PROBABILISTIC REPRESENTATION OF THE GAIA EXTRAGALACTIC REFERENCE FRAME

A.H. Andrei[*1,2,3] **, A. Fienga**[1]**, M. Assafin**[3]**, J.L. Penna**[2]**, D.N. da Silva Neto**[3]**,**
R. Vieira Martins[1,2,3]**, P. Soares Musumeci**[4,2]

[1]Observatoire de Paris, IMCCE / CNRS UMR 8028, 77 Av. Denfert Rochereau, 75014 Paris, France
[2]Observatório Nacional-ON/MCT, Brasil
[3]Observatório do Valongo-OV/UFRJ, Brasil
[4]Departamento de Física-UFF, Brasil

ABSTRACT

The Gaia extragalactic reference frame (GCRF) will be formed by about 500 000 quasars, to a positional precision reaching 50 μas. The GCRF is a key factor for several of the Gaia science goals, and it is important to access its sky and magnitude distribution before the mission launch. Despite recent in-depth investigations, such as the 2dF and SLOAN, there is no all-sky account to such. Hitherto, to the M $=$ 20 limit, the USNO B1.0 and GSC2 catalogues attain completeness. They can thus be used to generate an accurate probabilistic description of the GCRF. Starting from those catalogues, the main strategies that are brought into play for the task involve determination of the quasars loci in the proper motions space; in the colour/colour space; in the broad-band spectroscopy space; and in the probability space as given by the stars and extragalactic objects distribution models. Here we discuss such strategies and some of the obtained results.

Key words: Gaia; Reference frame; Quasar.

1. INTRODUCTION

The Gaia extragalactic celestial reference frame (GCRF) will be formed by about 500 000 quasars brighter than G $=$ 20, to a precision better than 100 μas. It will be defined at $|b|>20^\circ$ avoiding the galactic plane, and will further contain a clean sample, five to ten times smaller, free from stellar contaminants. The GCRF will be pivotal for many of the mission's science goals and represents an important improvement over the present representation of the ICRS. Although it is not possibile to obtain it before the start of mission, it is particularly interesting to obtain a realistic description of what the GCRF, and its clean sample, will look like in terms of the spatial, magnitude and colour distributions.

Assuming the current ideas on the large-scale distribution of matter, the GCRF mean spatial density should be of 20 QSOs deg^{-2}. Setting aside the Gaia stars, albeit extremely precise proper motions will be known for them, the GCRF QSOs density is still low for many of the modern very reduced field observations. Also, according to widely accepted determinations of the luminosity function, most (perhaps 87%) of constituent QSOs will have Z $<$ 2.2. QSOs clustering at low Z, by its turn, seems to be an observational evidence. Therefore, there are added reasons to try to mimic the GCRF from present observations, and to access its small-scale regime of homogeneity. Figure 1 exemplifies the small-scale inhomogeneities, from the densest region of the 2QZ-2dF (Croom et al. 2004) survey.

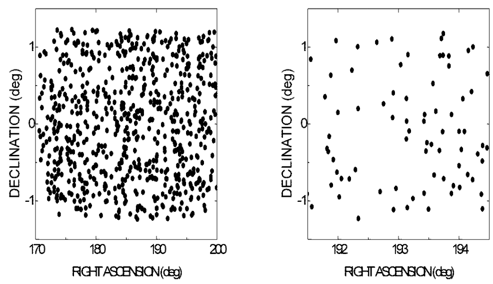

Figure 1. Plot (left) and zoom (right) from the densest region of the 2QZ–2dF quasar survey.

There is no complete all-sky quasar survey. However, there are all-sky plate surveys complete to V $=$ 20. which are the basis for the USNO B1.0 catalogue (Monet et al. 2003) and the not yet released GSC2.3 (STScI and OAT). In other words, the quasars forming the GCRF are contained in those catalogues. To single them out from the much larger stellar population, or more rigorously to maximize the combination of efficiency and completeness of the harvest is what is searched for in this project. A number of independent strategies are then combined to mimic the GCRF. The strategies are listed and some results are commented in the next Section. Conclusions and perspectives are summarized in the final Section.

[*]On leave from home institute

2. STRATEGIES

2.1. The Known Quasars

Besides the already mentioned 2QZ-2dF survey, several surveys contain statistically significant collections of quasars. Certainly among the most numerous are: the SLOAN DSS (Schneider et al. 2003), 16 713 QSOs selected by multi-colour technique; the VLBI standard calibrators, from the VLBA, VLA, and the ICRF, presently amounting to 3135 QSOs. Quasars can also be found in large numbers, and with reasonably sensitive search mechanisms from projects like the 2MASS (Cutri et al. 2003), 5 000 000 point sources; and FIRST (Becker et al. 2001), where there maybe are 250 000 radio QSOS, from which 35 000 POSS identifications are already obtained. Almost all of the results from those investigations are present in the $11^{\rm th}$ edition of the catalogue of quasars and active nuclei (Véron-Cetty & Véron 2003). The quasars from these different projects enable us to train the strategies to single out quasars from the astrometric catalogues. They also provide a clean sample for the mimicked GCRF. Finally, the V&V list enables us to define an optical reference frame directly tied to the HCRS and the ICRS, as discussed elsewhere in these proceedings (Andrei et al. 2005).

2.2. Proper Motions

Since the astrometric precision of the dense catalogues is not better than 100 mas, detected proper motions provide completeness but low efficiency for quasar filtering. Only 9.4% of the B1.0 entries for known QSOs show non-zero proper motion, and usually with relatively high error. In a pilot investigation on 38 fields already studied (Fienga & Andrei 2004), the significant proper motion criterion removed 25% of the field stars, to a spurious loss of 3.9% QSOs. Trial experiments including GSC2.2 proper motions remove an additional 20% of stellar sources. A further 20% were removed by applying proper motions reconstructed from the B1.0 and GSC2.2 positions. With the improved astrometry expected from the GSC2.3, the proper motion criteria is expected to cleanse away 50% of the stellar objects, at a potential loss of about 5% of true QSOs.

2.3. Morphology

Combining the B1.0 and GSC2 performance to separate compact and extended objects an efficiency of 97.95% can be expected, to the corresponding loss of 2% of true QSOs. The shadow of a host galaxy around the QSO would typically be 2 magnitudes dimmer, and would not extend beyond $1''$, while the pixel size used for the B1.0 is $0''.9$. Therefore, host galaxies do not pose a statistically relevant trouble for the GCRF simulation.

However, a pilot investigation is on course to assess their relevance for the Gaia actual observations. For this, let us assume that, in general, the host galaxy absolute magnitude should be brighter than -23.5; that the host galaxy luminosity increases proportionally to the strength of the central source; that also their size tends to follow the same rule; and that the QSO luminosity function indicates that up to $G = 20$ the QSO number density should peak for $z = 0.6$. Then both a statistically significant number of resolved host galaxies around QSOs, and a contaminant population of AGNs, must be expected among the Gaia detected quasars.

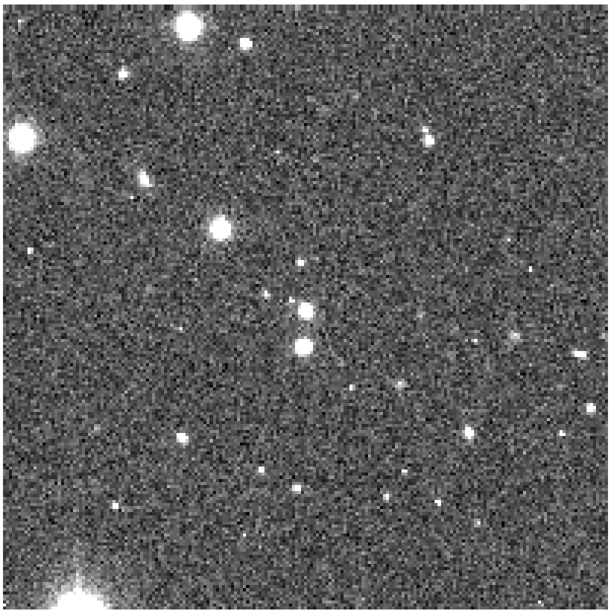

Figure 2. CCD image of the optical counterpart of the RORF source 1726+455h (in the center of the frame). It is a defining ICRF source, and yet it is a Seyfert galaxy and its PSF's SHARP exceeds the stellar average in the field.

This hypothesis was probed using the USNO data bank of optical images of RORF sources (Johnston et al. 1995). These are 532 CCD images, in 200×200 pixel frames, at scale $0''.7$ per pixel or $0''.4$ per pixel. An IRAF pipeline was established to investigate the quasar PSF relative to selected field stars. For safety, only frames in which there were at least 5 well defined field stars, and in that the quasar instrumental magnitude was brighter than 18 were retained. Using SHARP as the PSF measuring parameter, it was found that 17 RORF sources (out of 199) are not stellar like, at 2.5 σ. Out of these, 13 are found in the literature as AGNs other than QSO, Figure 2. To check the reliability of the method, the 30 QSOs found as most stellar-like were also searched for in the literature. As a result, just 3 of them are referred as AGNs other than QSO.

2.4. Colour/Colour Loci

Since the initial optical identifications of radio quasars, the ultraviolet excess was remarked. This prompted the classical method to separate quasars from stellar population using colour-colour diagrams. Though this method loses effectiveness for redshifts greater than 2.2, it was already seen that the vast majority of the GCRF quasars would lie closer than that limit. As the B1.0 and GSC2 catalogues bring magnitudes in the filters J, F, and N (the first two in slightly different passbands on the two epochs), the method can be used. Additionally, the 2MASS (Cutri et al. 2003) point source catalogue brings 3 near-infrared colours, for about half of the sources.

Table 1 lists the QSOs space in the B1.0 colour loci. It contains three different samples, radio selected standard quasars, near infrared bright quasars from 2MASS, and the all-sky varied quasar population from V&V. The QSO loci in the B1.0 colour space is similar for the three samples, suggesting the adequacy of the method. The corresponding histograms for all the distributions is associated to gaussian distributions to more than 98% for the B–R and R–I colours. Nevertheless, the standard deviations are seen to be large, pointing to the need of the other strategies of quasar searching, and to the need of an independent assessment of the quasar colours variation.

Table 1. QSOs loci in the B1.0 colour space, averages and standard deviations. Three samples of differently selected quasars are compared. Ordinals '1' and '2' refer to the first and second epoch catalogue plates; 'm' refers to the mean magnitude.

Sample	V&V	2MASS	VLBI
N	42279	4966	238
B1-R1	0.2 (σ 0.7)	0.3 (σ 0.7)	0.4 (σ 0.7)
B2-R2	0.3 (σ 0.8)	0.4 (σ 0.9)	0.6 (σ 0.8)
Bm-Rm	0.3 (σ 0.5)	0.3 (σ 0.6)	0.5 (σ 0.6)
R1-I	0.5 (σ 0.9)	0.3 (σ 0.9)	0.3 (σ 1.0)
R2-I	0.6 (σ 0.8)	0.4 (σ 0.8)	0.3 (σ 0.8)
Rm-I	0.5 (σ 0.6)	0.3 (σ 0.6)	0.3 (σ 0.8)

The catalogues colour/colour criteria were tested in four 2dF regions, where the QSOs count is presumably complete. The regions are 1 deg^2, around δ=0°, centred at α 11^h, 12^h, 13^h, 14^h. There are 14 804 B1.0 objects with at least 2 colours. Applying the B1.0 colour/colour criteria only, 55.2% of objects are removed and 93.4% of the QSOs are kept. From the B1.0 population, there are 3998 GSC2.2 objects with 2 magnitudes (J and F). On these, by applying the GSC2.2 colour criterion 71.7% of objects are recognized as stars. Again in the B1.0 population, there are 4999 2MASS objects presenting the J, K, and F magnitudes. By applying the colour/colour criteria on them, 96.9% are recognized as stars. Finally, by combining all the criteria, on the B1.0 population, 66.6% of objects are removed and 89.5% of the QSOs are retained.

The variation of the QSOs locus in the colour space can be studied with the quasar synthetic spectra library, (QSSL, Smette et al. 2004), developed at the University of Liege. Among other features, it contains spectra at narrow wavelength intervals, for different redshifts, slopes, and integrated flux. From the spectra, the catalogue magnitudes can be synthesized, by taking into account the plates exposure times, sensitivity, non-linearity of the emulsion response, and photometric calibration. Figure 3 shows the spectral variation of the B1.0 B–R colour.

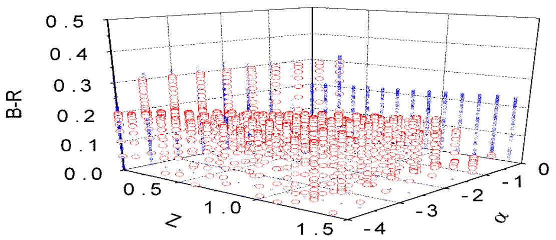

Figure 3. Variation of the QSOs B1.0 B-R colour as a function of the redshift (z) and slope (α), as synthesized from the QSSL.

2.5. Stellar Model

A number of stellar models are nowadays capable of providing accurate predictions of stellar counts (Castellani et al. 2001), and to generate representative star catalogues, at scales as detailed as 1deg^2, especially off the galactic plane. The numerical counts of the models put an important constraint on the available candidates from the catalogues. In this way the probability of finding a given density of QSOs in a given 1deg^2 region of the sky becomes significantly better defined. Stellar models also offer an independent assessment of the colours from the catalogues, giving an opportunity to correct systematic biases from the plates photometry.

Both aspects were verified in a 2QZ 1deg^2 region, centred at $\alpha = 11^h\ 9^m\ 4^s.9$ and $\delta = -0°\ 6'\ 19''$. It happens to be an under average density region, counting 6 QSOs. The B1.0 catalogue counts 2368 objects, of which 87 are quasar candidates by the colour/colour criteria. The stellar population was synthesized by the Besançon model (Robin & Creze 1986), to match the B1.0 number of objects. It resulted then that 102 stellar objects would fall into the QSO B1.0 colour/colour space. Therefore the region can correctly be identified as suspected of containing few quasars, and the colour/colour criteria can be adjusted to match the actual distribution. This process can be iteratively done, in steps of varying sizes, over all the GCRF domain. Table 2 lists the B1.0 and Besançon model colour/colour centre of loci in the studied region.

2.6. Quasar Space Distribution Models

Along about the same lines as the previous strategy, models and hypothesis would enable a probability distribution of QSOs to be drawn, for structures in scales of the order of 10 deg^2: QSOs and galaxies clustering, QSOs clustering on redshift; Universe models. The work on this item is still at the planning phase.

Table 2. Averages and standard deviations for the BRI colour loci. The upper line lists the values from the population synthesized by the Besançon model. The next two lines bring the equivalent values obtained from the B1.0 magnitudes. These three lines refer to a 1 deg^2 region of the 2QZ survey. The last two lines list the B1.0 derived values for V&V quasars in four nearby regions (see text) representative of the 2QZ survey.

Sample	B-R	R-I
Besançon model	$+2.1$ (σ 1.6)	-0.5 (σ 1.4)
B1.0 - 1step	$+1.4$ (σ 0.8)	$+0.3$ (σ 0.5)
B1.0 - 2ndep	$+0.5$ (σ 0.6)	$+0.5$ (σ 0.5)
2QZ(B1.0-1step)	$+0.1$ (σ 0.5)	$+0.8$ (σ 0.4)
2QZ(B1.0-2ndep)	$+0.2$ (σ 0.5)	$+0.9$ (σ 0.6)

3. CONCLUSION AND PERSPECTIVES

This is the report of a work in progress. Some parts are all but completed, as the precise astrometry of the Véron-Cetty & Véron sources. Some other are well under way, as the B1.0 colour/colour analysis, and the host galaxies discussion. And, yet, the last of the presented strategies has barely been started.

The final goal is to produce and analyze a probabilistic representation of the GCRF, as well as a clean sample of truly identified quasars.

The clean sample will be obtained from the existing QSOs surveys in various wavelengths. It will contain more than 50 000 quasars and will furnish the paradigms for training the criteria of colour/colour diagrams, stellar and QSOs space distribution models.

Taking separately, the proper motion criteria is believed to remove 50% of the stellar objects, to the potential loss of 5% of true QSOs. The morphology criteria would remove 97.75% of the galaxies, to a corresponding loss of 2% of true QSOs. Host galaxies will not be a hindrance, but may be influential for the actual Gaia observations. The colour/colour criteria seems to able to remove 66% of the stars, while correctly finding 89.5% of the QSOs. The distribution and variation of the different QSO magnitudes can also be used to analyze the Gaia quasar population redshift and slope of the spectrum.

To narrow down the probability function, the strategy is to combine the set of independent criteria to single out the QSO population contained in the dense, complete to V = 20 catalogues B1.0 and GSC2.3.

ACKNOWLEDGMENTS

AHA thanks CNRS for contract QAF183803, MA acknowledges FAPERJ grant E-26/170.686/2004, DNSN is grateful to CNPq grant 303950/2003-0, RVM is thankful to CAPES grant 0449/04-0, PSM thanks a PIBIC/CNPq grant.

REFERENCES

Andrei, A.H., Fienga, A., Assafin, M., et al., 2005, ESA SP-576, this volume.

Becker, R.H., White, R.L., Gregg, M.D., et al., 2001, ApJSupp., 135, 227 - The FIRST Bright Quasar Survey

Castellani V., Degl'Innocenti S., Petroni S., Piotto G., 2001, MNRAS, 324, 167

Croom, S.M., Smith, R.J., Boyle, B.J., et al., 2004, MNRAS, 349, 1397 - The 2dF QSO Redshift Survey

Cutri, R.M., Skrutskie, M.F., Van Dik, S., et al., 2003, IPAC/CalTec - The 2MASS Point Source Catalogue

Fienga, A., Andrei, A.H., 2004, A&A, 420, 1163

Johnston, K.J., Fey, A.L., Zacharias, N. et al;, 1995, AJ, 110, 880 - A Radio Reference Frame (RORF)

Monet, D.G., Levine, S.E., Casian, B. et al, 2003, AJ, 125, 984 - The USNO B1.0 Catalogue

Robin, A., Creze, M, 1986, A&A, 157, 1, 71 - Stellar Populations in the Milky Way - A Synthetic Model (Besançon)

Schneider, D.P., Xiaohui, F., Hall, P.B. et al, 2003, AJ, 126, 2593 - The SLOAN Digital Sky Survey Quasar Catalogue

Smette, A., Claeskens J.F., Surdej, J., Royer, P., Vandenbulcke, L., 2004, personal communication - The Liege QSOs Synthetic Spectra Library

Véron-Cetty, M.P., Véron, P., 2003, A&A, 412, 399 - 11th edition