Mem. S.A.It. Vol. 85, 102

Memorie della

A statistical analysis of varying α data

L. Kraiselburd[1], S. Landau[2], and C. Simeone[2,3]

[1] Grupo de Astrofísica, Relatividad y Cosmología, Facultad de Ciencias Astronómicas y Geofísicas, Universidad de La Plata, Paseo del Bosque S/N 1900 La Plata, Argentina e-mail: lkrai@fcaglp.unlp.edu.ar
[2] Instituto de Física de Buenos Aires, CONICET - Universidad de Buenos Aires, Ciudad Universitaria - Pab. 1, 1428 Buenos Aires, Argentina, e-mail: slandau@df.uba.ar
[3] Departamento de Física, FCEN, UBA, Ciudad Universitaria - Pab. 1, 1428 Buenos Aires, Argentina, e-mail: csimeone@df.uba.ar

Abstract. We collect different groups of data of the variation of the fine structure constant to compare and verify the consistency between them using the Student test and Confidence Intervals. We separate data sets in smaller intervals based on a proposed criterion. Another statistical analysis is proposed that considers phenomenological models for the variation in α. The results show consistency for a certain amount of reduced intervals, in contrast to those obtained considering the mean values from the entire interval.

Key words. quasars: absorption lines – Cosmology: miscellaneous

1. Introduction

Since the large-number hypothesis (LNH) was proposed (Dirac 1937), the variation of fundamental constants has been an important subject of research. The leading astronomical method is based on the analysis of high-redshift quasar absorption systems. Most of the reported data are consistent with null variation of fundamental constants. The many multiplet method (MMM) (Dzuba et al. 1999; Webb et al. 1999), gains an order of magnitude in sensibility with respect to previously reported data. Only two research groups (Webb et al. and Srianand et al.), which have a considerable amount of data, have used this method in 2 different telescopes (Keck/HIRES and VLT/UVES) obtaining very different results:(Webb et al. 1999; Murphy et al. 2003) suggests $\alpha_0 < \alpha_{today}$, while (Srianand et al. 2004; Chand et al. 2004; Srianand et al. 2007) arrived to $\alpha_0 = \alpha_{today}$, and (Webb et al. 2011; King et al. 2012) suggets $\alpha_0 > \alpha_{today}$). In a previous paper (Landau and Simeone 2008) we have pointed out that results calculated from the mean value over a wide redshift range (or cosmological time scale) are at variance with those obtained considering shorter intervals. Here, we show a condensed version of our article (Kraiselburd et al. 2013) where we have reanalyzed the available data obtained with the many multiplet method with both telescopes, using the statistical tools and method introduced in (Landau and Simeone 2008). On top of that, we propose another statistical method for studying the discrepancy between Keck and VLT data considering three phenomenological models for the α variation: i) null variation, ii) time variation of α equal to the mean

Send offprint requests to: L. Kraiselburd value of each data set, and iii) spatial variation of α following the dipole model proposed by (King et al. 2012). In each case we compute the amount of data of each group that lie within the Gaussian distribution corresponding to each model.

2. Statistical tools

The corresponding procedure for testing the consistency of two independent experiments is a test for the difference between two population means, which involves a statistic defined in terms of two sample means and two sample variances (Student test) (for details see Landau and Simeone 2008; Kraiselburd et al. 2013). In practice, we used an algorithm that yields a level λ^* such that the obtained value of the statistic lies within the associated rejection region. Hence, at level λ the null hypothesis should be rejected when $\lambda^* \leq \lambda$. But, when one of the experiments includes for a given interval very few data (so that one cannot reasonably define a sample mean and a sample variance), the procedure to be followed involves a confidence interval. This latter one is constructed from the sample values of the experiment with a number of data that do make a statistical treatment possible. In such way, to test the consistency of a given observation 2 against observation 1, we constructed an interval I of confidence $100\,P\%$ from the values of group 1. Then, if the null hypothesis is true, $P = 1 - \lambda$ is the probability that the result of an observation of group 2 lies within this confidence interval, and the null hypothesis should be rejected at level λ when this is not the case.

In order to avoid any biasing associated with the arbitrary choice of location and size of the intervals, we have followed the procedure of (Landau and Simeone 2008): The first interval (to be considered to apply the test) starts at a redshift/angle a with fixed b width. The following i intervals start at redshift/angle $a + i * c$ with the same width; $c = 0.1$ for the selection according to redshift, and $c = 0.025$ when selecting according to angular position. After testing all these intervals, we changed the value of b and performed the same analysis again. In all cases we considered a minimum number of data in each interval as a necessary condition for applying the test.

A type II error is the failure to reject a false null hypothesis and β is the associated probability. While the probability λ of type I error can be fixed independent of the population or sample values, the calculation of β requires the choice of a definite alternative hypothesis; that is, the inequality $\mu_1 \neq \mu_2$ must be specialized as a definite equality $\mu_1 - \mu_2 = \delta$. β is built from the normal cumulative distribution functions Φ, which depend on: z_λ (obtained by inverting a normal N(0,1) distribution);δ; S_1^2, S_2^2; n_1, y n_2. The probability β measures, for a given alternative hypothesis and certain sample sizes, whether the data have led to a too conservative result or not. More precisely, a high value of β implies that the data variance is larger than the difference between the null hypothesis and a given alternative hypothesis, which makes it difficult to resolve them. Small sample sizes will often lead to such high values of β.

2.1. Test of the null, mean value and dipole models

We also propose another way to analyze data on varying α, which is similar to some of the analyses performed by (King et al. 2012). However, in this case, we add the data reported by Srianand (2013, private comm.) (which come from the reanalysis (Srianand et al. 2007) of 21 observations made by (Chand et al. 2004)) to the discussion and include the null hypothesis as a possible model. We assumed three phenomenological models for the α variation: i) null variation, ii) the value of α in the past was different from its actual value and is a fixed number, and iii) the variation in α follows the dipole model proposed by (King et al. 2012). We proceeded as follows: for each phenomenological model we constructed the normal distribution asociated with its mean value and standard deviation. Then, we calculated the amount of data from each group that lies within the 3 and 6 standard deviations of the proposed normal distributions. We tested each data group separately. For the null distribution we took the standard deviation associated with the mean value reported by (Srianand et al. 2007). For ii), we considered the mean value and standard deviation reported by each group. For the dipole model we have one distribution for each value of θ; and the value of the dipole coordinates is equal to the one obtained by (King et al. 2012) considering each group of data separately.

3. Results

We analyze results of applying the Student test and/or confidence interval method to recent astronomical data. We consider the following three groups of data to perform our statistical analysis:

Group I: Data obtained with the Keck by (Murphy et al. 2003) and (King et al. 2012) (141 data points).

Group II: Data obtained with the VLT by (King et al. 2012) (153 data points).

Group III: Data obtained with the VLT by Srianand (2013, private comm.), details in (Kraiselburd et al. 2013).

3.1. Redshift

To apply the Student test grouping the data according to redshift, we considered datasets where the lowest total number of data is equal to 12 ($n \geq 12$). The total redshift interval to be tested is $(0.440, 2.795)$; we applied the test to shorter intervals of equal width ($\Delta z = 0.35$). Fig. 1 shows the value of $\lambda*$ obtained from the comparison of data from group I with data from group II. Not all values of λ^* corresponding to the redshift interval $(1.470, 2.795)$ are higher than 0.025 and therefore this interval was discarded from the consistency interval. We also performed the same test and changed the value of the interval width Δz and obtained similar results. These results give evidence that the variation in α may be relevant at higher redshifts, as was pointed out by (Murphy et al. 2003). The value of β was obtained using the normal distribution (one-tail test, see section 2); we considered the alternative hypothesis equal to the mean value of $\frac{\Delta\alpha}{\alpha}$ obtained by each group: $\mu_1 = -0.6 \times 10^{-5}$ (Keck) and

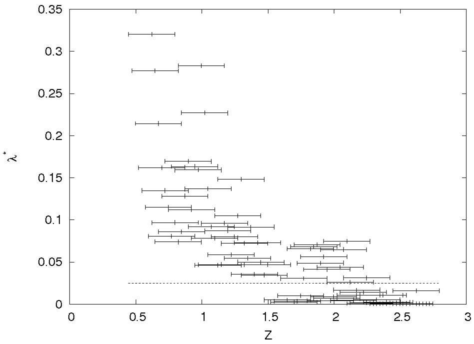

Fig. 1. Results of Student test comparing data from group I with data from group II. Data sets are selected according to redshift; $\lambda*$ is the calculated level of the test for each redshift interval (the dotted line indicates the $\lambda* \leq 0.025$ rejection region).

$\mu_2 = 0.2 \times 10^{-5}$ (VLT). In all cases the value of β was higher than 0.1, showing the difficulty to distinguish the null hypothesis from the alternative hypothesis with the present data sets. We also performed the same analysis, but constrained the number of data to $n \geq 15$ and $n \geq 18$ to obtain get lower values of β; and changing the width interval as described above. However, the analyses showed that the value of β did not change significantly, while increasing the lowest number of data in each data set involves ruling out many intervals from the possible intervals to be tested.

In cases where we had to build a confidence interval for a group of data and compared the results with each single reported value of another author we have chosen $\lambda = 0.025$. Again, we decided on the bin size as the shortest interval centered at the reported value, which contains n data. The criterion for analyzing the results was the same as in (Landau and Simeone 2008)

We calculated confidence intervals for the data from group II for redshift intervals centered on each value of data from group III with the same telescope and containing at least 12 data points. The confidence intervals calculated for $z = 1.348, 2.022$ do not overlap with the corresponding reported intervals. Therefore, the redshift intervals $(1.278, 1.419)$ and $(1.935, 2.110)$ should be discarded from the consistency interval because the data from group II used to calculate the confidence intervals belong to this interval. Therefore, we conclude that 19 data points from group III are consistent with 113 data points from group II over the redshift intervals $(0.142, 1.278)$, $(1.419, 1.935)$ and $(2.100, 2.429)$. In all cases β is higher than 0.1, which indicates how difficulty it is to distinguish the null hypothesis from the alternative hypothesis with the present data sets. We also calculated the confidence intervals for $n \geq 15$ and $n \geq 18$ in an attempt to improve the value of β. However, the values of β did not change significantly, while including more data in each confidence interval using the present data set leads to an enlarged redshift interval for which the confidence interval is calculated.

We also calculated confidence intervals to compare the data points obtained with the Keck Telescope (group I) with those obtained with the VLT (group III). The redshift interval $(1.233, 1.838)$ should be discarded because the confidence intervals calculated for $z = 1.348, 1.555, 1.657$ do not overlap with the corresponding reported intervals. Accordingly, the data points $z = 1.439, 1.637$ from group III should be discarded as well. Consequently, the redshift intervals $(0.179, 1.233)$ and $(1.838, 2.457)$ are consistent, where there are 6 and 7 data points from group III, and 54 and 28 from group I respectively. A similar analysis was performed previously (Landau and Simeone 2008), however, in the present work, we considered the enlarged errors reported by Srianand (2013, private comm.). The β values are again high, noting that more data points are needed to reduce the type II error probability.

3.2. Spatial variation

The selection according angular posistion was performed as follows: The data set were selected by their value of $\cos\theta = \boldsymbol{X} \cdot \boldsymbol{D}$, where $\boldsymbol{X}$ is the quasar position and $\boldsymbol{D}$ is the dipole direction obtained by (King et al. 2012) considering the Keck and VLT datasets. The angular distribution of available data is as follows: Keck data (group I) come from the region where $-1 \leq \cos\theta \leq 0.5$, while the VLT (group II and group III) reports data from the region where $-0.5 \leq \cos\theta \leq 0.9$. Accordingly, the Student test can be applied to reduced data sets from both telescopes. Furthermore, we have to reduce the lowest total value of each data set (to apply the test) to $n \geq 6$. Figure 2 shows the results of the Student test performed for reduced data set comparing group I with group II within the interval $-0.225 < \cos\theta < 0.225$. The width of the intervals for which the Student test was applied is $\cos\theta = 0.075$.

It follows from Fig 2 that not all values of λ^* within the interval $(0.075, 0.150)$ are higher than 0.025. Since data belonging to this interval were used to perform the Student test of other intervals, we analyzed these intervals again. For the interval $(0.05, 0.075)$ there is only one data set reported by Keck and four data sets from the VLT. Therefore we calculated a confidence interval with the VLT data and compared the result with the reported data from Keck, finding consistency within this interval. For the interval $(0.150, 0.225)$ the Student test can be applied because there are seven data sets from Keck and 12 data sets from VLT. We obtained $\lambda^* = 0.17$ for this interval. To complete our analysis we performed the Student test by changing the value of the width interval for which the test is applied arriving to similar results ($\cos\theta = 0.070$ and $\cos\theta = 0.080$).

Almost all values of β are higher than 0.1, showing the need for more data to reduce this value. We also calculated the values of β for other width intervals and found no significant difference with results shown.

For the interval where $\cos\theta > 0.225$, we calculated confidence intervals for the data from group II for intervals centered on each value of data from group I; and confidence intervals haven been calculated for the data from group I for intervals centered on each value of data from group II for the interval where $\cos\theta < -0.225$.

In summary, we analyzed the consistency over the interval $(-0.454, 0.497)$, comparing 87 data sets from group I with 121 data sets from group II. From the Student test and confidence interval calculation it follows that 68

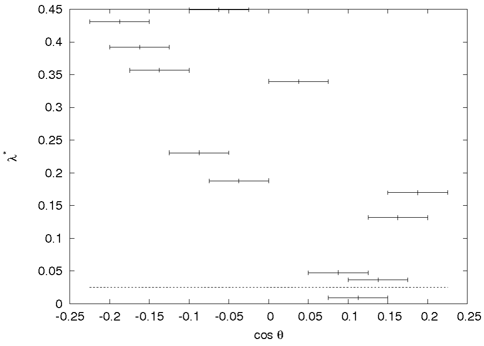

Fig. 2. Results of Student test comparing results obtained with by group I with results obtained by group II. Data sets are selected according to angular position, $\lambda*$ is the calculated level of the test for each interval (the dotted line indicates the $\lambda* \leq 0.025$ rejection region).

data sets from group I (79% of the analyzed data) are consistent with 112 data sets from group II (93% of the analyzed data) over the intervals $(-0.454, -0.429)$, $(-0.327, 0.05)$, and $(0.150, 0.497)$.

We have also calculated confidence intervals for the data from group II for $\cos\theta$ intervals (0.1 width) centered on each value of the data from group III with the same telescope and containing at least seven data points. The quantity of testable intervals is reduced to ten, because this is the number of quasars from which the Srianand (2013, private comm.) and (Srianand et al. 2007) data came from. The results obtained with this method differed enough with those obtained with the Sudent test for the same analyzed intervals. Therefore, we conclude that in this case the comparison between a single quasar data set and a confidence interval is not the appropiate tool for testing the consistency between groups of data. Thus, we performed a Student test using the same procedure as we applied for the comparison between Keck and VLT data reported by the group of Murphy et al. reducing the lowest total value of each data set to $n \geq 3$, and $\Delta\cos\theta = 0.075$. Even though it is not ideal to perform a Student test with such a reduced data set, we considered that it is a better tool for the analysis than the confidence interval comparison presented above. Results show that all intervals derived from the Srianand (2013, private comm.) data are consistent with those built from the (King et al. 2012) data, being the tested interval $(0.000, 0.625)$. Again, all β values obteined in this case are again very high.

To compare data from group I and group III we, again, calculated confidence intervals for the data from group I for $\cos\theta$ intervals (0.1 width) centered on each value of the data from group III. Only ten data points from group III ($\sim$ 45% of the data) can be used, and these in turn are able to form four intervals because some of them arise from the same quasar (29 data points from Keck can be compared (21% of the data)). The results show that the analyzed interval $(4.70\times10^{-4}, 0.380)$ is consistent. β values are very high in all cases, suggesting again that a greater quantity of data is required to improve these values.

3.3. Phenomenological models

In section 2.1 we have described another method for analyzing the different groups of data on varying α. The results are listed in Table 1. Although it can be noted that the three groups suit the dipolar model better than the other two models, there is still a large amount of data from group I and group II that is left out of both distributions. Furthermore, it should be noted data from the VLT (group II and group III) favor the dipole model over the other proposed phenomenlogical models, while data from the Keck telescope cannot distiguish between the dipole model and the null distribution.

4. Summary and conclusions

Using statistical methods explored in a previous paper (Landau and Simeone 2008) we have tested the consistency between different recent astronomical data that indicate a possible variation in α; adding this time, more data and grouped them according to redshift and angular position. We also proposed some

Table 1. Percentage of data within the SD of the mean. The dipole hypothesis implies a spatial variation in α following the dipole model proposed by (King et al. 2012), the coordinates of the dipole are those obtained by (King et al. 2012) considering each group of data separately; the null hypothesis consists of a null mean and a standard deviation given by (Srianand et al. 2007) and Srianand (2013, private comm.); and the mean value hypothesis contains the mean of the data group and the corresponding standard deviation.

Data Group	3σ			6σ		
	Dipole hyp	Null hyp	Mean value hyp	Dipole hyp	Null hyp	Mean value hyp
Group I	21%	21%	15%	45%	38%	30%
Group II	46%	7%	16%	71%	37%	30%
Group III	95%	33%	38%	95%	81%	81%

phenomenological models for the variation in α and computed the amount of data that lie within 3 and $6-\sigma$ of the asociated Gaussian distribution.

Although results of the statistical analyses by grouping in z show that the variation in α is more relevant at higher redshift, from the analysis by grouping the data in angular position, they show consistency over most of the analyzed intervals. In all cases, the high the probabilty of the type II error (β) indicates the imperious need for more data to reach strong conclusions. Finally, the analysis of Gaussian distributions of the proposed phenomenological models suggests that although one cannot rule out a possible variation in α, this may be due not only to the angular position but also to redshift. More data are required on this aspect as well.

Acknowledgements. Support for this work was provided by PIP 0152/10 CONICET. The authors are grateful to Raghunathan Srianand for providing enlarged errors of the VLT data.

References

Berengut, J. C., et al. 2012, A&A, 542, A118

Brandt, S. 1989, Statistical and computational methods in data analysis, (Elsevier, Amsterdam)

Brownlee, K. 1960, Statistical theory and methodology in science and engineering, (Wiley, New York)

Chand, H., et al. 2004, A&A, 417, 853

Devore, J. 1995, Probability and Statistics for Engineering and Sciences, (Duxbury Press, Belmont)

Dirac, P. A. M., 1937, Nature, 139, 323

Dzuba, V. A., et al. 1999, Phys. Rev. Lett., 82, 888–891

Kendall, M., et al. 1994, Kendall's Advanced Theory of Statistics, (Wiley, New York)

King, J. A., et al. 2012, MNRAS, 422, 3370-3414

Kraiselburd, L., Landau S., & Simeone C., 2013, A&A, 557, A36

Landau, S. J. & Simeone, C., 2008, A&A, 487, 857-863

Murphy, M. T., et al. 2003, MNRAS, 345, 609–638

Murphy, M. T., et al. 2007, Phys. Rev. Lett., 99, 23, 239001

Srianand, R., et al. 2004, Phys. Rev. Lett., 92, 121302

Srianand, R., et al. 2007, Phys. Rev. Lett., 99, 23, 239002

Webb, J. K., et al. 1999, Phys. Rev. Lett., 82, 884–887

Webb, J. K., et al. 2003, Ap&SS, 283, 577

Webb, J. K., et al. 2011, Phys. Rev. Lett., 107, 19, 191101