

# Homogenization in compiling ICRF combined catalogs

F. J. Marco<sup>1</sup>, M. J. Martínez<sup>2</sup>, and J.A. López<sup>1</sup>

<sup>1</sup> Dep. Matemáticas, IMAC, Universitat Jaume I, 12071 Castellón, Spain  
e-mail: [marco;lopez]@mat.uji.es

<sup>2</sup> Dep. Matemática Aplicada, IUMPA, Universidad Politécnica de Valencia, 46022 Valencia, Spain  
e-mail: mjmartin@mat.upv.es

Received 30 May 2013 / Accepted 6 August 2013

## ABSTRACT

*Context.* The International Astronomical Union (IAU) recommendations regarding the International Celestial Reference Frame (ICRF) realizations require the construction of radio sources catalogs obtained using very-long-baseline interferometry (VLBI) methods. The improvement of these catalogs is a necessary procedure for the further densification of the ICRF over the celestial sphere.

*Aims.* The different positions obtained from several catalogs using common sources to the ICRF make it necessary to critically revise the different methods employed in improving the ICRF from several radio sources catalogs. In this sense, a revision of the analytical and the statistical methods is necessary in line with their advantages and disadvantages. We have a double goal: first, we propose an adequate treatment of the residual of several catalogs to obtain a homogeneous catalog; second, we attempt to discern whether a combined catalog is homogeneous.

*Methods.* We define homogeneity as applied to our problem in a dual sense: the first deals with the spatial distribution of the data over the celestial sphere. The second has a statistical meaning, as we consider that homogeneity exists when the residual between a given catalog and the ICRF behaves as a unimodal pure Gaussian. We use a nonparametrical method, which enables us to homogeneously extend the statistical properties of the residual over the entire sphere. This intermediate adjustment allows for subsequent computation of the coefficients for any parametrical adjustment model that has a higher accuracy and greater stability, and it prevents problems related with direct adjustments using the models. On the other hand, the homogeneity of the residuals in a catalog is tested using different weights. Our procedure also serves to propose the most suitable weights to maintain homogeneity in the final results. We perform a test using the ICRF-Ext2, JPL, and USNO quasar catalogs.

*Results.* We show that a combination of catalogs can only be homogeneous if we configure the weights carefully. In addition, we provide a procedure to detect inhomogeneities, which could introduce deformities, in these combined catalogs.

*Conclusions.* An inappropriate use of analytical adjustment methods provides erroneous results. Analogously, it is not possible to obtain homogeneous-combined catalogs unless we use the adequate weights.

**Key words.** astrometry – celestial mechanics – methods: data analysis

## 1. Introduction

The International Celestial Reference Frame (ICRF; Ma 1998) was officially adopted to replace its predecessor, the FK5 Frame, as the realization of the International Celestial Reference System (ICRS; Arias 1995) at the 23th International Astronomical Union (IAU) General Assembly in 1997. Catalogs of radio source positions (RSC) derived from VLBI observations have been used by the IAU to establish the ICRF since 1998. The first realization of the ICRF, the ICRF1, consisted of 608 extragalactic radio sources and 294 candidate sources to make future improvements possible and 102 additional unstable sources.

Definitions of the ICRF are no longer related to the equinox or equator, but the IAU has recommended that the new reference frame should be consistent with the former FK5 system for the sake of continuity. The principal plane should be close to the mean equator at J2000.0 and the origin of right ascension should be close to the dynamical equinox at J2000.0. Feissel 1998 concluded that the final orientations of the ICRF axes were coincident with those of the FK5 J2000.0 system within the uncertainties of the latter catalog.

In 1999, the first extension of ICRF1, the ICRF-Ext1 (IERS 1999) was released with 59 new sources as candidate sources.

The positions and errors of the ICRF1 defining sources remained unchanged. The second extension of the ICRF1, the ICRF-Ext2 (Fey et al. 2004), added 50 new sources, and the positions of ICRF candidates and other sources were revised. The positions of the 212 defining sources were kept the same as those obtained in the ICRF1. It should be pointed out that both ICRF1 extensions were similarly obtained to the first realization, as a result of analysis of the VLBI observations at a single analysis center. The orientations of the ICRF-Ext2 axes and their uncertainties, however, still remained at the same level as for the original ICRF1.

The second realization of the ICRF, the ICRF2 (Fey et al. 2009) has 3414 objects (295 defining) and presents some advantages with respect to the previous realizations of the ICRF. The sky distribution of radio sources is improved by 1383 sources in the southern hemisphere, and the source categorization is subject to a more rigorous criterion. Only 97 defining sources of the ICRF1 remain as ICRF2 defining sources, and 39 unstable sources, some of which were ICRF1 defining sources, are identified as requiring special handling with their positions treated as arc parameters in VLBI data analysis. Statistical information on the evolution of the realization of the ICRS is given in (Yu 2013).

We focus our interest in the ICRF-Ext2, which is widely used for astrometrical measurements; thus, its accuracy and stability should be assured. This requires continuous maintenance as time elapses. First, the ICRF-Ext2 needs to be more densely defined because it includes a total of 717 sources and has a largely non-uniform distribution over the sky. Second, the defining sources have to be monitored to verify whether they are still proper and stable to be used in a future realization of the ICRS.

With regard to the first question above, attempts were made to improve the accuracy of the celestial reference frame by constructing combined catalogs after the appearance of the first VLBI radio sources catalog. Different methods were used to obtain a combined RSC with (Arias et al. 1988) being the first of them, and others were proposed by Walter (1989a, 1989b) and Yatskiv (1990). In accordance with Sokolova (2007), appropriate use of individual catalogs with common sources provides an improvement of each individual catalog. Among the general aims regarding this subject, we can point out the comparative study between catalogs and ICRF defining sources by considering a rigid frame (only rotations), avoiding deformations, and considering the densification of the ICRF in a twofold sense: the increased amount of defining sources and the densification in its wider sense of the catalog.

In all Cat-ICRF comparison processes, we intend to separate the residuals into two components: systematic errors (signal) and random errors (noise). During the study process of each individual catalog and the subsequent process of obtaining a combined catalog, we may find some problems that could provide unsatisfactory results. To study systematic errors, a parametrical model is a common choice and then the least squares method is indiscriminately used to obtain the coefficients of the chosen model, either geometrical (rotation or rotation plus deformation) or analytical (spherical harmonics development of Fourier-Legendre, for instance). Related to this, it may be extremely important that:

- a) The nonhomogeneity in the spatial distribution since it causes functional orthogonality does not turn into algebraical orthogonality in the discrete case. This problem is especially serious when we estimate high order harmonics.
- b) The mean quadratic error is the sum of the mean squared bias and the variance. An artificial decrease of the bias (If the information is biased, the model must consider this possibility) implies an increase in the variance, and this could introduce artificial and undesirable deformations.

For suitable treatment of these possible complications, we use a double meaning for the word “homogeneous”. A first meaning refers to the spatial distribution of the data. Another one is applied to the remaining “remainder” after the geometric adjustment of the Cat-ICRF residuals that must behave as a unimodal Gaussian distribution with null mean. A detailed study of the first sense was developed in our papers (Marco 2004a) for HIPPARCOS-FK5, (Marco 2004b) for HIPPARCOS-Tycho2 and (Martínez 2009) for the estimation of the HIPPARCOS-FK5 spin. All these works were carried out in the context of a complete adjustment between catalogs with many common stars. In our present case, there are differences in the number of reference positions, because we have defining sources with higher quality and the rest are common objects. In the case of few common sources, the previously mentioned problems are worse, and we must be especially careful if we want to obtain the aims proposed above.

Let us go back to the decomposition for every catalog of the residual in the signal with the random part that must be normally

distributed. A combination of catalogs take us generally to a random part that consists of one Gaussian mixture, which indicates the origin of the different sources used in the compilation. This shows that a combined catalog is not generally homogeneous in the second sense given to the word.

As we establish later, there is a specific assignment of weights that provides a residual that is a pure Gaussian distribution after the adjustment. This problem itself is very difficult to study. We have been able to make use of several procedures, such as a Gauss-induced by least squares or search, using minimization of the distance, as in Sfikas (2005). We consider the result given by the (not necessarily the most efficient) method, where homogenization is (at least) assured in a radius equal to the standard deviation and is centered in the mean. This problem would not be complete if we do not provide an answer to the inverse problem: given a nonhomogeneously combined catalog, how can we find the different populations from which it originates? The answer is not evident. We respond to another problem that may answer the former one: given one Gaussian mixture, how could we obtain the summands in which it decomposes? This part deals with the study of the random component of the residuals. These two questions are considered in the following section. The presentation of the problem is covered in greater detail and more suitable terminology and answers are provided to both previous questions. In the Appendix, a rigorous proof is included regarding the convergence of methods in a more general case.

Concerning the previously mentioned decomposition of the mean-quadratic error in the mean-squared bias and the variance from a statistical point of view, several approaches to this problem exist. The usual one consists of minimizing empirical errors, choosing a parametrical adjustment model. In this case, the coefficients of the model are very susceptible to changes in the initial data, being sensitive to spatial distribution and to the individual errors that they contain. On the other hand, we propose an approach to the problem with better stability properties. It implies the use, as an intermediary, of a non parametrical adjustment. We see some additional advantages of using this approach, which is the aim of the third section. In the fourth section, we give a summary of the parametrical approach. In the fifth section, we consider all the previous considerations for the use of two quasar catalogs to improve the ICRF. This paper concludes with a brief summary of conclusions and a final Appendix.

## 2. Homogeneity and errors in combined catalogs

Let  $\{x_i, i = 1, 2, \dots, n\}$  be the ICRF positions, and let  $\{x_i^1, i = 1, 2, \dots, n\}$ ,  $\{x_i^2, i = 1, 2, \dots, n\}$  be the corresponding values in the two other catalogs. We work with the original positions  $\{x_i, i = 1, 2, \dots, n\}$  and the residuals  $\{y_i, i = 1, 2, \dots, n\}$  and  $\{z_i, i = 1, 2, \dots, n\}$  for Catalog1-ICRF or Catalog2-ICRF. Suppose that we have

$$y_i = m^{[1]}(x_i) + \varepsilon_i^{[1]}, z_i = m^{[2]}(x_i) + \varepsilon_i^{[2]}, \quad (1)$$

where each  $\varepsilon_i^{[j]}$ ,  $j = 1, 2$  is a normal random variable with  $N(0, \sigma_1^2)$ ,  $N(0, \sigma_2^2)$  respectively and  $m^{[1]}$ ,  $m^{[2]}$  models of adjustment. In general, any linear combination of random variables (RV henceforth) is not a pure Gaussian, but a Gaussian mixture. In the particular case of being a pure Gaussian, we say that we have a homogeneous combined catalog.

Thus, we want to find a new RV  $U$  and a model  $m$  from the RV  $Y, Z$  and the models  $m^{[1]}$ ,  $m^{[2]}$  so that the residual  $U - m(X)$  verifies that it is a pure Gaussian RV with null expectation and that  $\sigma_U^2 < \sigma_X^2 + \sigma_Y^2$ .

Taking  $m(X) = \alpha_1 m^{[1]}(X) + \alpha_2 m^{[2]}(X)$  and  $U = \alpha_1 Y + \alpha_2 Z$ , it is evident that

$$U - m(X) = \alpha_1 (Y - m^{[1]}(X)) + \alpha_2 (Z - m^{[2]}(X)), \quad (2)$$

and this is generally a Gaussian mixture.

In a first step, the usual procedure employed to obtain a densified catalog from others that are referred to a main catalog (with few but very good points from a qualitative point of view) is based on this briefly explained method. Two questions arise from this point:

(Question 2.1.) There are many ways to consider the models  $m^{[1]}$ ,  $m^{[2]}$ , and the weights. If we have chosen the models, which are the optimum weights to account for the two previous conditions?

(Question 2.2.) We consider the word homogeneity applied to our problem in a dual sense: the first deals with the spatial distribution of the data over the celestial sphere. The second has a statistical meaning as we consider that homogeneity exists when the residual between a given combined catalog and the ICRF behaves as a unimodal pure Gaussian. If we have a catalog obtained from another two catalogs, is it possible to know if its construction has been homogeneous?

In the following subsections we try to answer these two questions.

### 2.1. Compiling an accurate catalog from other catalogs

Question 2.1 can be solved in the desired sense: if we want to obtain a Gaussian residual by supposing  $Y$  is independent of  $Z$ , we have

$$\sigma_U^2 = \alpha_1^2 \sigma_Y^2 + \alpha_2^2 \sigma_Z^2. \quad (3)$$

If we take

$$\alpha_1 = \frac{\sigma_U^2}{2\sigma_Y^2}, \alpha_2 = \frac{\sigma_U^2}{2\sigma_Z^2}, \quad (4)$$

then, we obtain the first property with  $\sigma_U^2 = \sqrt{\frac{2}{\frac{1}{\sigma_Y^2} + \frac{1}{\sigma_Z^2}}}$ :

$$U - m(X) = \frac{1}{\sqrt{2\pi}\sigma_U} \exp\left(-\frac{1}{2\sigma_U^2}x^2\right). \quad (5)$$

It can be shown that the previous properties are satisfied, because we have obtained a normal residual (at least in a radius equal to the typical deviation and centered in the mean), and the variance  $\sigma_U^2$  verifies the desired second property:

$$\sigma_U^2 < \sigma_X^2 + \sigma_Y^2. \quad (6)$$

### 2.2. A compiled catalog and its homogeneity

To obtain a new compiled catalog, two or more catalogs are linearly combined so that the errors generally propagate to the final catalog as a Gaussian mixture. We now deal with the inverse problem: how can we find the weights and the variances of a sum of Gaussians?

To this aim, we take the function  $f$  defined as a linear combination of the Gaussian probability density function with different variances and null expectations (this is not necessary, but it is the norm for these kinds of problems). Now, we ask ourselves if a

residual is a sum of Gaussians and, in this case, how can it be determined. Let a two-Gaussian mixture distribution be

$$f(x) = \frac{c_1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{1}{2\sigma_1^2}(x-\mu_1)^2\right) + \frac{c_2}{\sqrt{2\pi}\sigma_2} \exp\left(-\frac{1}{2\sigma_2^2}(x-\mu_2)^2\right), \quad (7)$$

where  $\sigma_1^2, \sigma_2^2$  are the variances,  $\mu_1, \mu_2$  the mathematical expectations and  $c_1, c_2$  the weights, and where we suppose  $\sigma_1 < \sigma_2$ . In addition, the most usual situation in catalog problems is  $\mu_1 = \mu_2 = 0$ , and so, we make this assumption (Even when we consider a more general case, the method is still valid as we show in the Appendix). If we define  $\phi_0(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right)$ , then the  $k$ th-derivative is

$$\phi_0^{(k)}(x) = (-1)^k H_k(x) \phi_0(x), \quad (8)$$

where  $H_k$  is the  $k$ th-Hermite's polynomial (Ayant 1971). (The sign is included for convenience.) If we denote  $h_k$  as the corresponding  $k$ th-Hermite function, we obtain the following for the derivatives of the Gaussian mixture:

$$f^{(k)}(x) = \frac{c_1}{\sigma_1^{k+1}} h_k\left(\frac{x}{\sigma_1}\right) + \frac{c_2}{\sigma_2^{k+1}} h_k\left(\frac{x}{\sigma_2}\right). \quad (9)$$

We consider the properties of the Hermite functions:

$$\begin{aligned} h_{2r+1}(0) &= 0 \\ h_{2r}(0) &< 0 \text{ for } r \text{ odd} \\ h_{2r}(0) &> 0 \text{ for } r \text{ even.} \end{aligned} \quad (10)$$

For  $k = 2r$ , we see that the summand containing the least variance dominates:

$$\begin{aligned} \frac{f^{(k)}(x)}{\frac{c_1}{\sigma_1^{k+1}} h_k\left(\frac{x}{\sigma_1}\right)} &= 1 + \frac{\frac{c_2}{\sigma_2^{k+1}} h_k\left(\frac{x}{\sigma_2}\right)}{\left(\frac{\sigma_2}{\sigma_1}\right)^{k+1} h_k\left(\frac{x}{\sigma_1}\right)} \Rightarrow \\ \frac{f^{(k)}(0)}{\frac{c_1}{\sigma_1^{k+1}} h_k(0)} &= 1 + \frac{\frac{c_2}{\sigma_2^{k+1}}}{\left(\frac{\sigma_2}{\sigma_1}\right)^{k+1}} \xrightarrow{k \rightarrow \infty} 1. \end{aligned} \quad (11)$$

Since we assumed  $\mu_1 = \mu_2 = 0$  and  $\sigma_1 < \sigma_2$ , then it follows that  $f^{(2r)}$  has a local extreme (also absolute) at  $x = 0$ . To compute the estimation of the least variance, it suffices to apply

$$\widehat{\sigma}_1^2 \equiv \left| \frac{a_{k+1}}{a_k} \frac{f^{(2k)}(0)}{f^{(2k+2)}(0)} \right|, \quad (12)$$

where  $a_k$  is the independent term of  $H_{2k}$ . The computation of  $\widehat{c}_1$  follows from

$$\widehat{c}_1 \equiv \frac{f^{(2k)}(0)}{\phi_{0,\widehat{\sigma}_1}^{(2k)}(0)}, \quad (13)$$

where  $\phi_{0,\widehat{\sigma}_1}^{(2k)}(x) = \frac{1}{\sqrt{2\pi}\widehat{\sigma}_1} \exp\left(-\frac{1}{2\widehat{\sigma}_1^2}x^2\right)$ . If we apply the previous formulas to the new function,

$$f^{[1]}(x) = f(x) - \frac{1}{\sqrt{2\pi}\widehat{\sigma}_1} \exp\left(-\frac{1}{2\widehat{\sigma}_1^2}x^2\right), \quad (14)$$

then we obtain  $\widehat{\sigma}_2, \widehat{c}_2$

Though the method used to deduce the different summands of a two-Gaussian mixture is intuitive, we have introduced a brief formal proof of the most general case (considering any  $\mu_1$  and  $\mu_2$ , being them null or not) in the Appendix.

### 3. Parameter estimation with intermediate nonparametric regression

In this section, we first consider a decomposition of the error estimation in two terms: a first term involving the mean squared of the bias and a second term with the variance. For a more complete exposition of this subject, see Cucker & Smale (2002).

The second subsection deals with nonparametric estimation of the coefficients belonging to an developing approach. One can conclude that different coefficients could be computed with different bandwidths. We can also observe that it is possible to compute different integrals, while evaluating them over a small set of points (for example in a Gaussian quadrature). Thus, the reconstruction of the required function (its finite development adding the total energy of the function) may be carried out using different bandwidths to compute nonparametric approximations in a reduced set of points, where integral evaluations are required. If a predetermined minimum total energy is not reached, the coefficient is rejected, and the next coefficient is tested. Obviously, higher harmonics could not be computed depending on the properties of the initial data set. These properties of the method may be used to take advantage if a development in a certain functional basis is required. It is also useful for the case of a geometrical expression of the signal, as we see in Sect. 5.

#### 3.1. Decomposition of the error of the signal function

Let  $X, Y$  be random variables that we suppose are defined in a real compact interval such as  $I_1$  and  $I_2$ , respectively. We suppose that we are searching for a model  $y = f(x)$ . The least squares error of the model is given as  $\varepsilon(f) = \varepsilon_\rho(f) = \int_Z (y - f(x))^2 d\rho$ , where  $Z = I_1 \times I_2$  and  $d\rho$  is a measure of probability for the bidimensional RV  $(X, Y)$ . We suppose that we work with functions  $\varphi(x, y)$  to whom Fubini's Theorem is applicable. Thus, we name  $d\rho_X, d\rho(y|x)$  the marginal measure for  $X$  and the conditional of  $Y$  with respect to  $X$ . We can apply

$$\int_{I_1 \times I_2} \varphi(x, y) d\rho = \int_{I_1} \left[ \int_{I_2} \varphi(x, y) d\rho_{Y|X=x} \right] d\rho_X. \quad (15)$$

We return to RV and denote

$$m_\rho(x) = \int_{I_2} y d\rho_{Y|X=x} = E[Y|X = x]. \quad (16)$$

Obviously, the residual  $y - m_\rho(x)$  has a null mathematical expectation, and its variance is given by

$$\sigma^2(x) = \int_{I_2} (y - m_\rho(x))^2 d\rho_{Y|X=x}. \quad (17)$$

We denote the integral of this variance as  $\sigma_\rho^2 = \int_{I_1} \sigma^2(x) d\rho_X = \varepsilon(m_\rho)$ . That is, it represents the least square error of the regression function (with respect to the density  $\rho$ ). It is clear that we obtain the decomposition of the error that is associated with an approximation  $\widehat{m}$  of  $f$ , which provides the following after some algebraic operations:

$$\varepsilon(\widehat{m}) = \int_{I_1} (\widehat{m}(x) - m_\rho(x))^2 d\rho_X + \sigma_\rho^2. \quad (18)$$

In this identity (18), we see that the first term depends on the regression function  $m_\rho$  and the approximation  $\widehat{m}$ . The last term is independent of the unknown; it depends only on the measure  $\rho$  and is the lower bound of the error, since it does not depend on

the approximation as it is stated. In general, one can suppose that  $m_\rho$  belongs to some functional space that admits a development in the appropriate convergent series. On the other hand, different approaches to the estimated function  $\widehat{m}$  are possible. The main question that appears concerns the bias. One can make many assumptions about  $m_\rho$  and try different methods to find  $\widehat{m}$ . We get  $m_\rho$  approaching the different densities by means of a nonparametrical method. Nevertheless, other choices are possible. This selection requires a consideration regarding bias and variance in the approximation process (see following subsection). For the function  $\widehat{m}$ , we select a complete Hilbert space (in such a way that it has a functional basis) or Banach space with a basis or dictionary (not necessarily free with a possible redundant generator system. An infinite series may be expressed in more than one way. This is not our case, because we work with a finite sum of terms) to consider a nonlinear,  $m$ -term approximation (sum of the  $m$  terms of the basis or the dictionary, depending on the situation). The method to obtain the coefficients is considered below, and we get a rule to diminish statistical bias.

The classical approach to the regression problem may be considered a particular case of the previous one. After selecting the working functional space  $H$ , we consider

$$\min_{m \in H} \frac{1}{n} \sum_{i=1}^n (m(x_i) - y_i)^2, \text{ where } (x_i, y_i), 1 \leq i \leq n \text{ represents the}$$

finite sample. The decomposition formulas are theoretically similar, but the implementation is different, and there are very important differences between both methods. We prefer the former approach and we explain its advantages in the following section. In addition, we can mention other advantages, such as a) a more robust approach; b) the lack of need to choose an a priori model and c) an approximation of the density function (which generalizes the assignment of weights). It is interesting to remark that the classical method may provide wrong results, if the abscises are, for instance, inhomogeneously distributed. Another problem consists of requiring the recomputation of the coefficients in a model of adjustment when its order is increased, because the coefficients of lower orders are not preserved. The main problem of nonparametrical methods is related with the bandwidth selected value. Nevertheless, there are very accurate methods to solve this problem, considering the complete statistical information of the data. Another problem is the missed geometrical-analytical sense. The earlier problem is overcome when we take the nonparametrical method as an intermediate adjustment, because it contains complete statistical information of the data by means of estimating the density function.

#### 3.2. Mixed methods

The range of the covariant values (the sample of the RV  $X$  in this Sect. 3) could be an inhomogeneous distribution. If we nevertheless take a nonparametric adjustment, it is possible to use it to compute the estimation of the function in an arbitrary point. This implies that we could extend the function values regularly over the complete domain. To highlight its advantage, we consider a simple example: if we suppose, for example, that the function  $f$  belongs to a separable Hilbert space with a complete orthonormal basis  $\{\phi_i > 0, i \geq 0\}$ , then we have a unique expansion

$$f(x) = \sum_{i \geq 0} a_i \phi_i(x). \quad (19)$$

The computation of each coefficient could be performed using  $a_i = \langle f | \phi_i \rangle$  with  $\langle | \rangle$  being the inner product in the space and

$\widehat{f}$  the nonparametric (intermediate) estimation. The inner product could be approximated by a numerical integration method to be as accurate as necessary. If we consider, for example, a Gauss-type method, we must estimate the value of the function in a few number of points. Furthermore, there are different methods, which are specially prepared to estimate the optimal bandwidth for each point, which must be taken. For more information about nonparametrical methods, see Wand & Jones (1995). As previously mentioned, it is equally possible to estimate the coefficients of a geometrical model of adjustment, as we see in Sect. 5.

There is another advantage when a coefficient of a harmonic must be computed. Not only is this due to the former property but also this procedure does not make any supposition regarding the order of the actual expansion. In this sense, our method is not linear and it is possible to expand it to a great number of problems.

### 3.3. Nonparametrical estimation of coefficients of minimum energy in the Fourier case

We suppose developments  $m(x) = \sum c_k \varphi_k(x)$  and  $\widehat{m}(x) = \sum \widehat{c}_k \varphi_k(x)$ . Along this subsection we neglect, unlike the previous section, the dependence of the density  $\rho$ . The mathematical expectation of the coefficients is given by

$$E[\widehat{c}_k | X_1, \dots, X_n] = \int_{\mathbb{R}^n} \widehat{c}_k f(x_1, \dots, x_n) dx_1, \dots, dx_n, \quad (20)$$

where  $X_1, \dots, X_n$  is a set of independent random variables distributed identically like  $X$ , and  $f$  is its joint density function. We apply the relation between the coefficients, the basis, and the function:

$$\int_{\mathbb{R}^n} \left\{ \int_{\mathbb{R}} \widehat{m}(x) \varphi_k(x) \right\} f(x_1, \dots, x_n) dx_1, \dots, dx_n = \int_{\mathbb{R}} \left\{ \int_{\mathbb{R}^n} \widehat{m}(x) f(x_1, \dots, x_n) dx_1, \dots, dx_n \right\} \varphi_k(x) dx, \quad (21)$$

where we have applied Fubini's theorem. By definition, we have this transformation:

$$\int_{\mathbb{R}} E[\widehat{m}(x) | X_1, \dots, X_n] \varphi_k(x) dx = c_k + \int_{\mathbb{R}} E[\widehat{m}(x) - m(x) | X_1, \dots, X_n] \varphi_k(x) dx. \quad (22)$$

If we take a local polynomial kernel regression of third order and a Fourier series is considered, then we have

$$E[\widehat{c}_k | X_1, \dots, X_n] = c_k \left( 1 + \alpha \frac{(kh)^4}{24} \right), \quad (23)$$

where  $\alpha$  is a constant depending only on the used kernel,  $h$  is the bandwidth, see Fan & Gijbels (1992), and  $k$  is the order of the considered harmonic. This expression makes it clear: if  $h_1$  is a good selection for the  $c_1$  term, then  $(kh)^4 \simeq h_1^4$  can be used to approach the  $c_k$  coefficient with the same reliability,  $h \simeq \frac{h_1}{k}$ . For numerical purposes, this property may be of interest.

## 4. Parametrical regression methods

The different adjustment methods have been widely described in other papers and we do not repeat them here. We only mention

that different parametrical methods provide different adjustment models, following different hypotheses:

a) Geometrical model: there is an infinitesimal rotation between the initial and the final positions (Walter & Sovers 2000) if we denote  $\Delta\alpha = \alpha^{[1]} - \alpha$ ,  $\Delta\delta = \delta^{[1]} - \delta$ :

$$\Delta\alpha \cos \delta = -\varepsilon_x \cos \alpha \sin \delta - \varepsilon_y \sin \alpha \sin \delta + \varepsilon_z \cos \delta$$

$$\Delta\delta = \varepsilon_x \sin \alpha - \varepsilon_y \cos \alpha. \quad (24)$$

b) Rotation plus deformation model: the vector field of the residuals may be developed in vectorial spherical harmonics using first order truncation; see Marco (2004a).

$$\Delta\alpha \cos \delta = -\varepsilon_x \cos \alpha \sin \delta - \varepsilon_y \sin \alpha \sin \delta + \varepsilon_z \cos \delta$$

$$+ d_{1,1} \cos \alpha - d_{1,-1} \sin \alpha$$

$$\Delta\delta = \varepsilon_x \sin \alpha - \varepsilon_y \cos \alpha - d_{1,1} \sin \alpha \sin \delta$$

$$- d_{1,-1} \cos \alpha \sin \delta + d_{1,0} \cos \delta. \quad (25)$$

c) Surface spherical harmonics model: the residuals come from a function of integrable square over the unitary sphere. Then the residuals allow the widely used surface spherical harmonic developments, see Heiskanen (1967).

$$\Delta\alpha \cos \delta = \sum_{n,m} a_{n,m} Y_{n,m}(\alpha, \delta)$$

$$\Delta\delta = \sum_{n,m} b_{n,m} Y_{n,m}(\alpha, \delta). \quad (26)$$

These methods present some advantages and some problems. The main advantage is that they provide a (primary) qualitative explanation for the errors, and it is also possible to include other parameters with physical meaning.

In contrast, there are some initial problems: in many cases, the selected model is not representative of the problem and even more importantly, an inhomogeneous spatial distribution of the set of data implies instability to compute the parameters directly. Even using high orders, the results are not realistic if we have few stars. If higher order harmonics appear in the spherical harmonics case, the usual method could be wrong.

## 5. Using two different quasar catalogs to improve the ICRF-Ext2

As a numerical application of the previous sections, we have considered two input catalogs: the USNO and JPL, which contain 189 common defining sources from the ICRF-Ext2 catalog. They were submitted by JPL (Caltech/NASA Jet Propulsion Laboratory, USA), and USNO (US Naval Observatory, USA). A brief description of the input catalogs is given in Table 1 (Sokolova & Malkin 2007).

This section is divided into two subsections. The first one concerns the process of building a combined catalog from the two chosen input catalogs. In the second subsection, we propose a procedure of acceptance from candidate sources to reference sources. We check that the candidate sources selected using our method from the ICRF-Ext2 are indeed included in the ICRF2 as defining or non defining sources.

### 5.1. The use of JPL and USNO catalogs to build a combined catalog

In this study, we considered only the sources which have at least 15 observations in two sessions. We have not included some

**Table 1.** Input catalogs.

Center	Software	Time Span month/year	$N.$ delay	$N.$ sources
JPL	MODEST	10/1978 01/2005	3 575 847	734(2)
USNO	Calc/Solve	09/1979 01/2005	4 252 684	943(207)

**Notes.** The last column shows the number of sources in the catalog and the number of reference sources used to tie the orientation of the catalog to the ICRF.

**Table 2.** Rotation and deformation parameters for the USNO and JPL in  $\mu\text{as}$ .

	USNO	JPL
$\varepsilon_x$	105.90	34.54
$\varepsilon_y$	-1.19	6.87
$\varepsilon_z$	-13.04	-82.02
$d_{1,-1}$	0.21	-19.73
$d_{1,1}$	35.09	-29.33
$d_{1,0}$	10.18	10.28

reference sources in our calculus that present oddly high residuals. All values are given in  $\mu\text{as}$ . The values of the initial arithmetic mean  $\mu$  and standard deviation  $\sigma$  are as follows: for the USNO catalog, they are  $\mu_\alpha = 10.2$  and  $\sigma_\alpha = 246.8$ ,  $\mu_\delta = -5.2$  and  $\sigma_\delta = 307.5$ . For the JPL catalog they are  $\mu_\alpha = -32.6$  and  $\sigma_\alpha = 341.5$ ,  $\mu_\delta = -12.1$  and  $\sigma_\delta = 354.4$ . It seems that JPL is not completely homogeneous as defined in Sect. 2 and it could require a more careful study. We have included it in our adjustment, because it serves to show the strength of our method.

With respect to the study of the residuals, we have chosen to carry out a preliminary kernel nonparametric (KNP henceforth) adjustment for the  $\Delta\alpha \cos \delta$  and  $\Delta\delta$  in both catalogs and a vectorial spherical harmonics (VSH henceforth) of first order for the adjustment model (25). Then, we apply our mixed method as it has been explained in Sects. 3.1 and 3.2 by taking  $m_\rho = [\Delta\alpha \cos \delta, \Delta\delta]^t$ ,  $\widehat{m} = [m_\alpha, m_\delta]^t$  for each catalog in the first summand of the expression (18). The existence of deformations has required the use of a correction for each catalog given by

$$\min_{C^i} \int_{S^2} \left\{ \left[ (\Delta\alpha \cos \delta)^{(i)} - m_\alpha^{[i]}(\alpha, \delta) \right]^2 + \left[ (\Delta\delta)^{(i)} - m_\delta^{[i]}(\alpha, \delta) \right]^2 \right\} dS, \quad (27)$$

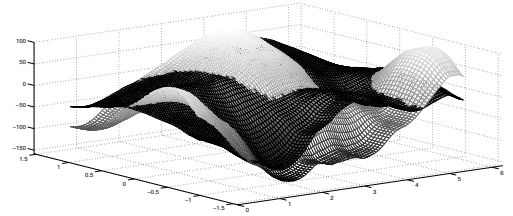
where  $C^i$  are the coefficients of the models  $m_\alpha^{[i]}$  and  $m_\delta^{[i]}$  with  $i = 1$  (USNO) and  $i = 2$  (JPL). The results for the coefficients of the VSH of first order are listed in the Table 2, where it can be seen that JPL provides high coefficients for the deformation. This must be considered in future studies. The coefficients listed in Table 2 are obtained by applying condition (27) and solving the resulting normal equations. They consist of different integrals, which are evaluated over a grid of equidistant points over the sphere where the KNP values for  $\Delta\alpha \cos \delta$  and  $\Delta\delta$  are computed.

Next, we consider only the rotations. We subtract the corrections provided by the rotations to the initial position to obtain the intermediate catalogs  $USNO^1$  and  $JPL^1$ :  $Cat^1 = Cat - correction$ , where these corrections depend only on the rotations. The adjustment itself is given by

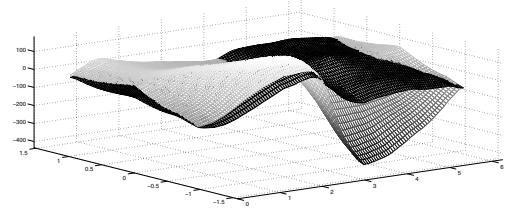
$$\begin{aligned} cat_1^1 - (ICRF - Ext2) &= m^{[1,1]} + \varepsilon^{[1,1]} \\ cat_2^1 - (ICRF - Ext2) &= m^{[2,1]} + \varepsilon^{[2,1]}. \end{aligned} \quad (28)$$

**Table 3.** Mean squared errors over the sphere after rotation adjustment.

	USNO		JPL		Combined catalog
	Initial	Final	Initial	Final	
$\mu_\alpha$	7.4	7.4	-1.4	-1.4	
$\sigma_\alpha$	61.1	36.4	99.5	53.7	35.6
$\mu_\delta$	-14.7	-14.7	-8.3	-8.3	
$\sigma_\delta$	113.9	41.5	87.3	77.0	41.0



**Fig. 1.** Differences in (ICRF-Ext2)-USNO for  $\Delta\alpha \cos \delta$  (in  $\mu\text{as}$ ). The clear surface represents the initial differences, the dark surface represents the differences after the correction given by the rotations. Notice the rank for differences in  $\Delta\alpha \cos \delta$ .

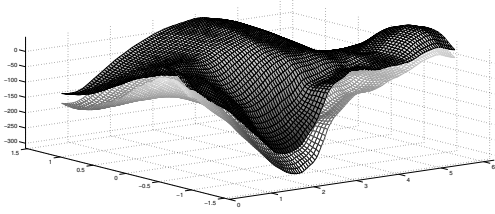


**Fig. 2.** Differences in (ICRF-Ext2)-USNO for the  $\Delta\delta$  (in  $\mu\text{as}$ ). The clear surface represents the initial differences, the dark surface represents the differences after the correction given by the rotations.

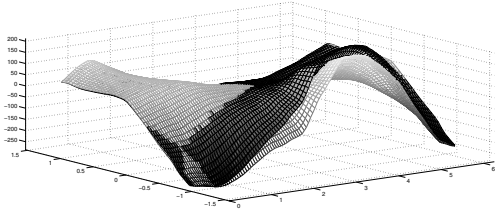
We use now the term WRMS that denotes weighted root mean squared. In our case, the function uses the weights assigned by the KNP adjustment. With regard to the WRMS in the entire sphere where we have used numerical integration and a KNP adjustment, the initial WRMS in right ascension is 61.1 and the final value is 36.4. The initial WRMS in declination is 113.9, and the final value is 47.9 for the USNO. The initial WRMS in right ascension is 99.5, and the final value is 61, while the initial WRMS in declination is 87.3, and the final value is 78.7 for the JPL. A clear summary of these data may be seen in Table 3.

The residual noise obtained after the adjustment (for each  $\Delta\alpha \cos \delta$ ,  $\Delta\delta$ ) may be observed in Figs. 1 to 4.

Prior to proceeding with the proposal of a combined catalog, we consider an example to show that our method detects inhomogeneities (in the second sense given in Sect. 2) in building a



**Fig. 3.** Differences in (ICRF-Ext2)-JPL for the  $\Delta\alpha \cos \delta$  (in  $\mu\text{as}$ ). The clear surface represents the initial differences, the dark surface represents the differences after the correction given by the rotations.



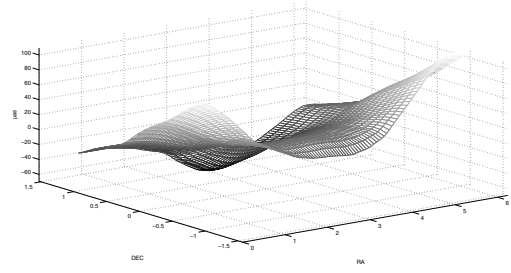
**Fig. 4.** Differences in (ICRF-Ext2)-JPL for the  $\Delta\delta$  (in  $\mu\text{as}$ ). The clear surface represents the initial differences; the dark surface represents the differences after the correction given by the rotations.

combined catalog. Taking  $0.5m^{[1,1]} + 0.5m^{[2,1]}$  only for  $\Delta\alpha \cos \delta$  residuals belonging to the USNO and the JPL catalogs, the corresponding error is  $0.5\varepsilon^{[1,1]} + 0.5\varepsilon^{[2,1]}$ , which distributes as a Gaussian mixture with weights  $c_1 = c_2 = 0.5$ , standard deviations  $\sigma_1 = 36.4$ ,  $\sigma_2 = 53.7$ , and mathematical expectations  $\mu_1 = 7.4$ ,  $\mu_2 = -1.4$ . We use the procedures described in Sect. 2. As the use of high orders derivative has been necessary, we have introduced some variations. To ensure the results, we build a third order local KNP over equally spaced points and consider this function over more points to progressively increase the order of the local KNP (depending on the required precision to stabilize the desired decimal digit). Alternatively, we have applied formulas (12) and (13) for  $k = 0, 2, 4, 6$  and an Aitken approximation to obtain  $\mu^1 = 7.4$ ,  $\sigma^1 = 36.3$ ,  $c^1 = 0.5$ ,  $\mu^2 = -1.4$ ,  $\sigma^2 = 53.7$ , and  $c^2 = 0.5$ . This means that we have obtained the two populations used in the compilation of the combined catalog. We conclude that the construction of the compiled catalog has not been homogeneous in the second definition as given in Sect. 2.

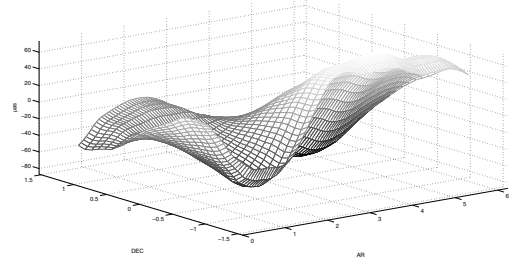
Returning to the aim of compiling a combined catalog, we take the weights proposed by (4). The final residuals may be seen in Figs. 5 and 6. The final WRMS for the residuals are 35.6 for the right ascension and 41 for the declination.

## 5.2. Method to densify the reference

Finally, we propose a procedure to densify the reference frame. We work with the candidate sources of the ICRF-Ext2, but it could be also applied to other sources. For this purpose, we apply a simple test to discover if a source belongs to a determined population. The procedure is as follows: the obtained combined catalog provides residuals for  $\Delta\alpha \cos \delta$  and  $\Delta\delta$  distributed as  $N(0, 35.6)$  and  $N(0, 41.0)$  respectively. The corresponding density functions are Gaussians, and we can consider, for example, a 30% of the area centered on the corresponding mean (the radius of each interval is determined by each standard deviation). This area is bounded by the function and two abscises that represent



**Fig. 5.** Residuals in the final combined catalog for the  $\Delta\alpha \cos \delta$  (in  $\mu\text{as}$ ).



**Fig. 6.** Residuals in the final combined catalog for the  $\Delta\delta$  (in  $\mu\text{as}$ ).

errors in  $\mu\text{as}$ . If the expected error for a candidate source is comprised in the mentioned interval, it is reasonable for this candidate to be considered as a new reference source. In any case, a more strict condition is considered in the next steps.

Considering the previous guidelines, each final function of residuals for  $\Delta\alpha \cos \delta$  and  $\Delta\delta$  is a surface. Thus, we can project its contour lines in the  $XY$  plane. In particular, we focus on the contour lines signaled by the extremes of the intervals mentioned in the previous paragraph. This may also be applied individually for each catalog, and each individual candidate source of the catalogs. The steps to follow are

Step 1: we compute  $cat_1^C - (ICRF - Ext2) - m_{G_1}^{[1]}$ , where the  $C$  refers to the candidate sources and  $G_1$  is the rotations provided by the USNO catalog (in  $\Delta\alpha \cos(\delta)$  and  $\Delta\delta$ ). As previously mentioned, each function of residuals is represented by a surface. The valid surface levels from the point of view given in the previous guidelines are the ones included in the interval  $(-29, 44) \mu\text{as}$  for  $\Delta\alpha \cos \delta$  and  $(-55, 27) \mu\text{as}$  for  $\Delta\delta$ .

Step 2: we compute  $cat_2^C - (ICRF - Ext2) - m_{G_2}^{[2]}$ , where  $C$  refers to the candidate sources and  $G_2$  is the rotations provided by the JPL catalog. For this catalog, the resulting intervals are  $(-55, 52) \mu\text{as}$  for  $\Delta\alpha \cos \delta$  and  $(-85, 68) \mu\text{as}$  for  $\Delta\delta$ .

Step 3: now, we include the candidate sources in our contour map and we take the ones that are included in the intersection of the areas determined by all the contour lines as preliminary new reference sources.

Step 4: finally, we substitute the selected sources in the final residual function associated with the combined catalog. All the results are on the order of the corresponding typical deviation in right ascension and declination. Only the second source presents an error that is 1.5 times the final standard deviation in  $\alpha$ , and the fifth source is twice the final standard deviation in declination. We must consider that the selection criteria has been very strict, taking sources that simultaneously accomplish the

**Table 4.** Selected candidate sources.

	$\alpha$	$\delta$	ICRF designation	IERS name
Source 1	00h 06 m in 13.892889 s	-06° 23' 35.33532"	ICRF J000613.8-062335	0003-066
Source 2	01h 13 m in 43.144949 s	02° 22' 17.31635"	ICRF J011343.1+022217	0111+021
Source 3	01h 49 m in 22.370886 s	05° 55' 53.56866"	ICRF J014922.3+055553	0146+056
Source 4	10h 24 m in 44.809590 s	19° 12' 20.41561"	ICRF J102444.8+191220	1022+194
Source 5	18h 00 m in 24.765366 s	38° 48' 30.69753"	ICRF J180024.7+384830	1758+388

condition given in the previous steps in both right ascension and declination and for both catalogs at the same time. This is why the results of both discrepancies are not relevant when we consider the final residual conjunctly in  $\Delta\alpha \cos \delta$  and  $\Delta\delta$ .

The verification using new catalogs may lead to the definitive choice of the source. In our case, we obtain the candidate sources listed in Table 4, where the names and the proposed position are given. It is remarkable that sources 1, 2 and 3 have been included among the 922 ICRF2 non-defining sources, while sources 4 and 5 have been included among the 295 defining sources of the ICRF2 (Fey et al. 2009).

## 6. Conclusions

The compilation of a large quasar catalog from other input catalogs is an important trend in the astronomical context. In this sense, it is interesting to study the initial catalogs, compared to the main catalog (in our case, the ICRF-Ext2 Catalog), to determine its inner homogeneity. To obtain a homogeneous combined catalog, it is necessary to use adequate weights. Reciprocally, it is possible to determine if a combined catalog has been obtained homogeneously or not. To this aim, we have studied both subjects in Sect. 2, which have been applied in Sect. 5.

A mixed method using nonparametrical regression to build an intermediate catalog presents many advantages to obtain accurate parameters. These advantages are as follows: a) a more robust approach in the sense that the computation of the coefficients is carried out by means of functional scalar products over the sphere. The homogeneous grid selection preserves the functional orthogonality in the process of discretization. b) There is no need to choose an a priori model. c) The approximation of the density function by means of a KNP (which generalizes the assignment of weights). These subjects have been studied in Sect. 3, and we can remark that different parametrical methods (geometric or analytic) may be employed to study global and local properties of the used catalogs to obtain a very good compiled catalog. In Sect. 4, we have briefly described the most currently parametrical models of adjustment.

The mixed method with an adequate weights selection has provided us a combined catalog with little final residuals, as seen in text.

We have also proposed a method to decide if a given candidate source may be coherently accepted as a new reference source with the obtained results for the combined catalog. The five sources proposed using our method have been included in the ICRF2, which includes two new defining sources and supports the reliability of our methods. Further studies may be carried out to explore other cases. For example, one can conduct a deeper study using surface or vectorial spherical harmonics of high orders. We can also include other input catalogs in the computation, which would not affect the method.

*Acknowledgements.* Part of this work was supported by a grant P1-1B2009-07 from Fundacio Caixa Castello BANCAIXA and a grant P1-0611455.01/1 from Bancaja.

## Appendix A:

Lemma 1: let  $f, g$  be continuous and derivable in a given interval with the possible exception of a finite number of points. Let us assume that  $f(a) = g(a)$  and  $f'(a) = g'(a)$ , which verifies:

$$f''(x) + \alpha(x)f(x) = 0 \quad (\text{A.1})$$

$$g''(x) + \beta(x)g(x) = 0. \quad (\text{A.2})$$

If we suppose that  $\alpha(x) > \beta(x)$ , then

a) For every subinterval  $[a, c]$  where  $f(x), g(x) > 0$ , we have  $f(x) < g(x)$ .

b) For every subinterval  $[a, c]$  where  $f(x), g(x) < 0$ , we have  $f(x) > g(x)$ .

c) Let the interval  $[a, d]$  contain point  $c$ , where  $f(c) = 0$ . We suppose that  $f(x) > 0$  if  $x < c$ ,  $f(x) < 0$  if  $x > c$ ,  $g(x) > 0$  for  $x > a$ ,  $f'(d) = 0$  and  $g'(x) < 0$  if  $x > a$ . Then  $-g(x) \leq f(x) \leq g(x)$  for  $a < x < d$ .

Proof of a):

If we multiply the first equation by  $g(x)$ , the second by  $f(x)$ , and then we subtract, we get

$$f''(x)g(x) - g''(x)f(x) = (\beta(x) - \alpha(x))f(x)g(x). \quad (\text{A.3})$$

Given any  $x$  where  $a < x \leq c$ , we integrate both members of the equality:

$$\int_a^x [f''(x)g(x) - g''(x)f(x)] dx = \int_a^x (\beta(x) - \alpha(x))f(x)g(x) dx < 0 \quad (\text{A.4})$$

and we obtain:

$$f'(x)g(x) - g'(x)f(x) < 0. \quad (\text{A.5})$$

Thus,  $\frac{f(x)}{g(x)}$  is a strict decreasing function that reaches a value one for  $x = a$ . It is then necessary that  $f(x) < g(x)$  for  $x > a$ , because both functions are positive in the subinterval.

Proof of b) We follow the development of Lemma 1 until we arrive at

$$0 < \frac{f(x)}{g(x)} < 1, \quad (\text{A.6})$$

since we assume that  $g(x) < 0$ , then  $g(x) < f(x)$ .

Note: let us suppose the differential equation  $\tilde{g}''(x) + \beta(x)\tilde{g}(x) = 0$ , where  $\tilde{g}(a) = -g(a)$ ,  $\tilde{g}'(a) = -g'(a)$ . Then  $\tilde{g}(x) = -g(x)$  in the interval of definition. It is evident that this result is equally true for function  $f$ .



Proof of c) For the subinterval  $[a, c]$  we can apply a). For the subinterval  $[c, d]$ , we consider function  $\tilde{g}(x)$  from the former note. Following the steps of the proof given in a), we have for each  $x$ , where  $c < x < d$ :

$$\int_x^d [f''(x)\tilde{g}(x) - \tilde{g}''(x)f(x)] dx = \int_x^d (\beta(x) - \alpha(x))f(x)\tilde{g}(x)dx < 0. \quad (\text{A.7})$$

In  $(c, d)$ , we have  $f(x) < 0, \tilde{g}(x) < 0$ . Therefore,

$$[f'(d)\tilde{g}(d) - \tilde{g}'(d)f(d)] - [f'(x)\tilde{g}(x) - \tilde{g}'(x)f(x)] < 0. \quad (\text{A.8})$$

Hence,

$$[f'(x)\tilde{g}(x) - \tilde{g}'(x)f(x)] > -\tilde{g}'(d)f(d) > 0, \quad (\text{A.9})$$

because  $f(d) < 0$  and  $\tilde{g}'(d) > 0$ . With this, we deduct that  $\frac{f(x)}{\tilde{g}(x)}$  is decreasing in  $(c, d)$  with the quotient null in  $x = c$  and both expressions are negative, reaching the maximum at  $x = d$ . Thus,  $\frac{f(x)}{\tilde{g}(x)} > 0$  for  $c < x < d$ . In addition, if  $\frac{f(x)}{\tilde{g}(x)} \geq 1$ , then  $f(x) \leq \tilde{g}(x)$ , and we would obtain  $\tilde{g}(c) \geq 0$ , which is absurd. Consequently,

$$0 < \frac{f(x)}{\tilde{g}(x)} < 1 \rightarrow f(x) > \tilde{g}(x) \text{ for } c < x < d. \quad (\text{A.10})$$

Thus, function  $f(x)$  verifies  $-g(x) \leq f(x) \leq g(x)$  between two consecutive extremes of  $f$ . This argument may be extended to all real domain.

Corollary 1: we assume the same general conditions about  $f$  and  $g$  from the former Lemma. If the first differential equations have been verified with  $\beta > 0, g > 0$ , and  $\beta \leq \alpha$  and we have the additional conditions  $f(a) = g(a), f'(a) = g'(a)$  or  $f(a) = -g(a), f'(a) = -g'(a)$  then  $|f(x)| \leq g(x)$  for  $x \in [a, b]$ .

Lemma 2: let  $z_n(x) = \exp(-\frac{1}{4}x^2) \frac{H_n(x)}{H_n(0)} \equiv \frac{h_n(x)}{h_n(0)} \exp(\frac{1}{4}x^2)$  where  $n = 2r$  (which includes the particular case  $n = 4r$ ). Then, it is verified that

$$z_n''(x) + \left(n + \frac{1}{2} - \frac{1}{4}x^2\right)z_n(x) = 0 \text{ con } z_n(0) = 1, z_n'(0) = 0.$$

Proposition 1: let  $g(x) = \exp(-\frac{1}{\alpha}x^2)$  with  $\alpha > 0$ . Then we have for any  $x$  that  $\left|\frac{h_n(x)}{h_n(0)}\right| \leq g(x)$  for a certain  $\alpha > 0$ . For example,  $\alpha = 2\sqrt{5}$ . The value of  $\alpha$  may be provided with more accuracy:  $\alpha$  is a value such that the inflection point of  $h_2(x)$ , as indicated in  $(\sqrt{3+6\sqrt{3}})/2 \approx 3.3165$ , is a common point for both functions, which provides  $\alpha \approx 4.4234$ , a very close value to  $2\sqrt{5} \approx 4.4721\dots$

Proof: considering that function  $g$  verifies that

$$g''(x) + \left(\frac{2}{\alpha} - \frac{4}{\alpha^2}x^2\right)g(x) = 0 \quad (\text{A.11})$$

and  $g(0) = 1, g'(0) = 0$ , we apply corollary 1 if

$$\frac{2}{\alpha} - \frac{4}{\alpha^2}x^2 \leq \frac{1}{2} + n - \frac{1}{4}x^2, \quad n = 2r, r \geq 1. \quad (\text{A.12})$$

The least value of the second member of the inequality is obtained for  $n = 2$ , (when  $x$  is fixed), and we obtain the sufficient condition:

$$\frac{2}{\alpha} - \frac{4}{\alpha^2}x^2 \leq \frac{5}{2} - \frac{1}{4}x^2. \quad (\text{A.13})$$

This may be immediately verified: if we take, for example,  $\alpha = 2\sqrt{5}$ , the inequality verifies for  $x \leq 41$ . For higher values of  $x$ ,  $g(x)$  and  $z_n(x)$  are negligible for any desired accuracy.

Proposition 2: let  $h_{2r,1}$  the first positive zero of the Hermite function  $h_{2r}(x)$ . We have that (Ayant 1971)

$$0 < h_{2r,1} < \frac{\pi}{\sqrt{8r+2}}. \quad (\text{A.14})$$

Theorem 1: let

$$f(x) = c_1 \exp(-0.5x^2) + \frac{c_2}{\sigma} \exp\left(-0.5\left(\frac{x-\mu}{\sigma}\right)^2\right), \quad (\text{A.15})$$

$$c_1 + c_2 = \frac{1}{\sqrt{2\pi}}, c_1 > 0, c_2 > 0$$

and let us suppose that  $\sigma > 1$ . Then if we denote  $\xi_r$  as the point on the  $x$ -axis where the higher relative extreme is in absolute value of the  $2r$ -th derivative of  $f$ , we have that

$$\lim_{r \rightarrow +\infty} \xi_r = 0. \quad (\text{A.16})$$

Proof: to enable the proofs, we take  $4r$ , instead of  $2r$ , because we can work with positive maximum without loss of generality in this case. Let us suppose that  $f^{(4r)}(\xi_r) > 0$  and also that  $\xi_r > 0$ . First, note that

$$f^{(4r)}(x) = c_1 h_{4r}(x) + \frac{c_2}{\sigma^{4r+1}} h_{4r}\left(\frac{x-\mu}{\sigma}\right) \quad (\text{A.17})$$

and, dividing by  $c_1 h_{4r}(0)$ , we have

$$\frac{f^{(4r)}(x)}{c_1 h_{4r}(0)} = \frac{h_{4r}(x)}{h_{4r}(0)} + \frac{c_2}{c_1} \frac{1}{\sigma^{4r+1}} \frac{h_{4r}\left(\frac{x-\mu}{\sigma}\right)}{h_{4r}(0)}. \quad (\text{A.18})$$

With this, we have the following boundary for any  $x$ :

$$\left|\frac{f^{(4r)}(x)}{c_1 h_{4r}(0)}\right| \leq 1 + \frac{c_2}{c_1} \frac{1}{\sigma^{4r+1}} \rightarrow 1 \text{ para } r \rightarrow +\infty, \quad (\text{A.19})$$

because  $\sigma > 1$  and  $h_{4r}(x)$  reaches the absolute maximum at  $x = 0$ . This indicates that the first summand in the decomposition of  $f^{(4r)}(x)$  is dominant. We see that this idea is confirmed in the announced asymptotic result. Given the former considerations, it is immediate that, there is a  $r_1$  for any given  $\varepsilon_1 > 0$  such that

$$\left|\frac{c_2}{\sigma^{4r+1}} \frac{h_{4r}\left(\frac{x-\mu}{\sigma}\right)}{h_{4r}(0)}\right| < \varepsilon_1 \quad (\text{A.20})$$

for  $r \geq r_1$  with an independence of  $x$ . In particular, we have

$$\left|\frac{f^{(4r)}(\xi_r)}{h_{4r}(0)} - c_1 \frac{h_{4r}(\xi_r)}{h_{4r}(0)}\right| < \varepsilon_1. \quad (\text{A.21})$$

We have  $f^{(4r)}(\xi_r) > 0$ , and the previous bound assures that the quotient  $\frac{h_{4r}(\xi_r)}{h_{4r}(0)}$  is positive. Considering the first proposition, we know that  $\frac{h_{4r}(\xi_r)}{h_{4r}(0)} \leq \exp\left(-\frac{1}{\alpha}\xi_r^2\right)$  for a certain  $\alpha > 0$ , so

$$\frac{f^{(4r)}(\xi_r)}{h_{4r}(0)} - c_1 \frac{h_{4r}(\xi_r)}{h_{4r}(0)} \geq \frac{f^{(4r)}(\xi_r)}{h_{4r}(0)} - c_1 \exp\left(-\frac{1}{\alpha}\xi_r^2\right). \quad (\text{A.22})$$

Now, let us suppose that  $\xi_r^2 > h_{4r,1}^2$  with  $h_{4r,1}$  being the first positive zero of  $h_{4r}(x)$ . Then

$$\xi_r^2 > h_{4r,1}^2 \rightarrow \exp\left(-\frac{1}{\alpha}h_{4r,1}^2\right) > \exp\left(-\frac{1}{\alpha}\xi_r^2\right) \rightarrow -c_1 \exp\left(-\frac{1}{\alpha}\xi_r^2\right) > -c_1 \exp\left(-\frac{1}{\alpha}h_{4r,1}^2\right). \quad (\text{A.23})$$

From (A.22) and (A.23),

$$\begin{aligned} \frac{c_2}{\sigma^{4r+1}} \frac{h_{4r}\left(\frac{\xi_r - \mu}{\sigma}\right)}{h_{4r}(0)} &= \frac{f^{(4r)}(\xi_r)}{h_{4r}(0)} - c_1 \frac{h_{4r}(\xi_r)}{h_{4r}(0)} \\ &> \frac{f^{(4r)}(\xi_r)}{h_{4r}(0)} - c_1 \exp\left(-\frac{1}{\alpha} h_{4r,1}^2\right). \end{aligned} \quad (\text{A.24})$$

Because we suppose that  $\xi_r$  is the absolute positive maximum of  $f^{(4r)}(x)$ , we can deduct that there is a  $\varepsilon(r)$  value such that:

$$\begin{aligned} \varepsilon_1 &> \frac{c_2}{\sigma^{4r+1}} \frac{h_{4r}\left(\frac{\xi_r - \mu}{\sigma}\right)}{h_{4r}(0)} > \Xi_r \equiv \\ &\frac{f^{(4r)}(\xi_r)}{h_{4r}(0)} - c_1 \exp\left(-\frac{1}{\alpha} \xi_r^2\right) > \\ \varepsilon(r) &> \frac{f^{(4r)}(h_{4r,1})}{h_{4r}(0)} - c_1 \exp\left(-\frac{1}{\alpha} h_{4r,1}^2\right) \equiv \Gamma_r, \end{aligned} \quad (\text{A.25})$$

where we have applied (A.24). The  $\varepsilon(r)$  value separates the different values of  $\Xi_r$  and  $\Gamma_r$ . Two contrasting situations may appear: either  $\lim_{r \rightarrow +\infty} \varepsilon(r) = \varepsilon_0 \neq 0$  or  $\lim_{r \rightarrow +\infty} \varepsilon(r) = 0$ , which corresponds, as we see, with the case  $0 < \xi_r < h_{4r,1}$  for  $r > r_1$ . Let us check that we already have the desired result. The first case is incompatible with the free choice of  $\varepsilon_1$ , which could be taken as  $\varepsilon_1 < \varepsilon_0$ . Thus, it should accomplish that

$$\lim_{r \rightarrow +\infty} \varepsilon(r) = 0 \quad (\text{A.26})$$

and then

$$\lim_{r \rightarrow +\infty} (\Gamma_r - \Xi_r) = 0 \quad (\text{A.27})$$

with the limits being finite in both expressions. Consequently, we consider Proposition 2 and we have:

$$\begin{aligned} \lim_{r \rightarrow +\infty} \Gamma_r &= \lim_{r \rightarrow +\infty} \left[ \frac{f^{(4r)}(h_{4r,1})}{h_{4r}(0)} - c_1 \exp\left(-\frac{1}{\alpha} h_{4r,1}^2\right) \right] \\ &= \lim_{r \rightarrow +\infty} \left[ \frac{f^{(4r)}(0)}{h_{4r}(0)} - c_1 \right], \end{aligned} \quad (\text{A.28})$$

which implies that  $\lim_{r \rightarrow +\infty} \frac{f^{(4r)}(0)}{h_{4r}(0)}$  is finite, and

$$\begin{aligned} \left[ \lim_{r \rightarrow +\infty} \frac{f^{(4r)}(0)}{h_{4r}(0)} \right] - c_1 &= \lim_{r \rightarrow +\infty} \Xi_r \\ &= \lim_{r \rightarrow +\infty} \left[ \frac{f^{(4r)}(\xi_r)}{h_{4r}(0)} - c_1 \exp\left(-\frac{1}{\alpha} \xi_r^2\right) \right]. \end{aligned} \quad (\text{A.29})$$

The existence of finite limits allow us to rewrite them as

$$\begin{aligned} 0 &\leq \lim_{r \rightarrow +\infty} \frac{f^{(4r)}(\xi_r) - f^{(4r)}(0)}{h_{4r}(0)} \\ &= \lim_{r \rightarrow +\infty} c_1 \left[ \exp\left(-\frac{1}{\alpha} \xi_r^2\right) - 1 \right] \leq 0. \end{aligned} \quad (\text{A.30})$$

The limit on the left is equal or greater to zero because it is finite, the denominator is positive, and  $\xi_r$  is the point where  $f^{(4r)}(x)$  is maximum (and positive). The second member of the equality is equal to or less than zero, because  $\exp\left(-\frac{1}{\alpha} \xi_r^2\right) \leq 1$  y  $c_1 \geq 0$ . In conclusion, only one possibility remains:

$$\lim_{r \rightarrow +\infty} \exp\left(-\frac{1}{\alpha} \xi_r^2\right) = 1 \rightarrow \lim_{r \rightarrow +\infty} \xi_r = 0 \quad (\text{A.31})$$

which we wanted to prove.

Theorem 2: let

$$\begin{aligned} f(x) &= c_1 \exp(-0.5x^2) + \frac{c_2}{\sigma} \exp\left(-0.5\left(\frac{x - \mu}{\sigma}\right)^2\right) \\ c_1 + c_2 &= \frac{1}{\sqrt{2\pi}}, c_1 > 0, c_2 > 0, \end{aligned} \quad (\text{A.32})$$

and let us suppose that  $\sigma < 1$ . If we denote  $\xi_r$  as the point on the  $x$ -axis where the higher relative extreme is in absolute value of the  $2r$ -th derivative of  $f$ , we have that:

$$\lim_{r \rightarrow +\infty} \xi_r = \mu. \quad (\text{A.33})$$

## References

- Arias, E. F., Feissel, M., & Lestrade, J. F. 1988, Annual Report for 1987, Observatoire de Paris, Paris, D-113
- Arias, E. F., Charlot, P., Feissel, M., Lestrade, & Lestrade, J. F. 1995, A&A, 303, 604
- Ayant, Y., & Borg, M. 1971, Fonctions spéciales à l'usage des étudiants en physique (Paris: Dunod)
- Cucker, F., & Smale, S. 2002, Bull. A.M.S., 39, 1
- Fan, J., & Gijbels I. 1992, Ann. Statist. 20, 4, 2008
- Feissel, M., & Mignard, F. 1998, A&A, 331, L33
- Fey, A. L., Ma, C., Arias, E. F., et al. 2004, AJ, 127, 3587
- Fey, A., Gordon, D., & Jacobs, C. 2009, IERS Technical Note 35, The Second realization of the International Celestial Reference Frame by Very Long Baseline Interferometry, Frankfurt am Main 2010: Verlag des Bundesamts für Kartographie und Geodäsie
- IERS 1999, 1998 IERS Annual Report, Observatoire de Paris, D. Gambis (Editor)
- Heiskanen, W. A., & Moritz H. 1967, Physical Geodesy (San Francisco: Freeman)
- Ma, C., Arias, E. F., Eubanks, T. M., et al. 1998, AJ, 116, 516
- Marco, F., Marco, F. J., Martínez, M. J., & López, J. A. 2004a, A&A, 418, 1159
- Marco, F., Marco, F. J., Martínez, M. J., & López, J. A. 2004b, AJ, 127, 549
- Martínez, M. J., Marco, F. J., & López, J. A. 2009, Pub. ASP, 121, 167
- Sfikas et al. 2005, in ICANN 2005, LNCS 3697, eds. W. Duch et al., 835
- Sokolova, Ju., & Malkin, Z. 2007, A&A, 474, 2, 665
- Walter, H. G. 1989a, A&A, 210, 455
- Walter, H. G. 1989b, A&A, 79, 283
- Walter, S., & Sovers, O. J. 2000, Astrometry of fundamental catalogs. The evolution from optical to radio reference frames (Berlin: Springer-Verlag)
- Wand, M. P., & Jones, M. C. 1995, Kernel Smoothing (London: Chapman and Hall)
- Yatskiv, YA S., & Kuryakova. 1990, in Inertial Coordinate System on the Sky, Proc. 141th, eds. J.H. Lieske, & V. K. Abalakin, IAU Symp., 295
- Yu, M., Manchester, R. N., Hobbs, G., et al. 2013, MNRAS, 429, 688