

An attempt to check homogeneity at large scales by using galaxy samples

B. Lukács and N. Fáy

CRIP RMKI, Bp. 114, Pf. 49, 1525 Budapest, Hungary

Received 21 December 2006 / Accepted 9 March 2007

ABSTRACT

Aims. A statistical method is applied to galaxy samples to search for possible repetitions on 100 Mpc scales.

Methods. The method is based on autocorrelations, and is constructed in such a way that it does not detect structures of changing scales.

Results. The signal for non-randomness is moderate but significant. This result argues for homogeneity well above 300 Mpc scales.

Key words. large-scale structure of Universe – distance scale – cosmological parameters

1. Introduction

Astronomical observations of the last 20 years have shown that the Universe has a structure on large scales i.e. even at scales exceeding 10 Mpc. The bubble structure outlined by the distribution of galaxies indicates a scale of the magnitude of 100 Mpc. Thus a question has arisen as to what scales (if any) does the Universe become homogenous, in accordance with the deeply embedded principle of cosmology. Provided that the distribution of the bubble walls is not random but recurring, albeit with a spread along the characteristic length, we determine a regular distribution, i.e. a distribution homogenous at scales well above 100 Mpc.

Our goal is to check the very large-scale symmetries of the Universe at scales larger than the largest local structures. The largest bound structures are thought to exist at 30–50 Mpc, the “walls” exist at 130 Mpc scales, and homogeneity may appear from, say, 300 Mpc. Thus, the decisive scales are $r > 300$ Mpc, $z > 0.1$.

Thus, for homogeneity, it is not sufficient that no structure is observed at larger scales. Even fully disordered matter is compatible with an inhomogeneous space-time structure. Exact solutions demonstrating this are rare because people seldom look for nonsymmetric solutions, however, see e.g. the Kustaanheimo solution in Kramer et al. (1980), where matter has a homogeneous isotropic energy density, but the metric is inhomogeneous. For the opposite case, when the matter has less symmetry than the metric, see Lukács & Perjés (1977) where the metric is time-independent, but the electromagnetic field tensor is not (while T_{ik} is).

Still, for fluid solutions the relation between symmetries of the metric and those of the material fields is rather strong. In addition, it is much easier to detect asymmetries in the density than in the metric.

Astronomers have been rather sceptical about the existence of regularity even at 100 Mpc scales. Now evidences are accumulating for such more or less regular structures, see e.g. Saar et al. (2002) and citations therein. Our results indicate that the degree of regularity strongly depends on the actual cosmology, so on Ω . The regularity, at least that detected by our method,

seems strongest for $0.2 \leq \Omega \leq 0.4$, and, while some regularity is still detectable at $\Omega = 1$, individual samples show very little significant regularity, with the exception of the Broadhurst sample in the GP directions (Broadhurst et al. 1990).

Here we report some results of a statistical analysis performed on SDSS and deep pencil beam galaxy samples. Section 2 gives the results for random and semi-random sequences, and Sect. 3 defines the true samples. Section 4 evaluates the samples with the result that they are significantly non-random, but some repetitive patterns are present. We do not know the mechanism behind this non-random pattern, but the existence of it enables us to deduce something about the homogeneity of the space-time at scales >300 Mpc (see Sect. 5). Section 6 provides our conclusions.

2. A statistical procedure to detect regularities: the tests

The direct way to decide homogeneity vs. inhomogeneity well above 100 Mpc scales would be to collect statistically rich galaxy counts much beyond the said scale, more or less equally populated at different distances and in several directions. With such samples, simple autocorrelation calculations would suffice. However, such ideal samples are unavailable. We would need pencil-beam surveys because we cannot expect isotropic structures centered on the Earth and otherwise a 3-dimensional statistical analysis would be needed. Pencil-beam surveys are available, but deep pencil-beam surveys were performed in the first half of the 90s and then attention shifted elsewhere. Even these deep pencil-beam samples are not homogeneously populated: being pencil-beams, the near parts of them show a z^2 tendency in density, while the far parts are less and less populated, losing the fainter galaxies. In most directions no deep survey is available at all.

However, even non-optimal samples in large enough number treated with proper statistical methods may be sufficient using an indirect approach. We have elaborated a statistical procedure based on auto-correlation but providing an outlook beyond the limits of simple autocorrelation (Fáy & Lukács 2006). It

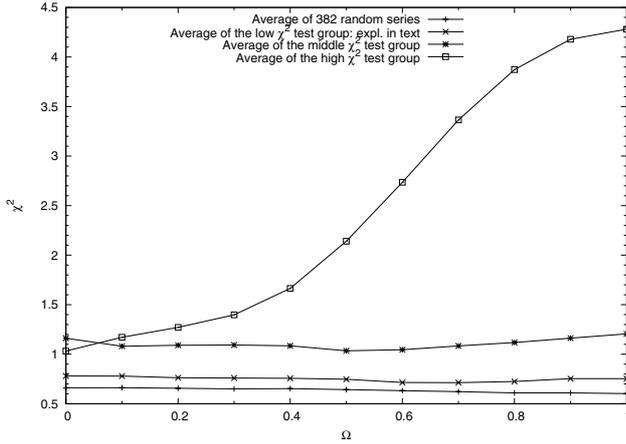


Fig. 1. Self-tests of the statistical procedure mentioned in Sect. 2. The average of truly random test results for each Ω is represented by the grey line with diamonds. The test results with recurrences are classified into 3 classes according to their maximal χ^2 over the whole Ω range: the upper boundary of the first class and the lower one of the second have been chosen as the 95% significance border for an individual sample. Even the curve of the second class (the first one above the significance border for one Ω ; middle triangles) shows a preference for $\Omega = 1$ (where the repetition as well as its dampening was created by the spreads in repetition frequency); as expected, the preference is rather explicit for Class 3 (large triangles).

gives information on the repetitive nature of these structures, and can search for the most regular models among the different cosmological ones. The procedure results in a χ^2 -type quantity (called χ^2 ; the exact definition can be found in Fáy & Lukács 2006) for the examination of the specific distribution in a specific cosmology. Based on self-tests, whenever the significance fulfils $0.37 < \chi^2 < 1.072$, the regularity of the specific sample is not significant. (The boundaries are the averages of the cosmology-dependent boundaries in Fáy & Lukács (2006). If the specific cosmology is selected, the Ω -dependent significance limits should be used.) Above 1.072 the distribution has regular features, detected by the method. Results below 0.37 would be difficult to interpret, but it would be unreasonable to conclude on their regularity; see e.g. Jánossy (1965). For the boundaries see the details in Fáy & Lukács (2006). We have tested our statistical procedure in cases where the test sample is by construction regular with recurrences in the cosmology $\Omega = 1$, albeit with some prescribed spread along the period length. As expected, increasing the spread the value of χ^2 decreases; thus it detected the standard cosmology as the most regular from among the set of cosmologies with greater and greater tolerance. Creating the test sequences at $\Omega = 1$ but evaluating them in all the 11 specific cosmologies we distinguished three ranges: $\chi_{\max}^2 < 1.072$ (this is when the relative spread is approximately from 30% to 50%); $1.072 < \chi_{\max}^2 < 3$ (this roughly corresponds to spreads from 20% to 30%); $3 < \chi_{\max}^2 < 6$ (relative spreads approximately from 10% to 20%). Here χ_{\max}^2 means maximum in the range $0 \leq \Omega \leq 1$. Having classified in such a way the semi-random test series, the class averages of them and the truly random ones are shown in the 4 curves in Fig. 1. The dependence on Ω of the averages demonstrates the sensitivity of the method in selecting the true cosmology $\Omega = 1$ (i.e. where the repetition was created), even in the second range, just above the significance level. This preference for the standard cosmology is even more explicit in the third range.

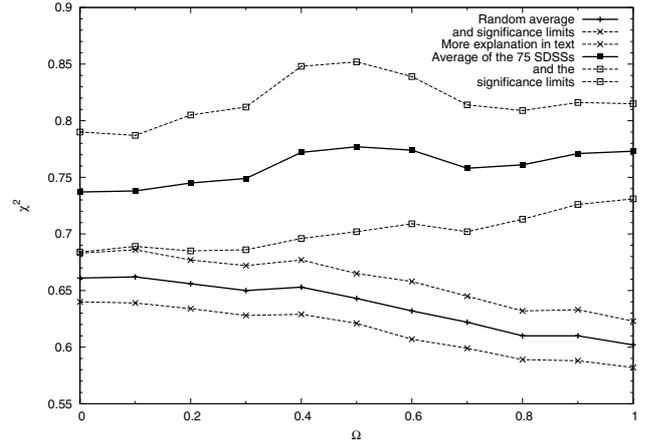


Fig. 2. The average of the χ^2 values for the 75 SDSS directions for 11 cosmologies, compared to the corresponding averages of 382 randomised samples, the 95% significance limits indicated (see in Fáy & Lukács 2006).

We summarize some results of Fáy & Lukács (2006) concerning the method and significance in the Appendix; for details see the original paper.

3. Samples

From the ongoing large scale projects mapping the entire sky, we evaluated the 2005 results of the Sloan Digital Sky Survey (SDSS). Since the length of the sample is important and the technique made it possible to duplicate the length by uniting 2 samples if on a straight line, we used those samples whose opposite pairs existed too. 150 samples (of the 826 original ones) met this condition, most in the equatorial plane and the remainder not farther than 10 degrees, so we used 75 double samples. The distribution of galaxies can be considered well populated only until $z = 0.2$, so the doubling was necessary.

A few measurements are available where the samples were deep enough, even if unilateral, such as that measuring into 4 directions (SA57, SA68, Lynx and Hercules) (Bershady et al. 1997) and that on narrow angle samples measuring in the direction of Sculptor (Bellanger & de Lapparent 1995).

There are two additional bilateral pencil beam samples of high importance obtained in the 1990s (Broadhurst et al. 1993; Broadhurst et al. 1990; Szalay 1999). These observations were performed in the direction of the galactic poles and in the direction of SA68.

4. Results for regularity

We used 75 double artificial pencil-beam samples created from the SDSS data as described above. These 75 double samples were not pre-selected in any other way than the applicability for forming double pencil beams. The total of the 75 samples is not random in the sense of our statistical test even if most individual samples are not significantly non-random.

We have calculated for each individual Ω value the average of the 75 χ^2 values of the samples of SDSS and compared the averages and their mean errors to the averages and mean errors obtained from the purely randomised process runs. (See Fig. 2, showing the significance limits around the averages; for the tabulated significance limits see Fáy & Lukács (2006).) The result is that the average of the χ^2 s of all the 75 samples is significantly

Table 1. A statistical processing of the 18 samples (of the 75 SDSS directions; identified with right ascensions & declinations) for which χ^2_{\max} is above the significance limit. *: weakly expressed.

Direction pair	χ^2 maximum	Related Ω	Characteristic size of ord. arrangement expressed in Mpc at $\Omega = 0.3$
347 0-167 0	2.809471	0.4	220
8 0-188 0	1.751462	0.2	185
341 0-161 0	1.625981	0.5	205
49 0-229 0	1.620588	0	235
6 0-186 0	1.619298	0.2	130 and 220
355 0-176 0	1.506357	0.6	215 *
350 -9-169 9	1.365507	1	140
39 0-219 0	1.347519	0.4	210
327 -7-148 7	1.309114	0	200
339 -9-159 8	1.248024	0.8	140 *
21 0-201 0	1.220668	0.4	185
335 -8-155 8	1.209302	0.9	220
319 -7-140 7	1.172419	1	140
331 0-151 0	1.146198	0.6	190 and 235 *
329 0-149 0	1.130218	0.8	185 *
51 0-230 0	1.115163	0	190
329 -8-148 7	1.109202	0	200
144 0-324 0	1.092834	0.3	185

higher than that of the random test series, at least for $\Omega \geq 0.1$. So the totality of the investigated Sloan samples is not random but shows some repetitive trends. Ω is expected somewhere near to 0.3. For $\Omega = 0.3$ the average of the 75 χ^2 values is 0.748, while for random series this value would be expected between 0.628 and 0.672 at a 95% significance level.

We categorised the 75 $\chi^2(\Omega)$ functions into the same 3 classes as described in Sect. 2 and obtained as follows: 57 distributions have been classified into the range below significance level, 18 distributions into range 2 above the significance level and no SDSS distribution was classified into range 3. A few details of the 18 significantly non-random samples are described in Table 1.

The present statistical results + Occam's Razor would enable us to say something about Ω even if its value were not known elsewhere. We use only those samples that seem to be significantly non-random individually; the statistical significance of this second evaluation is not totally quantitative. Investigating the Ω -dependence of the average of the 18 χ^2 values in range 2 (cf. Fig. 3) a maximum point occurs at $\Omega = 0.4$ (and in itself the highest regularity would single out this value, not too far from the recently accepted one).

Now let us turn to the results of the 5 unilateral samples in a similar way. There, 4 samples show regularity. Some aspects of these results are summarised in Table 2. The average of the χ^2 values vs. Ω shows a maximum at $\Omega = 0.3$ (cf. Fig. 3) which is near the Ω value preferred e.g. by the comparison of the BOOMERANG, CHANDRA and far SN observations (Allen et al. 2002).

The bilateral pencil beam measurements show higher maxima. Both of the samples are classified into range 3. Table 3 shows the results while Fig. 4 represents the result of averaging, indicating a maximum at $\Omega = 0.2$ and 0.3, again in the dynamically preferred range.

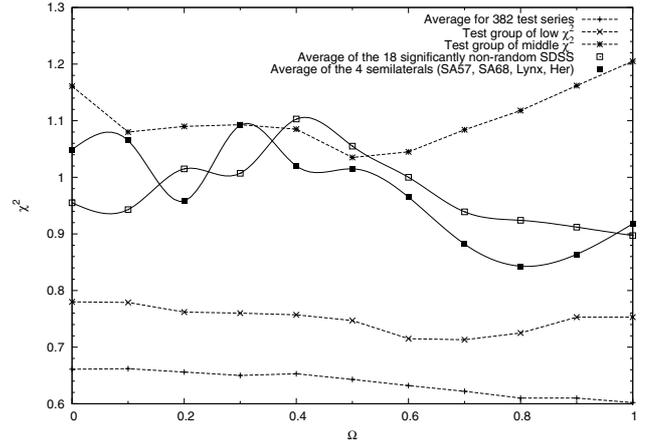


Fig. 3. The averages of the χ^2 values for 11 cosmologies. Symbols are as follows: thick black line with heavy dots for the 18 SDSS samples exceeding the significance level compared to the related self-test results; gray line with squares the average calculated from the 4 semi-lateral distributions exceeding the significance level (Hercules, Sculptor, Lynx and SA57) giving a maximum of χ^2 at the cosmology of $\Omega = 0.3$. The other 3 lines are the test results for comparison, the same as Fig. 1. The SDSS averages of the significant 18 samples provide a maximum at $\Omega = 0.4$.

Table 2. A statistical processing of the 4 semi-lateral deep pencil-beam samples (of 5), for which χ^2_{\max} is above the significance limit. *: weakly expressed.

Direction pair	χ^2 maximum	Related Ω	Char. size in Mpc at $\Omega = 0.3$
Lynx 130.25 44.75	1.183244	0	110 *
SA57 196.5 29.6	1.429499	0.5	150 and 170 *
Hercules 259.5 50	1.253967	0.1	130
Sculptor 5.25 -30.1	1.118754	0.3	100 and 135 *

Table 3. A statistical analysis of the 2 bi-lateral pencil-beam galaxies. The highly significant non-random character of them is conspicuous compared to the other two classes. Our opinion is that the qualitative difference in significance (compared to the samples used for Tables 1 and 2) is caused by the more uniform population and the much larger depth of the GP and SA68-anti-SA68 samples.

Direction pair	χ^2 maximum	Related Ω	Char. size in Mpc at $\Omega = 0.3$
Galactic poles	5.390166	0.8	140
SA68 3.75 15.5	5.212735	0	130

5. On the homogeneity of space

The above results can be employed to draw suggestions for the large-scale spatial symmetries of spacetime. The Einstein equations establish a strong relation among symmetries of matter and those of geometry.

In our case we concentrate on homogeneity, and specifically on homogeneity of geometry. Pencil-beam samples are inappropriate for checking homogeneity of the density, and other samples are not deep enough. Going farther in z (or distance) we are continuously losing the fainter galaxies, partly because they cannot be observed, partly because the observer selects brighter objects. It would be very assumption-dependent to try to correct this bias.

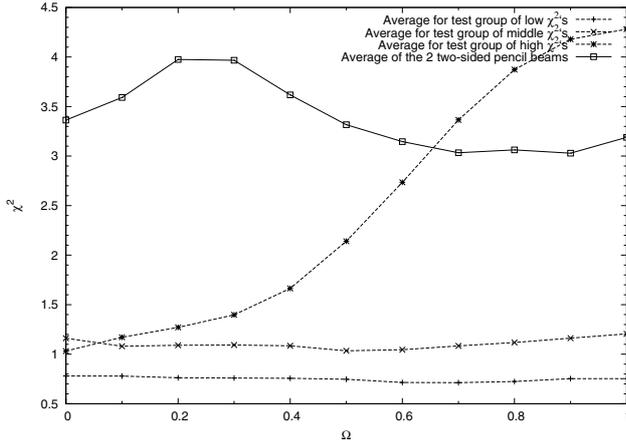


Fig. 4. The averages of the χ^2 vs. Ω of the 2 bilateral pencil-beam samples (GP and SA68-antiSA68) thick black line with heavy dots; compared to the related self-test results (thin lines; as in Fig. 1). The self-test results prefer the standard cosmology, while the average of the bilateral samples has a substantial maximum between $\Omega = 0.2$ and $\Omega = 0.3$.

However, if a definite structure is observed (so, if the galaxy distribution is inhomogeneous) at some scale which is still substantially smaller than the length of the samples, then we can check the behaviour of this scale of inhomogeneity; and while this signal is somewhat indirect for the geometry, the inevitably distance-dependent population of the samples does not falsify the check (but, of course, amplifies the noise).

Obviously now an approx. 140 Mpc inhomogeneity scale is detected, and nobody seriously considers homogeneity below 300 Mpc. Our analysis shows a repetition in the pencil-beam galaxy samples, which means a constant scale length of repetition (the method has been so constructed that it would not detect a repetition with variable scale). Then we must clarify 2 questions: is this constant repetition length an artefact created by a false cosmological constant Λ ; and if the constancy of the repetition length is real, what does it say for the large-scale homogeneity of the geometry?

For Question 1) we can tell that via Occam's Razor it is more prudent not to assume that a nonhomogeneous Universe mimics a more symmetric one with a false cosmological constant. Of course, a monotonously changing length scale could seem approximately constant by using another Ω , but Fig. 2 demonstrates that the average of the χ^2 's of all the 75 SDSS samples significantly differs from that of the random sequences again independently of Ω (excepting a marginal overlap at the 95% level for $\Omega = 0$). Table 1 of the 18 individually significant "repetitive" SDSS samples gives highly diverse Ω s for the maximal "regularities", and this in itself might be suspicious. However, Fig. 3 shows that the average of the χ^2 's of the 18 samples is much higher than those of the random sequences for any Ω , and the difference is maximal about $\Omega = 0.4$, which is not an improbable cosmology at all. For Question 2) see the argumentation below.

Spacetimes filled with fluid generally tend to inherit the symmetry of matter. While the question cannot be treated as a generality, let us take a spacetime filled with fluid, where $k = 0$ and both the matter and the geometry have six spacelike conformal symmetries. Then

$$ds^2 = e^{2\nu(u,r,\theta,\phi)} \{-du^2 + dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2)\} \quad (5.1)$$

$$\rho = \rho(u, x, y, z) \quad (5.2)$$

$$p = p(u, x, y, z) \quad (5.3)$$

$$u^i = e^{-\nu}(1, 0, 0, 0) \quad (5.4)$$

where u stands for the timelike coordinate. If, for simplicity, we assume that the isotropy is 3 true Killing symmetries, then the conformal factor can depend only on u and r . The remaining Einstein equations were evaluated in Lukács & Mészáros (1984). If ρ and p do not depend on r , then we have an alternative.

Either $\nu' = 0$, or not. The first case is the usual homogeneous and isotropic Robertson-Walker line element. In the second case there is only one nonhomogeneous solution, for which

$$e^{2\nu(u,r)} = [1/2(r^2 - u^2) + Au + B]^{-2} \quad (5.5)$$

where $A = 0$ can be achieved by a trivial shift in u . However even this only nonhomogeneous solution gives

$$\rho + p = 0 \quad (5.6)$$

$$\rho = \text{const.}, \quad (5.7)$$

which is a vacuum solution with a Λ term, not a fluid solution. For any fluid solution, conformal Robertson-Walker symmetries with homogeneous energy density and pressure lead to true homogeneity of the geometry.

This strongly suggests that at large scale (where direct observations are difficult) a homogeneous matter distribution would lead to homogeneous geometry, and our repetition with constant scale suggests homogeneity at larger scale.

6. Conclusions

The criteria of our selection of SDSS samples was independent of any nontrivial internal property of the samples; we used all data points having an opposite counterpart. The resulting set of samples is significantly non-random, so something in the Universe seems to have a nontrivial structure on a 140 Mpc scale. But then a constant length of scale seems to belong to this structure, which suggests homogeneity on larger scales. The first statement should be checked with more data. Other statistical studies using different methods confirm this statement. Saar et al. (2002) detect regularity with a 120–140 Mpc characteristic length at $\Omega = 0.3$ (and earlier Einasto et al. (1997a) or Einasto et al. (1997b) were able to detect regularity even at $\Omega = 1$, with 115–120 Mpc repetition length). So we think these results mutually confirm each other.

The effective lengths of the available (double) SDSS samples are not even an order of magnitude above the scale of the nontrivial structure; however the GP and SA68 samples are deeper and they show even clearer structures. According to the result of the previous section, such results are compatible with a Universe whose geometry is homogeneous at larger scales, but would be hard to be interpreted in a nonhomogeneous one. Therefore we think that the evaluation of the pencil-beam galaxy samples has yielded an observational argument for large-scale homogeneity.

For this non-trivial structure of constant length scale the signal is strongest for $0.1 < \Omega < 0.5$. Note that similar Ω values also have been deduced from different cosmological arguments in recent years.

The recurrence pattern is rather complicated: we detect 2 scales, which, at $\Omega = 0.3$, are between 130 and 140 Mpc and 185 and 235 Mpc. Since our investigation is purely statistical, we do not hypothesize about the origin of this double pattern. Saar et al. (2002) detect the first mode at the same length; the second one they sometimes detect but not in all samples,

while we never detected it in the deep unilateral samples and it is generally larger in both studies.

In 1997, when Einasto et al. (1997a) were able to detect statistical regularity in whole-space galaxy cluster distributions, the result was somewhat unexpected. Figure 2 demonstrates that the regularities are much more explicit with the new, moderate, values of Ω . While at $\Omega = 1$ all SDSS data are significantly regular, no individual sample is, nor is any of the deeper pencil-beam samples, excepting the GP of Broadhurst et al. (1990). As can be seen from Fig. 3, regularities of individual samples are clearer at $\Omega = 0.3$.

Acknowledgements. We acknowledge use of the SDSS (SLOAN) galaxy survey data.

Funding for the Sloan Digital Sky Survey (SDSS) has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Aeronautics and Space Administration, the National Science Foundation, the U.S. Department of Energy, the Japanese Monbukagakusho, and the Max Planck Society. The SDSS web site is <http://www.sdss.org>

The SDSS is managed by the Astrophysical Research Consortium (ARC) for the Participating Institutions. The Participating Institutions are The University of Chicago, Fermilab, the Institute for Advanced Study, the Japan Participating Group, The Johns Hopkins University, Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, University of Pittsburgh, Princeton University, The United States Naval Observatory, and the University of Washington.

We acknowledge the valuable technical help of Zs. Regály (Konkoly Observatory, Budapest).

Appendix A: Observation of recurrences

The details of the method are described in Fáy & Lukács (2006).

The samples are 1 dimensional, meaning a pencil-beam sample, of galaxies vs. redshifts z . If the samples were originally wider than cca. 1 degree, we form more narrow samples.

Now we choose a definite cosmology, i.e. an individual Ω and calculate the comoving coordinate x from redshift z . Evaluations are done separately for different Ω values, so “improbable” ones do not disturb the results of probable ones. We obtain the autocorrelation function $r(\Delta)$. Of course $r(0) = 1$, and later substantial r values indicate some recurrences/repetitions; however $r(\Delta)$ is not yet a unique number but a function, so it is not easy to find a criterion as a signal of repetition. To arrive at such, we take a tentative repetition length D , and form the averages

$$r_{av}(x) = 1/(n-1) \sum_{i=2}^n r(x + (i-1)D) \quad (\text{A.1})$$

the relevant quantity is $r_{av}(0)$. In (A.1) the summation starts from $n = 2$ to remove the uninformative and automatic high value $r(0) \equiv 1$. If $r_{av}(0)$ is high, then both the autocorrelations are substantial and the repetition distance is real. This D is an output value and it is given in the last columns of Tables 1–3.

However the autocorrelation can be calculated only if the data have been ordered first into bins: the number of bins is arbitrary and according to our numerical experience $r_{av}(0)$ is very sensitive on the chosen bin number. Therefore we calculate $r_{av}(0)$ for many different bin numbers and then average again.

Now $r_{av}(0)$ is the average formed from autocorrelation coefficient(s), so if the original sample was uncorrelated then the expected value of $r_{av}(0)$ is 0 and its mean squared error is cca. $1/N$ where N is the bin number. Thus, the squared average of $r_{av}(0)/(1/\sqrt{N})$ is a χ^2 -type quantity. This is our “ χ^2 ” in this paper.

While the expectation value of χ^2 is by construction of the order of unity, it is better to determine it and the mean deviation from test runs. The results can be found in Fáy & Lukács (2006), and we use the significance limits determined there.

The method is sensitive only to repetitions at constant lengths. If there is repetition but at growing or shrinking intervals, the many averagings obscure the signal. Thus we can detect non-randomness only if we are at (or near to) the “true” cosmology.

References

- Allen, S. W., Schmidt, R. W., & Fabian, A. C. 2002, MNRAS, 334, L11
 Bellanger, C., & de Lapparent, V. 1995, ApJ, 455, 103
 Bershad, M. A., Koo, D. C., Kron, R. G., et al. 1997, ApJS, 109, 45
 Broadhurst, T. J., Ellis, R. S., Koo, D. C., & Szalay, A. S. 1990, The Large-Scale Distribution of Galaxies and the Galactic Walls, Nature, 343, 726
 Broadhurst, T. J., Ellis, R. S., Ellman, N., et al. 1993, ASP Conf. Ser., 51, 112
 Einasto, J., Einasto, M., Gottlöber, S., et al. 1997a, Nature, 385, 139
 Einasto, J., Einasto, M., Frisch, P., et al. 1997b, MNRAS, 289, 801
 Fáy, N., & Lukács, B. 2006, On a Possible Indicator of Homogeneity of the Universe [arXiv:gr-qc/0605081]
 Jánossy, L. 1965, Theory and Practice of the Evaluation of Measurements (Oxford: Clarendon Press)
 Kramer, D., Stephani, H., Hertl, E., MacCallum, M., & Schmutzer, H. (ed.) 1980, Exact Solutions of Einstein’s Field Equations (Cambridge: Cambridge University Press)
 Lukács, B., & Perjés, Z. 1977, Time-Dependent Maxwell Fields in Stationary Geometry, Proc. 1st Marcel Grossmann Meeting, Trieste (North-Holland), 281
 Lukács, B., & Mészáros, A. 1984, Ap&SS, 114, 211
 Saar, E., Einasto, J., Toomet, O., et al. 2002, A&A, 393, 1
 Szalay, A. S. 1999, Phil. Trans. Roy. Soc., 357, 117