H_{NN}(\Lambda)) |\mathcal{J}(H \rightarrow \{\Lambda\})| \int_{\mathbb{V}_N} dO. \quad (\text{D.36})$$

where $\int_{\mathbb{V}_N} dO = \frac{2^N \pi^{N^2/2}}{\Gamma_N(N/2)}$ that equals $2^3 \cdot 2\pi^2$ if $N = 3$. Dividing for a factor 2^N to guarantee the uniqueness of the eigendecomposition, fixing the sign of the first row of the matrix O one obtains the right normalization factor. The change of variables between entries and eigenvalues ($H \rightarrow O\Lambda O^\top$) must be one-to-one. But eigenvectors are defined up to a phase, e.g. if \mathbf{v} is a real eigenvector, so is $-\mathbf{v}$.

Vandermonde determinant

Let us prove Eq. (D.33) for any real symmetric matrix H . Since it can be diagonalized, we can write $H = O\Lambda O^\top$ where Λ is the diagonal matrix of the eigenvalues. Differentiating H as in analogy with Eq. (D.12) we have

$$\delta H = (\delta O)\Lambda O^\top + O(\delta\Lambda)O^\top + O\Lambda(\delta O^\top), \quad (\text{D.37})$$

and using the same considerations as in the last section, we obtain

$$\delta \hat{H} = (\delta\Omega)\Lambda - \Lambda(\delta\Omega) + \delta\Lambda, \quad (\text{D.38})$$

with $\delta H = O(\delta \hat{H})O^\top$. Furthermore $\delta\Omega = O^\top \delta O$ is an antisymmetric matrix. Writing

$$d\hat{H}_{ij} = d\Omega_{ij}(\lambda_j - \lambda_i) + d\lambda_i \delta_{ij}, \quad (\text{D.39})$$

we can get the following differential relations

$$\frac{d\hat{H}_{ij}}{d\lambda_k} = \delta_{ij}\delta_{ik}, \quad \frac{d\hat{H}_{ij}}{d\Omega_{kl}} = \delta_{ik}\delta_{jl}(\lambda_j - \lambda_i). \quad (\text{D.40})$$

Then the Jacobian matrix is ready and we can compute the so-called Vandermonde determinant, see E

$$|\det \mathcal{J}| = \prod_{i < j} |\lambda_i - \lambda_j|. \quad (\text{D.41})$$

Marginal pdf of Zel'dovich tensor

Is useful to compute the individual probability distribution of each eigenvalue [6, 54]

$$\begin{aligned}\varrho(\lambda_1) &= \frac{\sqrt{5}}{12\pi\sigma_2} \left\{ 20\frac{\lambda_1}{\sigma_2} \exp\left(-\frac{9\lambda_1^2}{2\sigma_2^2}\right) - \sqrt{2\pi} \exp\left(-\frac{5\lambda_1^2}{2\sigma_2^2}\right) \operatorname{erf}\left(\sqrt{2}\frac{\lambda_1}{\sigma_2}\right) \left(1 - 20\frac{\lambda_1^2}{\sigma_2^2}\right) \right. \\ &\quad \left. - \sqrt{2\pi} \exp\left(-\frac{5\lambda_1^2}{2\sigma_2^2}\right) \left(1 - 20\frac{\lambda_1^2}{\sigma_2^2}\right) + 3\sqrt{3\pi} \exp\left(-\frac{15\lambda_1^2}{4\sigma_2^2}\right) \operatorname{erf}\left(\frac{\sqrt{3}\lambda_1}{2\sigma_2}\right) \right. \\ &\quad \left. + 3\sqrt{3\pi} \exp\left(-\frac{15\lambda_1^2}{4\sigma_2^2}\right) \right\}, \\ \varrho(\lambda_2) &= \frac{\sqrt{15}}{2\sqrt{\pi}\sigma_2} \exp\left(-\frac{15\lambda_2^2}{4\sigma_2^2}\right), \\ \varrho(\lambda_3) &= -\frac{\sqrt{5}}{12\pi\sigma_2} \left\{ 20\frac{\lambda_3}{\sigma_2} \exp\left(-\frac{9\lambda_3^2}{2\sigma_2^2}\right) + \sqrt{2\pi} \exp\left(-\frac{5\lambda_3^2}{2\sigma_2^2}\right) \operatorname{erfc}\left(\sqrt{2}\frac{\lambda_3}{\sigma_2}\right) \left(1 - 20\frac{\lambda_3^2}{\sigma_2^2}\right) \right. \\ &\quad \left. - 3\sqrt{3\pi} \exp\left(-\frac{15\lambda_3^2}{4\sigma_2^2}\right) \operatorname{erfc}\left(\frac{\sqrt{3}\lambda_3}{2\sigma_2}\right) \right\}.\end{aligned}\tag{D.42}$$

with σ_2 being the mass variance defined in Eq. (3.37). Note that the distribution of λ_2 is Gaussian despite that $\lambda_2(\vec{q})$ is not a Gaussian random field. Each of these distribution as been computed from the definition of marginal pdf

$$\varrho(\lambda_i) = \int d\lambda_j \int d\lambda_k \varrho(\lambda_i, \lambda_j, \lambda_k); \quad i \neq j \neq k.\tag{D.43}$$

These eigenvalues distributions lead to

$$\begin{aligned}p(\lambda_1 > 0) &= \frac{23}{25}, \quad p(\lambda_2 > 0) = \frac{1}{2}, \quad p(\lambda_3 > 0) = \frac{2}{25}, \\ \langle \lambda_1 \rangle &= \frac{3}{\sqrt{10\pi}}\sigma_2, \quad \langle \lambda_2 \rangle = 0, \quad \langle \lambda_3 \rangle = -\frac{3}{\sqrt{10\pi}}\sigma_2, \\ \sigma_{\lambda_1}^2 &= \frac{13\pi - 27}{30\pi}\sigma_2^2, \quad \sigma_{\lambda_2}^2 = \frac{2}{15}\sigma_2^2, \quad \sigma_{\lambda_3}^2 = \frac{13\pi - 27}{30\pi}\sigma_2^2,\end{aligned}\tag{D.44}$$

Appendix E

Vandermonde determinant

The Vandermonde determinant is defined as follows

$$\Delta_N(\mathbf{x}) := \prod_{i < j}^N (x_j - x_i) = \det \begin{pmatrix} 1 & \dots & 1 \\ x_1 & \dots & x_N \\ \vdots & & \vdots \\ x_1^{N-1} & \dots & x_N^{N-1} \end{pmatrix}. \quad (\text{E.1})$$

This implies that the Vandermonde determinant is a completely anti-symmetric polynomial in N variables. More generally we can always replace a n -th degree variable x_i^n with a n -th degree polynomial $\pi_n(x_i) = a_k x_i^k + \dots$ in the variable x_i , without affecting the structure of the determinant, namely

$$\Delta_N(\mathbf{x}) := \frac{1}{a_0 a_1 \cdots a_{N-1}} \det \begin{pmatrix} \pi_0(x_1) & \dots & \pi_0(x_N) \\ \pi_1(x_1) & \dots & \pi_1(x_N) \\ \vdots & & \vdots \\ \pi_{N-1}(x_1) & \dots & \pi_{N-1}(x_N) \end{pmatrix}, \quad (\text{E.2})$$

where we didn't consider lower terms in x_i , since they could be anything. That's why, usually, orthogonal polynomials cover an important role in this section, as they usually are the polynomials π_n . So one could choose them to be Hermite or Laguerre polynomials.

Appendix F

Additional calculations

Power spectra

In a linear perturbation regime, the following linearized continuity equation rules

$$\dot{\delta} + \vec{\nabla} \cdot \vec{u} = 0, \quad (\text{F.1})$$

where δ is the density contrast and \vec{u} is the peculiar velocity. Since the density contrast grows linearly following

$$\dot{\delta} \approx H\delta, \quad (\text{F.2})$$

if we write \vec{u} in units of H^{-1} we have

$$\delta \approx -\vec{\nabla} \cdot \vec{u} = -\nabla^2 \psi, \quad (\text{F.3})$$

recalling that $\vec{u} = \vec{\nabla} \psi$. The last equation, in Fourier space, reads

$$\tilde{\delta} = k^2 \tilde{\psi}, \quad (\text{F.4})$$

and therefore

$$P_\delta(k) = \langle |\tilde{\delta}|^2 \rangle = k^4 \langle |\tilde{\psi}|^2 \rangle = k^4 P_\psi(k). \quad (\text{F.5})$$

Exponential correlation function

Power spectrum

Starting from the correlation function

$$\xi_\delta(r) = A e^{-r/R}, \quad (\text{F.6})$$

we are interested in calculating the 3d-Fourier transform, namely the power spectrum

$$\begin{aligned}
P_\delta(k) &= \int_k \xi_\delta(r) e^{-i\vec{k} \cdot \vec{r}} \\
&= A \int_{-1}^{+1} d\mu \int_0^{2\pi} d\varphi \int_0^{+\infty} dr r^2 e^{-r/R - ikr\mu} \\
&= 2\pi A \int_{-1}^{+1} d\mu \frac{2}{(R^{-1} + ik\mu)^3} \\
&= 4\pi A \left[-\frac{1}{2ik(R^{-1} + ik\mu)^2} \right]_{-1}^{+1} \\
&= \frac{2\pi A}{ik} \left(-\frac{1}{(R^{-1} + ik)^2} + \frac{1}{(R^{-1} - ik)^2} \right) \\
&= \frac{8\pi AR^{-1}}{(R^{-2} + k^2)^2},
\end{aligned} \tag{F.7}$$

where we used Eq. (B.8).

Moments of the initial density-fluctuations power spectrum σ_n^2

Having defined

$$\sigma_n^2 = \frac{1}{2\pi^2} \int_0^{+\infty} dk k^{2n-2} P_\delta(k), \tag{F.8}$$

Let us calculate, from Eq. (3.37)

$$\begin{aligned}
\sigma_1^2 &= \frac{1}{2\pi^2} \int_0^{+\infty} dk \frac{8\pi AR^{-1}}{(R^{-2} + k^2)^2} \\
&= \frac{4A}{\pi R} \left[\frac{\frac{R^{-1}k}{R^{-2}+k^2} + \arctan(Rk)}{2R^{-3}} \right]_0^{+\infty} \\
&= \frac{2AR^2}{\pi} \left(\frac{\pi}{2} - 0 \right) \\
&= AR^2.
\end{aligned} \tag{F.9}$$

$$\begin{aligned}
\sigma_2^2 &= \frac{1}{2\pi^2} \int_0^{+\infty} dk k^2 \frac{8\pi AR^{-1}}{(R^{-2} + k^2)^2} \\
&= \frac{4A}{\pi R} \left[\frac{1}{2} \left(\frac{\arctan(Rk)}{R^{-1}} - \frac{k}{R^{-2} + k^2} \right) \right]_0^{+\infty} \\
&= \frac{2A}{\pi R} \left(\frac{\pi R}{2} - 0 \right) \\
&= A.
\end{aligned} \tag{F.10}$$

Appendix G

Numerical considerations

All the $N \times N$ random covariance matrices have been sampled generating particles uniformly at a given number density ρ in a box of length L , where $\rho = \frac{N}{L^3}$. In order to better compare the eigenvalue distributions of different matrix classes, we chose to write all the random matrices such as the entry on the main diagonal is equal to unity. To do that is enough to rescale each entry of the matrix by the diagonal term, since for any autocorrelation matrix, the diagonal terms are all the same. The obtained random matrices can be then numerically diagonalized with MATLAB®. The diagonalization procedure can be parallelized to improve the performance.

Acronyms

KFT Kinetic Field Theory

RMT Random Matrix Theory

ERM Euclidean Random Matrix

CDM Cold Dark Matter

WL Wishart Laguerre

GOE Gaussian Orthogonal Ensemble

GUE Gaussian Unitary Ensemble

GSE Gaussian Symplectic Ensemble

pdf probability distribution function

cpdf cumulative probability distribution function

jpdf joint probability distribution function

i.i.d. independent and identically distributed

CLT Central Limit Theorem

List of Figures

1.1	Cosmic Microwave Background	5
1.2	Simulation of cosmic web	8
1.3	Zel'dovich flow	9
1.4	Evolution of cosmic webs from random field.	9
2.1	Gaussian random field	15
2.2	Gaussian random field with $P(k) \sim k^{-\alpha}$	16
3.1	Illustration of the idea of KFT	18
3.2	Non-linear power spectrum $\mathcal{P}(k)$ in KFT.	28
4.1	2×2 GOE matrix eigenvalues.	32
4.2	Semicircle law	33
4.3	Marčenko-Pastur law.	36
4.4	Analytical finite $N = 8$ GOE matrix	39
4.5	Gaussian and exponential ERM	40
4.6	$f(r) \sim \frac{\sin(r)}{r}$ ERM	40
5.1	Power spectrum and momentum functions of CDM, with their fit.	43
5.2	Eigenvalue distribution of $C_{\delta\delta}$	45
5.3	Eigenvalues distribution of $C_{\delta\delta}$ at different densities.	45
5.4	High-density limit of $C_{\delta\delta}$ and with Gaussian and exponential power spectrum. . .	47
5.5	High-density limit of $C_{\delta\delta}$ and with CDM power spectrum.	48
5.6	Analytical expression of $2 \times 2 C_{pp}$	49
5.7	2×2 eigenvalues distribution of \mathbf{C}_{pp}	51
5.8	\mathbf{C}_{pp} eigenvalues distribution at different N and ρ	52
5.9	Contribution of $a_2(r)$ to \mathbf{C}_{pp}	53
5.10	Sample counts of $3 \times 3 C_{ p p }$ covariance matrix.	53
5.11	High-density limit of \mathbf{C}_{pp}	54
5.12	Low-density limit of \mathbf{C}_{pp} and C_{pp} with CDM power spectrum.	55
5.13	Single eigenvalues distribution for \mathbf{C}_{pp}	56
5.14	Single eigenvalues distribution for the $C_{ p p }$ ERM.	56

5.15 Δ at different densities	57
--	----

Bibliography

- [1] M. Skrutskie, R. Cutri, R. Stiening, M. Weinberg, S. Schneider, J. Carpenter, C. Beichman, R. Capps, T. Chester, J. Elias, *et al.*, “The two micron all sky survey (2mass),” *The Astronomical Journal*, vol. 131, no. 2, p. 1163, 2006.
- [2] J. A. Peacock, M. Colless, I. Baldry, C. Baugh, J. Bland-Hawthorn, T. Bridges, R. Cannon, S. Cole, C. Collins, W. Couch, *et al.*, “Studying large-scale structure with the 2df galaxy redshift survey,” *arXiv preprint astro-ph/0204239*, 2002.
- [3] D. G. York, J. Adelman, J. E. Anderson Jr, S. F. Anderson, J. Annis, N. A. Bahcall, J. Bakken, R. Barkhouser, S. Bastian, E. Berman, *et al.*, “The sloan digital sky survey: Technical summary,” *The Astronomical Journal*, vol. 120, no. 3, p. 1579, 2000.
- [4] P. A. Ade, N. Aghanim, M. Arnaud, F. Arroja, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. Banday, R. Barreiro, *et al.*, “Planck 2015 results-xvii. constraints on primordial non-gaussianity,” *Astronomy & Astrophysics*, vol. 594, p. A17, 2016.
- [5] N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. Barreiro, N. Bartolo, S. Basak, *et al.*, “Planck 2018 results-vi. cosmological parameters,” *Astronomy & Astrophysics*, vol. 641, p. A6, 2020.
- [6] A. Doroshkevich, “Spatial structure of perturbations and origin of galactic rotation in fluctuation theory,” *Astrophysics*, vol. 6, no. 4, pp. 320–330, 1970.
- [7] J. M. Bardeen, J. Bond, N. Kaiser, and A. Szalay, “The statistics of peaks of gaussian random fields,” *Astrophysical Journal, Part 1 (ISSN 0004-637X)*, vol. 304, May 1, 1986, p. 15-61. SERC-supported research., vol. 304, pp. 15–61, 1986.
- [8] P. Peebles, “Large-scale background temperature and mass fluctuations due to scale-invariant primeval perturbations,” 1982.
- [9] F. Bernardeau, S. Colombi, E. Gaztanaga, and R. Scoccimarro, “Large-scale structure of the universe and cosmological perturbation theory,” *Physics reports*, vol. 367, no. 1-3, pp. 1–248, 2002.
- [10] S. P. Das and G. F. Mazenko, “Field theoretic formulation of kinetic theory: basic development,” *Journal of Statistical Physics*, vol. 149, pp. 643–675, 2012.

- [11] S. P. Das and G. F. Mazenko, "Newtonian kinetic theory and the ergodic-nonergodic transition," *Journal of Statistical Physics*, vol. 152, pp. 159–194, 2013.
- [12] S. P. Das and G. F. Mazenko, "Field Theoretic Formulation of Kinetic Theory: Basic Development," *Journal of Statistical Physics*, vol. 149, pp. 643–675, Nov. 2012.
- [13] G. F. Mazenko, "Smoluchowski dynamics and the ergodic-nonergodic transition," *Physical Review E*, vol. 83, no. 4, p. 041125, 2011.
- [14] M. Bartelmann, F. Fabis, D. Berg, E. Kozlikin, R. Lilow, and C. Viermann, "A microscopic, non-equilibrium, statistical field theory for cosmic structure formation," *New Journal of Physics*, vol. 18, no. 4, p. 043020, 2016.
- [15] M. Bartelmann, E. Kozlikin, R. Lilow, C. Littek, F. Fabis, I. Kostyuk, C. Viermann, L. Heisenberg, S. Konrad, and D. Geiss, "Cosmic structure formation with kinetic field theory," *Annalen der Physik*, vol. 531, no. 11, p. 1800446, 2019.
- [16] M. Bartelmann, J. Dombrowski, S. Konrad, E. Kozlikin, R. Lilow, C. Littek, C. Pixius, and F. Fabis, "Kinetic field theory: Non-linear cosmic power spectra in the mean-field approximation," *SciPost Physics*, vol. 10, no. 6, p. 153, 2021.
- [17] C. Pixius, S. Celik, and M. Bartelmann, "Kinetic field theory: perturbation theory beyond first order," *Journal of Cosmology and Astroparticle Physics*, vol. 2022, no. 12, p. 030, 2022.
- [18] M. Mézard, G. Parisi, and A. Zee, "Spectra of euclidean random matrices," *Nuclear Physics B*, vol. 559, no. 3, pp. 689–701, 1999.
- [19] G. Parisi *et al.*, "Euclidean random matrices: solved and open problems," *Application of Random Matrices in Physics. NATO Science Series*, vol. 206, pp. 219–260, 2006.
- [20] J. Wishart, "The generalised product moment distribution in samples from a normal multivariate population," *Biometrika*, pp. 32–52, 1928.
- [21] E. P. Wigner, "Random matrices in physics," *SIAM review*, vol. 9, no. 1, pp. 1–23, 1967.
- [22] E. P. Wigner, "Characteristic vectors of bordered matrices with infinite dimensions i," *The Collected Works of Eugene Paul Wigner: Part A: The Scientific Papers*, pp. 524–540, 1993.
- [23] P. A. Ade, N. Aghanim, M. Alves, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, H. Aussel, C. Baccigalupi, *et al.*, "Planck 2013 results. i. overview of products and scientific results," *Astronomy & Astrophysics*, vol. 571, p. A1, 2014.
- [24] S. Konrad and M. Bartelmann, "Kinetic field theory for cosmic structure formation," *La Rivista del Nuovo Cimento*, vol. 45, no. 11, pp. 737–799, 2022.
- [25] O. Hahn, "Collisionless dynamics and the cosmic web," *Proceedings of the International Astronomical Union*, vol. 11, no. S308, pp. 87–96, 2014.

- [26] Y. B. Zel'Dovich, "Gravitational instability: An approximate theory for large density perturbations.,," *Astronomy and Astrophysics*, Vol. 5, p. 84-89, vol. 5, pp. 84–89, 1970.
- [27] T. Brinckmann, M. Lindholmer, S. Hansen, and M. Falco, "Zeldovich pancakes in observational data are cold," *Journal of Cosmology and Astroparticle Physics*, vol. 2016, no. 04, p. 007, 2016.
- [28] T. Tao, *Topics in random matrix theory*, vol. 132. American Mathematical Society, 2023.
- [29] M. L. Mehta, *Random matrices*. Elsevier, 2004.
- [30] M. Mehta and R. Matrices, "the statistical theory of energy levels," 1967.
- [31] B. Eynard, T. Kimura, and S. Ribault, "Random matrices," *arXiv preprint arXiv:1510.04430*, 2015.
- [32] É. Brézin, V. Kazakov, D. Serban, P. Wiegmann, and A. Zabrodin, *Applications of random matrices in physics*, vol. 221. Springer Science & Business Media, 2006.
- [33] C. Ganter and W. Schirmacher, "Euclidean random matrix theory: low-frequency non-analyticities and rayleigh scattering," *Philosophical Magazine*, vol. 91, no. 13-15, pp. 1894–1909, 2011.
- [34] T. Grigera, V. Martín-Mayor, G. Parisi, and P. Verrocchio, "Vibrational spectrum of topologically disordered systems," *Physical Review Letters*, vol. 87, no. 8, p. 085502, 2001.
- [35] T. Grigera, V. Martín-Mayor, G. Parisi, and P. Verrocchio, "Vibrations in glasses and euclidean random matrix theory," *Journal of Physics: Condensed Matter*, vol. 14, no. 9, p. 2167, 2002.
- [36] T. S. Grigera, V. Martin-Mayor, G. Parisi, P. Urbani, and P. Verrocchio, "On the high-density expansion for euclidean random matrices," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2011, no. 02, p. P02015, 2011.
- [37] V. Martin-Mayor, M. Mézard, G. Parisi, and P. Verrocchio, "The dynamical structure factor in topologically disordered systems," *The Journal of Chemical Physics*, vol. 114, no. 18, pp. 8068–8081, 2001.
- [38] A. Amir, Y. Oreg, and Y. Imry, "Mean-field model for electron-glass dynamics," *Physical Review B*, vol. 77, no. 16, p. 165207, 2008.
- [39] A. Amir, Y. Oreg, and Y. Imry, "Electron glass dynamics," *Annu. Rev. Condens. Matter Phys.*, vol. 2, no. 1, pp. 235–262, 2011.
- [40] A. Amir, J. J. Krich, V. Vitelli, Y. Oreg, and Y. Imry, "Emergent percolation length and localization in random elastic networks," *Physical Review X*, vol. 3, no. 2, p. 021017, 2013.
- [41] E. Bogomolny, O. Bohigas, and C. Schmit, "Spectral properties of distance matrices," *Journal of Physics A: Mathematical and General*, vol. 36, no. 12, p. 3595, 2003.

- [42] S. Ciliberti, T. S. Grigera, V. Martín-Mayor, G. Parisi, and P. Verrocchio, “Anderson localization in euclidean random matrices,” *Physical Review B*, vol. 71, no. 15, p. 153104, 2005.
- [43] B. Huang and T.-M. Wu, “Localization-delocalization transition in hessian matrices of topologically disordered systems,” *Physical Review E*, vol. 79, no. 4, p. 041105, 2009.
- [44] A. Goetschy and S. Skipetrov, “Euclidean random matrices and their applications in physics,” *arXiv preprint arXiv:1303.2880*, 2013.
- [45] G. Livan, M. Novaes, and P. Vivo, “Introduction to random matrices theory and practice,” *Monograph Award*, vol. 63, pp. 54–57, 2018.
- [46] A. Edelman, “Eigenvalues and condition numbers of random matrices,” *SIAM journal on matrix analysis and applications*, vol. 9, no. 4, pp. 543–560, 1988.
- [47] E. P. Wigner, *Statistical properties of real symmetric matrices with many dimensions*. Princeton University, 1957.
- [48] F. J. Dyson, “Statistical theory of the energy levels of complex systems. i,” *Journal of Mathematical Physics*, vol. 3, no. 1, pp. 140–156, 1962.
- [49] V. A. Marchenko and L. A. Pastur, “Distribution of eigenvalues for some sets of random matrices,” *Matematicheskii Sbornik*, vol. 114, no. 4, pp. 507–536, 1967.
- [50] S. Ghosh and A. Pandey, “Skew-orthogonal polynomials and random-matrix ensembles,” *Physical Review E*, vol. 65, no. 4, p. 046221, 2002.
- [51] A. M. Mathai, P. Moschopoulos, and G. Pederzoli, “Distance between random points in a cube,” *Statistica*, vol. 59, no. 1, pp. 61–81, 1999.
- [52] H. Weyl, *The classical groups: their invariants and representations*. No. 1, Princeton university press, 1946.
- [53] R. J. Muirhead, *Aspects of multivariate statistical theory*. John Wiley & Sons, 2009.
- [54] J. Lee and S. F. Shandarin, “The cosmological mass distribution function in the zeldovich approximation,” *The Astrophysical Journal*, vol. 500, no. 1, p. 14, 1998.

Acknowledgements

Arrived at this point I would like to express my gratitude to my supervisor Matthias Bartelmann, who gave me the opportunity to work on this project and share experiences within the Cosmology group.

I would also like to thank Professor Tilman Enss, who agreed to correct this thesis.

I am grateful to my family for their support and encouragement throughout my studies.

I deeply thank Alessandro, Matteo and Luca for sharing this fantastic abroad experience with me. I bring with me all the discussions we had in the Mensa and all the great moments that you made me have.

I express my profound thanks to my girlfriend Matilde who accompanied me in my journey through the thesis and chose the proper colormaps for the plots.

I thank Irene for having been the best partner for the exercises.

I also want to thank Heinrich and Richard for helping me with the German version of the abstract.

Many thanks to my flatmate Valentin, who helped me since the first day to settle in Germany and who managed to teach me some Deutsch.

I am grateful to Sara, Giulia C., Giulia B., I thank Federico who has always been with me even from Munich.

Erklärung:

Ich versichere, dass ich diese Arbeit selbstständig verfasst habe und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Heidelberg, den 23.04.2024

Davide Mapelli