The projected mass distribution and the transition to homogeneity

José Gaite

Applied Physics Dept., ETSIAE, Univ. Politécnica de Madrid, E-28040 Madrid, Spain
E-mail: jose.gaite@upm.es

Abstract. A statistical analysis of the angular projection of the large-scale stellar mass distribution, as obtained from the Sloan Digital Sky Survey (data release 7) with the stellar masses of galaxies, finds values of the clustering length $r_0$ and the power-law exponent $\gamma$ of the two-point correlation function that are larger than the standard values, on account of the presence of very massive galaxies. A multifractal cosmic-web model with a transition to homogeneity at about $10\,h^{-1}\text{Mpc}$ is still good, but there is considerable uncertainty in both this scale and the correlation dimension.

Keywords: cosmic web, superclusters, galaxy clusters, redshift surveys

Contents

1 Introduction 2
2 The scale of transition to homogeneity in redshift surveys 3
3 Fractal projections 5
4 The angular correlation of galaxies 6
5 Angular analysis of the stellar mass distribution 8
  5.1 Complete magnitude-limited sample 9
  5.2 Samples with redshift cutoffs 10
6 Regarding the calculation of $\sigma_8$ 12
7 Conclusions and discussion 13
1 Introduction

It is claimed that we are entering the era of precision cosmology, in which the fundamental parameters of the universe are known within a few percent precision and only remains to progressively refine them [1, 2]. In addition to the global parameters of the FLRW model, we have the parameters that determine the primordial density perturbations and hence the formation of large scale structure; in particular, the amplitude of the primordial density perturbations, usually measured in terms of the present linear-theory mass dispersion on a scale of $8 \, h^{-1}\text{Mpc}$, named $\sigma_8$. The scale of $8 \, h^{-1}\text{Mpc}$ comes from Peebles’ observation that galaxy counts on this scale have a rms fluctuation $\delta N/N$ approximately equal to one [3]. In general, the mass dispersion over the length $R$, $\sigma(R)$, defines a scale such that $\sigma(R) = 1$, which separates the strong clustering regime from the large-scale quasi-homogeneous mass distribution.

A similar scale is the clustering length of galaxies $r_0$, such that $\xi(r_0) = 1$, where $\xi(r)$ is the reduced galaxy-galaxy correlation function. This function is well approximated by the power law

$$\xi(r) = \left(\frac{r_0}{r}\right)^\gamma,$$

for distances not much larger than $r_0$ [3]. The scale such that $\delta N/N = 1$ is related to $r_0$ by a $\gamma$-dependent factor [3]. Historically, the power-law form of $\xi(r)$ was deduced from the angular positions of galaxies, yielding $\gamma = 1.8$ and $r_0 = 4.7 \, h^{-1}\text{Mpc}$ [4, 5]. Nowadays, in spite of the availability of galaxy redshift surveys, the angular positions are still useful, because they are more precise and more abundant, allowing better statistics. In particular, the angular correlation function is used, in combination with other data, for recent precision values of $\sigma_8$ [2].

The amplitude of the primordial density perturbations is not theoretically constrained and determines the present scale of transition to homogeneity. The transition to homogeneity has been the subject of numerous studies, many of them motivated by the bold proposal that no such transition exists [6–8]. This proposal, namely, that the mass distribution is fully scale invariant, at all scales in a Newtonian cosmology, has been called the “fractal universe” [3]; which is rather a misnomer, because fractal geometry describes the properties of mass distributions on infinitesimal scales rather than on large scales. Indeed, Eq. (1.1), for any magnitude of $r_0$, corresponds to a fractal distribution on scales $r \ll r_0$. At any rate, the scale and form of the transition to homogeneity determine the size of the largest structures that can be observed [9]. An examination of the literature shows values of this scale that go from $5 \, h^{-1}\text{Mpc}$ to more than $20 \, h^{-1}\text{Mpc}$ [3, 10–14]. In contrast, the present precision values of $\sigma_8$ are assigned a relative error of about one percent [2]. This small error justifies the adjective “precise”, but such precision is surprising.

In this context, we find it useful to make a new analysis of the angular correlation function, for its intrinsic interest and to assess its suitability for a study of the transition to homogeneity. The main novelty is that we add an important ingredient: the stellar masses of galaxies. Therefore, we study the projected stellar mass distribution, specifically, as a projected fractal mass distribution. Given the broad range of galaxy masses, Pietronero [7] already argued that these masses must be taken into account and, moreover, that a full multifractal analysis of galaxy clustering is necessary. Such analysis was initiated by Pietronero and collaborators [8, 15]. Meanwhile, the large-scale structure has been known to consist not just of clusters but of a web structure [16, 17], which can be well described as a multifractal [18, 19]. However, no study of the projected mass distribution exists yet.
Our galaxy data come from the Sloan Digital Sky Survey, data release 7 (SDSS DR7) [20], as provided by the New York University Value-Added Galaxy Catalog (NYU-VAGC) [21]. The SDSS DR7 has been well studied and, in particular, there is a careful analysis of its angular two-point correlation function [22]. Here we study the scaling properties of the SDSS DR7 in the angular projection, but using the stellar masses of the galaxies [23, 24], like in our recent multifractal analysis of the same data [18]. We chiefly focus on the correlation dimension, deduced from the two-point correlation function.

We begin in Sect. 2 with a short discussion of the transition to homogeneity in galaxy redshift surveys and a re-analysis of the scaling of the correlation function found in Ref. [18]. Then, we proceed to our main subject, namely, the properties of projections of fractal distributions, showing the relevance of the two-point angular correlation function for angular projections (Sect. 3). We connect with the standard theory of the angular correlation function of galaxies, to reinterpret it in terms of the projected mass distribution (Sect. 4). After this theoretical study, we can undertake the analysis of the angular projection of the SDSS DR7 (Sect. 5). The introduction of galaxy masses causes an unexpected problem that we try to solve in Sect. 5.2. We end with some remarks about the calculation of $\sigma_8$ (Sect. 6) and a general discussion (Sect. 7).

2 The scale of transition to homogeneity in redshift surveys

To extract information from galaxy redshift surveys, it is necessary to construct volume limited samples. If a volume limited sample is sufficiently deep, it is possible to obtain the scale of homogeneity, although there is no universally accepted definition of this scale. Although fractal properties belong in the strongly inhomogeneous regime, the scale of homogeneity is necessary for an accurate calculation [18]. The most suitable method of determining that scale consists in the evaluation of $(\delta M/M) (R)$, namely, the scale dependent rms fluctuation of the stellar mass, and choosing an appropriate value. Let us review this question.

We have recalled above Peebles’ use of $(\delta N/N) (R)$, the scale-dependent fluctuation of galaxy counts in cells [3]; but it is more accurate to consider $(\delta M/M) (R)$, replacing galaxy counts by total stellar mass. Therefore, to choose a definite scale of transition to homogeneity, it seems natural to take $(\delta M/M) (R) = 1$, but one must be aware that a mass distribution with such fluctuations is far from being Gaussian and therefore is not homogeneous at all. In fact, it is argued in Ref. [18] that a suitable criterion for homogeneity is the mass variance $(\delta M/M)^2 (R) = 0.1$. Whether or not the scale so determined is close to the scale defined by $\delta M/M = 1$ depends on how sharp is the transition to homogeneity; and it may not be sharp.

In this regard, it is useful to consider the precise form of the function $(\delta M/M) (R)$. If the two-point correlation function of the stellar mass distribution follows a power law, Eq. (1.1), so does $(\delta M/M) (R)$; namely,

$$\frac{\delta M}{M} = C(\gamma)^{1/2} \left( \frac{r_0}{R} \right)^{\gamma/2},$$

where the form of the function $C(\gamma)$ depends on the shape of the cell. For a sphere of radius $R$, it is given by Peebles [3] and fulfills $C(\gamma)^{1/2} \in (1.1, 1.5)$ for $\gamma \in (1, 2)$ (the natural range). Therefore, the scale such that $\delta M/M$ equals a given number is a definite function of $\gamma$. For $\delta M/M = 1$, $R$ is only a little larger than $r_0$, whereas, for $\delta M/M = 0.1^{1/2} = 0.32$, it is between $4.7r_0$ and $12r_0$ (the larger, the smaller is $\gamma$). In conclusion, $r_0$ and the scale such that $\delta M/M = 1$ are roughly equivalent scales but the scale where one can observe real homogeneity
is several times larger, especially, for low values of $\gamma$. The standard values $r_0 = 5.4 h^{-1}\text{Mpc}$ and $\gamma = 1.8$ are thus compatible with the presence of relatively large structures; for example, cosmic voids of size $\sim 30 h^{-1}\text{Mpc}$ [3]. Actually, even larger structures can be observed, provided that $\xi(r)$ does not fall too rapidly for $r \gg r_0$ [9].

At any rate, the values of $r_0$ and $\gamma$ obtained from redshift surveys may not be precise, because the scaling range in which Eq. (1.1) is measured is necessarily small. Notice, for example, the broad range of values of the correlation dimension $D_2 = 3 - \gamma$ in the table of Ref. [10], which are quoted from various sources. Since Eq. (1.1) is obtained from a linear fit of a log-log graph that yields both $\gamma$ and $r_0$, the range of values of $r_0$ has a similar breadth. In fact, one sees values of the scale of homogeneity that go from $5 h^{-1}\text{Mpc}$ to $30 h^{-1}\text{Mpc}$ or more [11–14]. Of course, some of these references are not calculating $r_0$ but some other scale that is necessarily larger. However, even if one just tries to find $r_0$, there are two common ways of doing it that can give quite different results. One can employ for the power-law fit the reduced correlation function

$$\xi(r) = \frac{\langle \varrho(r) \varrho(0) \rangle}{\langle \varrho \rangle^2} - 1 = \frac{\langle \delta \varrho(r) \delta \varrho(0) \rangle}{\langle \varrho \rangle^2}$$

(2.1)
or just the correlation function $1 + \xi(r)$. If $\xi \gg 1$, then the two functions are equivalent, but this condition is not fulfilled in the full range to fit. Let us illustrate this point with an argument and an example.

For a fractal analysis, it is argued that the average density is not a primary concept and, therefore, it is natural to employ the correlation function, specifically, in the form of conditional density, namely, $\Gamma(r) = \langle \varrho(r) \varrho(0) \rangle / \langle \varrho \rangle$, or also $\Gamma^*(r) = \int_0^r \Gamma(r') 4\pi r'^2 dr' / (4\pi r'^3/3)$ [8]. This is the measure employed in several works that find a large scale of homogeneity, e.g., the analyses of the SDSS in Refs. [11, 13]. In our multifractal analysis of the stellar mass distribution in the SDSS-DR7 [18], the scale of homogeneity is a prerequisite and it is found from the scaling of the second moment of the coarse-grained density in cells of volume $v$, namely,

$$\mu_2(v) = \langle \varrho_v^2 \rangle / \langle \varrho_v \rangle^2,$$

which can be expressed as the corresponding average of the correlation function. The power-law fit of $\mu_2(v)$ in Ref. [18] is equivalent to a power-law fit of the correlation function [and also equivalent to a power-law fit of $\Gamma(r)$ or $\Gamma^*(r)$]. But one can instead fit $\mu_2(v) - 1$, which is equivalent to fitting $\xi(r)$. Notice that $\mu_2 - 1 = (\delta M/M)^2$. Both the two options are displayed in Fig. 1, which we now explain.

We test scaling in the volume limited sample in the redshift range $[0.003, 0.013]$ called VLS1 in Ref. [18] (where a full description of the sample can be found). The solid blue lines in Fig. 1 are the graphs of $\mu_2$ and $\mu_2 - 1$, The fit to the power law

$$\mu_2 = (v/v_0)^{D_2/3-1}$$

in the interval $v \in [3, 190] \text{Mpc}^3/h^3$ yields $v_0 = 2000 \text{Mpc}^3/h^3$ and $D_2 = 1.43$ (and $\gamma = 3 - D_2 = 1.57$). An analogous fit to the power law

$$\mu_2 - 1 = (v/v_0)^{D_2/3-1}$$

yields $v_0 = 850 \text{Mpc}^3/h^3$ and $D_2 = 1.20$ ($\gamma = 1.80$). Both fits are represented by dashed red lines in Fig. 1. The latter fit, with $\gamma = 1.80$ and $v_0^{1/3} = 9.5 \text{Mpc}/h$, basically agrees with the classic values of $\gamma$ and $r_0$. However, both fits are questionable, because the fractal
regime takes place for $\mu_2 - 1 \approx \mu_2 \gg 1$ and, arguably, the fits extend beyond the small scales where that condition holds. Therefore, some intermediate values of $v_0$ and $\gamma$ should be more appropriate, although they are difficult to specify, due to the small range available.

In Ref. [18], we have studied two other volume limited samples, namely, VLS2, with redshift range $z \in [0.02, 0.03]$, and VLS3, with $z \in [0.04, 0.06]$. But they have reduced scaling ranges. In fact, the scaling ranges are so reduced that any fit is quite uncertain. Although it may seem that deeper volume limited samples should provide larger scaling ranges, this is not so. The deeper the volume limited sample, the more luminous the galaxies in it and the smaller their number density, so that discreteness effects take over growing ranges of the smaller scales. It is to be remarked that this problem does not preclude a partial fractal analysis. For example, a relevant part of the multifractal spectrum $f(\alpha)$ can be calculated from VLS2 and VLS3 [18]. On the other hand, in VLS1, the coarse multifractal spectrum for $v = 1.53 \cdot 10^3$ (Mpc/$h$)$^3$ is fairly reliable, in spite that $\mu_2$ is as low as 1.5. In general, we find that the multifractal spectrum $f(\alpha)$ is more robust than the spectrum of dimensions $D_q$, although the two functions provide equivalent information from the mathematical standpoint.

At any rate, we have to cope with the limitations in the fractal analysis of current redshift surveys and admit that it cannot provide yet precise values of some relevant quantities, in particular, $D_2$ and $v_0$. Although their values often appear with small relative errors, these errors correspond to particular fits and do not represent the real total errors, which are considerably larger, as shown by the fact that, often, the confidence intervals do not overlap. Therefore, it is useful to analyze angular surveys. In particular, if we discard the redshift and just consider the angular positions in the SDSS-DR7, we do not have to construct volume limited samples and we have a large number of galaxies available. However, we have to study first the properties of angular projections.

3 Fractal projections

The study of fractal projections has tradition in mathematics [25, 26]. In cosmology, the properties of angular projections of fractal sets have been considered in regard to the possibility of a fractal universe with no transition to homogeneity [6, 8, 27]. Of course, the study of the angular projection of the distribution of galaxies, that is to say, of angular galaxy surveys, is old and predates the study of redshift surveys [4, 5]. In fact, Peebles [3] uses the
properties of the angular distribution of galaxies as a powerful argument against a fractal universe with no transition to homogeneity. This argument will also be valid here.

We have to consider, first of all, the generalization of the theory of projection of fractal sets to projection of fractal measures, which has been developed more recently [26]. This is necessary because the large-scale mass distribution has to be treated as a measure rather than as a set. To be precise, it is a multifractal mass distribution of non-lacunar type; namely, the support of the mass distribution is the full volume. These concepts are explained in detail in Ref. [19]. The non-lacunarity is observed in the dark matter distribution obtained in cosmological N-body simulations and is also very likely in the stellar mass distribution. In this regard, we recall that Durrer et al. [27] intended to show that the angular projection of a fractal set can have a vanishing lacunarity, with various arguments, mainly, that the projection of a fractal with dimension 2 or higher is non-fractal, but also other arguments related to galaxy properties (apparent sizes and opacities). If the three-dimensional stellar mass distribution is already non-lacunar, then that issue is no longer meaningful. Furthermore, a small lacunarity would also give rise to a non-lacunar projection, because it is certain that the fractal dimension of the support of the distribution is considerably larger than two [18].

The study of projections of multifractal mass distributions is not reduced to the behavior of lacunarity. If we characterize a mass distribution by its dimension spectrum $D_q$, the question of how this spectrum behaves under projections has been partially answered [28]. The answer is the natural generalization of the standard result for fractal sets [25, 26]: in particular, the dimension of the projection of a fractal set onto a plane (or a smooth surface) equals the dimension of the set if it is smaller than two and it is two otherwise (this statement has to be qualified for non-random fractals, which can have special projections along some directions). This result is still valid for the dimension spectrum $D_q$ of a mass distribution, with the restriction that $1 < q \leq 2$ [28]. The restriction is due to technical reasons and may not apply to some types of mass distributions, but we are mostly interested in that interval and, specifically, in the correlation dimension $D_2$. Since we expect that $D_2$ is in the range 1.2–1.4 for the stellar mass distribution (Sect. 2), it has to be preserved under angular projections. It must be noticed that some authors find that $D_2 \geq 2$ [10, table 1], but they use scale ranges so large that the corresponding fits are probably biased. Moreover, all those values refer to the galaxy number density.

We can measure $D_2$ in angular catalogs of galaxies by calculating the angular correlation function. The two-point angular correlation of galaxies is defined in analogy with its three-dimensional counterpart [3–5]. However, the usual definition refers to correlations of galaxy positions and we are interested in correlations of the stellar mass distribution; so we have to replace number density with mass density in Eq. 2.1. Let us recall and generalize some standard results about the two-point correlation of angular positions.

### 4 The angular correlation of galaxies

The reduced two-point correlation of angular positions, denoted $w(\theta_{12})$, can be expressed as an integral of the two-point correlation of ordinary positions $\xi(r_{12})$ over the radial coordinates $r_1$ and $r_2$. The integral can be simplified for a power-law $\xi$, Eq. (1.1), and in the small-angle approximation, $\theta \ll 1$ ($\theta$ is measured in radians). It gives

$$w(\theta) = K \theta^{1-\gamma} \left( \frac{r_0}{d_e} \right)^\gamma \int_{-\infty}^{\infty} \frac{dx}{(1 + x^2)^{\gamma/2}},$$

(4.1)
where \( K \) is a constant that depends on \( \gamma \) and the radial selection function, and \( d_* \) is the characteristic sample depth \([3]\). The integral over \( x \), which is a variable proportional to the relative radial coordinate, is left unevaluated because the integral can be divergent; namely, it is divergent for \( \gamma \leq 1 \). In this case, the present approximation fails, because the projected correlation function is dominated by pairs of points at large relative radial distances.

Let us remark that \( w(\theta) \) and Eq. (4.1) are commonly assumed to refer to correlations of galaxy positions but they are equally applicable to the stellar mass distribution or any mass distribution. Indeed, \( \gamma \leq 1 \) is equivalent to \( D_2 \geq 2 \), the case of non-fractal projection. In the case that \( \gamma > 1 \), the projected correlation function is dominated by pairs of points at small relative radial distances, so that the three-dimensional fractal structure, as a set of short-distance geometric properties, is preserved in the projection.

It is also useful to notice that \( K \) and the integral in Eq. (4.1), for \( \gamma > 1 \) and not too close to one, are factors of the order of unity, so the magnitude of \( w(\theta) \) is ruled by the quotient \( r_0/d_* \) \([3]\). If the characteristic sample depth is much larger than \( r_0 \), as is normal in deep surveys, the angular correlations are small, except for very small angles. This means that the projected fractal structure, which requires large fluctuations, that is to say, \( w(\theta) \gg 1 \), is only observable for very small \( \theta \). In contrast, \( \gamma \) and hence \( D_2 \) can perhaps be measured in a larger range of values of \( \theta \), provided that the power law in Eq. (4.1) holds in that range.

It is obvious that the cosmic structure appears blurred in angular images of the galaxy distribution, as illustrated, for example, by the image of the Lick survey \([3, \text{Fig. 3.9}]\). These images, on the one hand, are proof of large scale homogeneity, namely, of a small ratio \( r_0/d_* \); but, on the other hand, one can hardly distinguish the cosmic web features, which are quite sharp in slices of three dimensional redshift surveys. If the cosmic web were an ideal mathematical fractal distribution, its features would appear at very small angles, that is to say, zooming in on an angular image of the mass distribution. But the galaxies have discrete nature. In this respect, we should notice that the two-point radial distance is integrated down to zero in Eq. (4.1), whereas correlation functions are actually calculated as sums over pairs of points in a finite set. In particular, there is a minimal distance for some pair, while the density of projected far points on the angular location of that pair keeps growing as \( d_* \) grows. It is evident that uncorrelated far points must blur the small scale features and even obliterate them at some stage.

The cosmic web structure also fades in the three-dimensional mass distribution when its sampling by galaxies does not have a sufficiently high number density to probe scaling well below \( r_0 \), where \( \xi \) is large. We have already noticed this problem in Sect. 2, as regards our volume limited samples from SDSS-DR7. Of course, one can also analyze scales larger than \( r_0 \), where \( \xi < 1 \), namely, the scales of the transition to homogeneity, and perhaps find that a power law, like Eq. (1.1), provides a good fit in some range of scales. This power law, however, will be unrelated to the strong clustering, fractal regime, and it will just show the presence of a type of weak clustering regime called \textit{critical regime} in Ref. \([9]\), in analogy with the theory of critical phenomena in condensed matter physics.

In can be concluded from the preceding arguments that the analysis of angular catalogs can be deceptive, as regards the strong clustering regime. To see this most simply, let us consider a three-dimensional mass distribution that may be strongly clustered and that is randomly sampled by equivalent points with a given mean number density. If this number density is so low that the sample does not show strong clustering at all, then its angular projection cannot show strong clustering either, in spite that the mean angular number density grows without limit with the depth of the three-dimensional sample. Therefore, very
deep angular galaxy catalogs may not be probing the strong clustering regime. The criterion to ensure that an angular catalog is probing the strong clustering regime is, of course, that the range of scales where \( w(\theta) \gg 1 \) must not be dominated by discreteness effects. In particular, if a non-trivial power law fit has been obtained in a range of scales, this range must include a subrange of the smaller scales where \( w \gg 1 \). Let us examine if the calculations of the two-point correlation function of angular positions of galaxies fulfill this criterion and to what extent.

Peebles [3, Fig. 7.2] shows log-log plots of \( w(\theta) \) for the Zwicky, Lick and Jagellonian catalogs. Only the shallow Zwicky catalog, with limiting apparent magnitude \( m = 15 \), has a range of small \( \theta \) with \( w(\theta) > 1 \). However, the absolute value of the slope grows in that range, possibly tending to two, the value that corresponds to \( \gamma = 3 \) and \( D_2 = 0 \), that is to say, to a distribution of isolated points. Something similar happens in the plot of apparent magnitude slices of the APM catalog [3, Fig. 7.3]. Analyses of modern and hence deeper catalogs seldom show any \( w > 1 \). For example, the analysis of the SDSS-DR7 angular two-point correlation function by Wang et al only has a very small range of \( \theta \) with \( w(\theta) > 1 \) for the apparent-magnitude slice \( 17 < r \leq 18 \) [22, Fig. 15]. Among the cited calculations of \( w(\theta) \), only the one for the Zwicky catalog allows us to speak of a strong clustering regime, although the range with \( w > 1 \) does not allow to fit a power law and, moreover, already shows discreteness effects.

It seems that the analyses of angular catalogs have favored depth, on account of the larger projected mean number density, without considering that the distant and hence more luminous galaxies have low mean number density in three dimensions and are, therefore, less suitable for studying the smaller scales. For example, the analysis of the SDSS-DR7 by Wang et al [22] starts at apparent magnitude \( r > 17 \), although there is a large number of galaxies of lower magnitude in the SDSS-DR7. These data can be used. Therefore, we undertake in Sect. 5 an angular analysis of the SDSS-DR7, taking into account especially galaxies with low apparent magnitudes. Moreover, we employ the stellar masses in the NYU-VAGC, which allow us to calculate the two-point correlation function of the stellar mass distribution.

Finally, let us notice that the log-log plot of \( w(\theta) \) directly gives \( \gamma \) (assuming \( \gamma > 1 \)), according to Eq. (4.1), but not \( r_0 \). To calculate \( r_0 \), it is necessary to estimate the radial selection function and carry out some computations [4, 5], introducing uncertainties.

5 Angular analysis of the stellar mass distribution

We employ the same data as in Ref. [18] and also use the apparent magnitude range \( 12.5 < m_r < 17.77 \). The angular ranges are defined in the same way, namely, in terms of equal area coordinates \( sl \) and \( f \), with ranges \(-0.7604 < sl < 0.7934 \) and \(-0.02269 < f < 1.1996 \). These ranges amount to a total solid angle \( \Omega = 1.554 \cdot 1.222 = 1.899 \) steradians, which contains 529967 galaxies.

In the angular rectangle so defined, we employ coarse multifractal analysis, using a lattice of cells like in Ref. [18] but without dividing the radial coordinate; that is to say, we use a \( 4 \times 3 \) basic lattice and a sequence of binary subdivisions. The corresponding cells have an aspect ratio close to one, which is convenient. Given that the size of the cells of the basic lattice is somewhat small to appreciate the behavior for large angles, we add two coarser lattices, namely, a \( 3 \times 2 \) lattice and a \( 2 \times 2 \) lattice.

The goal of multifractal analysis is to obtain the dimension spectrum \( D_q \) [or \( f(\alpha) \), alternatively]. For the angular distribution of stellar mass, as a projected mass distribution, we know that \( D_q \) is preserved from the three-dimensional distribution in the interval \( 1 < q \leq 2 \),
as long as $D_q < 2$ (Sect. 3). In general, $D_q$ is non-increasing with $q$. From redshift-space calculations, we have that $D_2$ is in the range 1.2–1.4 (Sect. 2) and that $D_q$, for diminishing $q$, reaches 2 for some $q > 1$. Therefore, the most interesting dimension is the correlation dimension $D_2$. This dimension can be obtained from a power-law fit of the two-point correlation function of the projected stellar mass distribution.

We first analyze the complete set of 529967 galaxies (a magnitude-limited sample) to calculate the coarse-grained statistical moment

$$\mu_2(\Omega) = \frac{\langle \rho_2^2 \rangle}{\langle \rho \Omega \rangle^2},$$

where $\Omega$ is the cell solid angle (the angular area). According to Eq. (4.1), the expected scaling is

$$\mu_2(\Omega) - 1 \propto \Omega^{(\gamma - 1)/2}.$$  

We calculate $\mu_2(\Omega)$ taking into account the stellar masses of the galaxies and also suppressing them. The latter calculation is equivalent to the calculation of the two-point correlation function of galaxy angular positions. For the stellar mass, we find an unexpected result that leads us to restrict the complete magnitude-limited sample. In this way, we can solve the problem and, in addition, obtain a significant fractal range, such that $\mu_2 \gg 1$.

5.1 Complete magnitude-limited sample

In Fig. 2 are the log-log plots of $\mu_2(\Omega)$ and $\mu_2(\Omega) - 1$ for the stellar mass and galaxy number distributions ($\Omega$ is normalized to the total solid angle). Evidently, the effect of considering the stellar mass is very significant. The fits have slopes $(\gamma - 1)/2 = 0.99$ (stellar mass) and $(\gamma - 1)/2 = 0.41$ (galaxy number). The first fit holds in a very large range of scales, which also includes a large range with $\mu_2 \gg 1$, but all of this is hardly useful: it just tells us that $D_2 = 0$, that is to say, that we have the dimension of a set of isolated point-like masses. The second fit holds in a smaller but substantial range and gives a reasonable $\gamma = 1.82$ (compatible with the results of Ref. [22]). However, the range with $\mu_2 \gg 1$, where $\mu_2$ and $\mu_2 - 1$ tend to coincide and which is therefore the fractal range, also corresponds to $D_2 = 0$. In contrast, the value of $\mu_2 - 1$ is well below one in the range where $\gamma = 1.82$, that is to say, this range corresponds to a quasi-homogeneous distribution. As regards the transition to homogeneity, the left-hand and right-hand plots suggest very different scales.

![Figure 2](image-url)  

**Figure 2.** Second moment $\mu_2(\Omega)$ and variance $\mu_2(\Omega) - 1$ for the stellar mass (left) and number distributions (right), with the corresponding fits ($\Omega$ normalized).
The effect of the galaxy stellar masses is so important because of the distribution of these masses along the radial coordinate. Indeed, let us compare how galaxy number counts and stellar mass are distributed with respect to redshift. Both quantities are roughly distributed in a similar way, namely, there are few galaxies and little stellar mass at low redshift, but both quantities grow to reach their maxima at some intermediate redshift and then decrease again. This behavior is a consequence of the geometry of the survey and the apparent-magnitude cut, which are such that the available volume grows with $z$ while the number density of visible galaxies diminishes. However, while the volume growth is indifferent to the stellar masses, the apparent-magnitude cut is not, because the most luminous and hence most massive galaxies are concentrated at high $z$. Therefore, the stellar mass reaches its maximum at higher $z$ and the mean mass per galaxy strongly grows at high $z$. The difference in the distributions of galaxy numbers and masses of our sample is shown in Fig. 3.

Now we can understand why we have found the value $D_2 = 0$ that corresponds to a set of isolated point-like masses: it is due to a reduced number of very massive galaxies at high redshift that are practically isolated. They have an overwhelming effect on the statistical moments of the stellar mass distribution but a small effect on the moments of the galaxy number distribution.

Since the multifractal analysis of the stellar mass distribution gives very reasonable results in redshift space (Sect. 2 and Ref. [18]), we expect that some modification of the complete sample may solve the problem. It is possible to reduce the apparent-magnitude cut, to try to avoid the predominance of isolated massive galaxies at high $z$, but this simple recipe will not work: the problem lies in the construction of any apparent-magnitude sample, because some very massive galaxies at high $z$ have low apparent magnitudes. We have tried cuts at $m_r < 16$ (keeping 53237 galaxies) and at $m_r < 15$ (keeping 13156 galaxies), and indeed the $D_2 = 0$ distribution along the full range of scales persists. To deal with the problem in a natural way, we can instead put a redshift cutoff.

5.2 Samples with redshift cutoffs

A redshift cutoff effectively removes massive galaxies with low number density, as can be deduced from Fig. 3. The plot in the right-hand side suggests us to cut at $z = 0.34$ (although the roughness of the plot for $z > 0.34$ diminishes with a larger bin size). This redshift cut removes from the complete sample only 245 galaxies, but they amount to a 6.5% of
mass and have a drastic effect on the statistical moments of the stellar mass distribution: as can be seen in left-hand side of Fig. 4, now the plots of $\mu_2$ and $\mu_2 - 1$ for the stellar mass distribution look like the plots for the galaxy number distribution in the right-hand side of Fig. 2. However, the slope of the fit is somewhat larger in absolute value, namely, $(\gamma - 1)/2 = 0.49$, giving $\gamma = 1.98$. Again, in the range of this fit, $\mu_2 - 1 \ll 1$, while a fit in the range where $\mu_2 - 1 \gg 1$ shows that $\gamma$ tends to three. Therefore, this projected mass distribution is quasi-homogeneous, with small fluctuations that are power-law correlated (see Fig. 5).

![Figure 4](image)

**Figure 4.** Second moment $\mu_2$ and variance $\mu_2 - 1$ for the stellar mass of two samples with redshift cutoffs: (left) cutoff at redshift $z = 0.34$, with a power-law fit giving $\gamma = 1.98$; (right) cutoff at redshift $z = 0.025$, with a power-law fit giving $\gamma = 2.03$.

The projected mass distribution is quasi-homogeneous because the sample is deep and the cosmic web features are obliterated in the projection, as explained in Sect. 4. Of course, we can prepare shallower samples with lower redshift cutoffs. For this purpose, Fig. 3 offers no guide, but we can appeal to a heuristic argument: if we want to preserve the cosmic web features, we should not go too deep into the homogeneous regime. As said in Sect. 4, the criterion for a sample of equal-mass particles is that the mean distance between projected uncorrelated points is larger than the minimal distance between correlated points. However, it is hard to translate this qualitative criterion into a quantitative one, especially, for unequal masses. As $30 h^{-1}$Mpc is a safe scale for homogeneity (Sect. 2), we can try some small multiple of $z = 0.01$, e.g., $z = 0.025$. The corresponding sample contains 18168 galaxies. The plots of $\mu_2$ and $\mu_2 - 1$ are shown in the right-hand side of Fig. 4 and demonstrate a notable improvement.

Indeed, this sample is not quasi-homogeneous, because the power law holds where $\mu_2 \gg 1$, that is to say, in the strong clustering regime. In this case, the projected mass distribution is not featureless and one can perceive in it the cosmic web structure (Fig. 5). Nevertheless, the fit for $\mu_2 - 1$ gives $\gamma = 2.03$, almost the same as in the preceding case. As this value holds for strong clustering, we can interpret that it gives the correlation dimension $D_2 = 3 - \gamma \simeq 1$, which is smaller than expected. But we have to notice, like in Sect. 2, that it is necessary to consider, in the fractal regime, the convergence of the respective fits for $\mu_2 - 1$ and $\mu_2$. Unfortunately, $\mu_2$ hardly has a scaling range, and if we were to try a fit, it would give a considerably smaller value of $\gamma$ and, therefore, higher $D_2$. This would be fine, but the consequent uncertainty in $D_2$ leads us to conclude that the analysis of Sect. 2 in redshift space is preferable.

Besides the correlation dimension, we are interested in the scale of transition to homo-
geneity. We can find the angular scale of homogeneity by solving the equation \( \mu_2(\Omega) - 1 = 1 \). In the case of magnitude-limited samples, we can relate that angular scale to \( r_0 \) by employing Eq. (4.1), although the relation is not straightforward. When we introduce a redshift cutoff, the relation is more involved, because the constants \( K \) and \( d_* \) in Eq. (4.1) need to be reinterpreted. At any rate, we can compare the solution of the equation \( \mu_2(\Omega) - 1 = 1 \) for either the galaxy number distribution of the complete sample or the \( z = 0.34 \) stellar mass distribution (which only misses a few galaxies). The two solutions are \( \Omega = 3.1 \cdot 10^{-6} \) and \( 7.2 \cdot 10^{-6} \), respectively. The significantly larger value for the stellar mass distribution shows that the corresponding value of \( r_0 \) is larger as well.

6 Regarding the calculation of \( \sigma_8 \)

Although we have shown that the value of \( r_0 \) grows when we consider the galaxy masses, it is difficult to quantify this growth. Any procedure of calculating \( r_0 \) from a magnitude-limited sample with a redshift cutoff has to be complex and, moreover, we would not know exactly what cutoff to use. Therefore, it seems doubtful that one can employ angular projections to obtain a precise value of \( \sigma_8 \).

The alternative is to use the redshift information (provided that it is available) to construct volume limited samples. With a volume limited sample, for example, we can follow the procedure in Sect. 2 to obtain a range of \( v_0 \). Then, we know that \( r_0 \) is proportional to \( v_0^{1/3} \), although the proportionality factor is difficult to calculate, because the cells are not spherical but have a complex shape. Of course, one can use spherical cells, at the cost of losing some data. At any rate, a volume limited sample such as VLSI only contains a low range of galaxy masses [18, Table 1], which may not be sufficiently representative of the general population to obtain a precise value of \( r_0 \). If one tries to analyze deeper volume limited samples, then the scaling is poorer (as explained in Sect. 2).

The results of the analysis of angular projections still are the only support for a power law Eq. (1.1) holding in the quasi-homogeneous regime, that is to say, for \( r \gg r_0 \). Therefore, angular projections can give better information than volume limited samples in that regime. As \( \sigma_8 \) is the linear-theory mass dispersion, the quasi-linear regime is the right regime to calculate it. In fact, it is currently calculated from the analysis of a combination of several
types of data [2]. At any rate, a calculation of $\sigma_8$ with a relative error of about one percent is perhaps beyond our present capabilities.

7 Conclusions and discussion

The general conclusion, of course, is that we must question the standard description of galaxy clustering based on the power-law correlation function with canonical exponent $\gamma = 1.8$ and clustering length $r_0 \simeq 5h^{-1}\text{Mpc}$, which has remained unaltered for almost 50 years. These values were initially obtained by means of the analysis of the galaxy-galaxy angular correlation function and still now the largest scaling ranges are obtained with this method. However, we have shown that the results change after introducing the stellar masses of galaxies and therefore studying the stellar mass distribution. This distribution is in fact a proxy for the distribution of baryonic matter, which is very relevant in physical cosmology. The galaxy number distribution is obviously less relevant.

The striking result is that the angular correlation function of the projected stellar mass distribution obtained from a magnitude-limited sample is dominated by a small number of the largest masses and, therefore, has $\gamma = 3$ ($D_2 = 0$). We have shown this for the complete SDSS-DR7 sample (with a convenient angular window). Furthermore, it also holds for magnitude-limited subsamples of it.

As there is no way to avoid the effect of large galaxy masses in the analysis of magnitude-limited samples, one has to find a way to remove them. An analysis of the distribution of galaxy masses with redshift (which, of course, requires redshift data) shows that the largest masses are naturally concentrated in the higher redshift interval. Therefore, one can remove this interval; in particular, we find it natural to remove the interval $z > 0.34$. This implies, in fact, the removal of a very small number of galaxies but it recovers a power law with a reasonable value of $\gamma$; namely, we obtain $\gamma \simeq 2$, which is somewhat larger than the value with the mass information suppressed. It also yields a larger value of $r_0$, which is difficult to estimate. These larger values can be explained as the effect of many relatively massive galaxies with $z < 0.34$ that still remain in the sample.

We have also used a very low redshift cutoff, namely, $z = 0.025$, which removes the majority of galaxies and allows us to probe the strong clustering regime. Indeed, the cosmic web structure is preserved in the projection. Therefore, we have access to fractal properties and we can speak of the correlation dimension, which is calculated as $D_2 = 3 - \gamma$. The power law fit of the scaling of $\delta M/M$ yields $\gamma \simeq 2$, as in the case of the $z = 0.34$ cutoff, but now with a scaling range in the strong clustering regime, $\delta M \gg M$. However, this range is hardly sufficient and the value of $D_2$ is quite uncertain. If we trust that the same scaling will be extended to higher values of $\delta M/M$, when we have more data (larger apparent magnitudes), then we can ascertain that $D_2 \simeq 1$.

There is nothing wrong with $D_2 \simeq 1$, in principle. But it disagrees with other estimations also based on the stellar mass distribution. For example, the volume limited sample VLS1 that is analyzed in Sect. 2 yields the range 1.20–1.40. However, VLS1 only contains galaxies with small masses, whereas the restriction of the full sample to $z = 0.025$ leaves considerably larger masses. At any rate, we see a more fundamental problem with $D_2 = 1$. The multifractal analyses of the stellar mass distribution in SDSS-DR7 and of the dark matter distribution in cosmological simulations show that the minimum value of the local dimension is $\alpha_{\text{min}} = 1$, that is to say, the value at which the gravitational potential diverges [18, 19]. On the other hand, it is easy to prove, for a general multifractal, that $D_q > \alpha(q)$, if $q > 1$. Therefore, $D_2 > \alpha(2) > \alpha_{\text{min}} = 1$, in the case of the cosmic web.
As regards the transition to homogeneity, our general conclusion is that the introduction of galaxy masses pushes $r_0$ to larger values. Unfortunately, we have found no way of making reliable estimations from the analysis of the projected mass distribution. In redshift space, the analysis of VLS1 in Sect. 2 yields the range $[9.5, 12.5] \text{ Mpc}/h$ for the length scale $v_0^{1/3}$, which is a measure of the scale of transition to homogeneity. Presumably, deeper volume limited samples give larger values, but the uncertainty of the values must grow.

As a final reflection, we can wonder why one obtains very sensible results from the projected distribution without galaxy masses; in fact, more sensible than with galaxy masses. This in contrast with the results from volume limited samples; for example, the comparison of the multifractal spectra of VLS1, with or without masses, shows that the former is more reliable [18]. Surely, what we need is a systematic procedure of removing a definite set of the largest masses from a magnitude-limited sample, based on first principles and, preferably, without invoking redshift data.

One way or another, we can conclude that “precision cosmology” is not yet as precise as we would like it to be.

Acknowledgments

I thank C.A. Chacón-Cardona for the SDSS-DR7 file.

References

[1] B.J.T. Jones, Precision Cosmology: The First Half Million Years, Cambridge University Press (2017).
[2] O. Lahav and A.R. Liddle, The Cosmological Parameters, in C. Patrignani et al. (Particle Data Group), Chin. Phys. C 40 (2016) [2017 update] 100001.
[3] P.J.E. Peebles, Principles of Physical Cosmology, Princeton University Press (1993).
[4] H. Totsuji and T. Kihara, The Correlation Function for the Distribution of Galaxies, Publ. Astron. Soc. Japan 21 (1969) 221.
[5] E.J. Groth and P.J.E. Peebles, Statistical analysis of catalogs of extragalactic objects. VII - Two- and three-point correlation functions for the high-resolution Shane-Wirtanen catalog of galaxies, Astrophys. J. 217 (1977) 385–405.
[6] B.B. Mandelbrot, The fractal geometry of nature (rev. ed. of: Fractals, 1977), W.H. Freeman and Company (1983).
[7] L. Pietronero, The fractal structure of the universe: Correlations of galaxies and clusters and the average mass density, Physica A 144 (1987) 257–284.
[8] P.H. Coleman and L. Pietronero, The fractal structure of the Universe, Phys. Rep. 213 (1992) 311–389.
[9] J. Gaite, A. Domínguez and J. Pérez-Mercader, The fractal distribution of galaxies and the transition to homogeneity, Astrophys. J. 522 (1999) L5–L8.
[10] B.J. Jones, V. Martínez, E. Saar and V. Trimble, Scaling laws in the distribution of galaxies, Rev. Mod. Phys. 76 (2004) 1211–1266.
[11] F Sylos Labini, NL Vasilyev and YV Baryshev, Power law correlations in galaxy distribution and finite volume effects from the Sloan Digital Sky Survey Data Release Four, Astron. Astrophys. 465 (2007) 23–33.
[12] P. Sarkar, J. Yadav, B Pandey and S. Bharadwaj, The scale of homogeneity of the galaxy distribution in SDSS DR6, MNRAS 399 (2009) L128–31.
[13] A.O. Verevkin, Y.L. Bukhmastova and Y.V. Baryshev, The non-uniform distribution of galaxies from data of the SDSS DR7 survey, Astron Rep 55 (2011) 324–40.

[14] C.A. Chacón-Cardona, R.A. Casas-Miranda, J.C. Muñoz-Cuartas, Multi-fractal analysis and lacunarity spectrum of the dark matter haloes in the SDSS-DR7, Chaos, Solitons and Fractals 82 (2016) 22–33.

[15] F. Sylos Labini, M. Montuori and L. Pietronero, Scale invariance of galaxy clustering, Phys. Rep. 293 (1998) 61–226.

[16] M.J. Geller and J.P. Huchra, Mapping the Universe, Science 246 (1989) 897–903.

[17] L. Kofman, D. Pogosyan, S.F. Shandarin and A.L. Melott, Coherent structures in the universe and the adhesion model, The Astrophysical Journal 393 (1992) 437–449.

[18] J. Gaite, Fractal analysis of the large-scale stellar mass distribution in the Sloan Digital Sky Survey, JCAP 2018, 07, 010.

[19] J. Gaite, The Geometry and Formation of the Cosmic Web, Galaxies, submitted [arXiv:1810.02311].

[20] K.N. Abazajian et al, The Seventh Data Release of the Sloan Digital Sky Survey, ApJS 182 (2009) 543–558.

[21] M.R. Blanton et al, New York University Value-Added Galaxy Catalog: A Galaxy Catalog Based on New Public Surveys, AJ 129 (2005) 2562–2578. NYU Value-Added Galaxy Catalog web page.

[22] Y. Wang, R.J. Brunner and J.C. Dolence, The SDSS galaxy angular two-point correlation function, MNRAS 432 (2013) 1961–1979.

[23] G. Kauffmann et al, Stellar Masses and Star Formation Histories for 10^5 Galaxies from the Sloan Digital Sky Survey, MNRAS 341 (2003) 33–53.

[24] M.R. Blanton and S. Roweis. K-Corrections and Filter Transformations in the Ultraviolet, Optical, and Near-Infrared, AJ 133 (2007) 734–754.

[25] K. Falconer, Fractal Geometry (Second Edition), John Wiley and Sons, Chichester, UK, (2003), Chapter 17.

[26] K. Falconer, J.M. Fraser and Xiong Jin, Sixty years of fractal projections, in Fractal Geometry and Stochastics V, Progress in Probability 70, Birkhauser (2015), pg. 3–25 [arXiv:1411.3156].

[27] R. Durrer, J.-P. Eckmann, F. Sylos Labini, M. Montuori and L. Pietronero, Angular Projections of Fractal Sets, Europhysics Letters 40 (1997) 491–496.

[28] B.R. Hunt and V.Y. Kaloshin, How projections affect the dimension spectrum of fractal measures, Nonlinearity 10 (1997) 1031–1046.