GAIA DR2 - THE FIRST MAIN COURSE

M. Altmann[1,2]

RESUMEN

El catálogo Gaia DR2, publicado el 25 de abril de 2018 con 1.3 mil millones de estrellas, es la primera versión completa del catálogo con los 5 parámetros astrométricos y fotometría medidos. Por tanto, el impacto de este conjunto de datos es revolucionario, aún cuando la publicación de la primera versión limitada de 2016 fue altamente exitosa. En mi presentación, haré una descripción general de los datos dentro de Gaia DR2, sus fortalezas, con algunos casos de ejemplo, y también sus limitaciones, en particular de sus datos astrométricos. Finalmente comentaré sobre las expectativas que se tienen para las próximas versiones de Gaia.

ABSTRACT

Gaia DR2, released on April 25th 2018 is the first full Gaia release with 1.3 billion stars having measured 5-parameter astrometry and 3-band photometry. Therefore the impact of this dataset is revolutionary, following the already highly successful but limited first data release from 2016. In my presentation, I will give an overview of the data within Gaia DR2, its strengths, some exemplary show cases, and also some limitations of this data and astrometric data overall. Finally I will look ahead to the next Gaia releases.

Key Words: astrometry — Galaxy: structure — space vehicles — stars: kinematics and dynamics

1. INTRODUCTION

Gaia, ESA's 1+billion star astrometric mission has meanwhile completed more than 5 years of measurements, and two data releases have been made available to the public since then. The first data release dating from 2016 (Gaia Collaboration et al. 2016), was more like an appetizer, since, though containing positions for more than 1 billion objects, only 2 million of those, appearing also in the Hipparcos and Tycho catalogues, had the full astrometry available. Nonetheless this was a dataset of unprecedented astrometric quality and quantity. For the remaining objects, hybrid catalogues such as HSOY (Altmann et al. 2017) were assembled, using the Gaia DR1 data as second epoch, and other ground-based sources as first epoch, yielding somewhat improved proper motions[3]. These were significant improvements to the previous completely ground based work, however no match for full Gaia Astrometry. Gaia DR2, released on April 25, 2018, by contrast was the first "full" release, with the vast majority of stars having proper motions and parallaxes in addition to positions, as well as colours, and a limited amount of spectroscopy. For a general description of the Gaia

project and Gaia DR1 we refer to Altmann (2015) and Altmann & Bouquillon (2018)[4], which in a way form Part 1 and 2 of a loose series of papers based on Gaia presentations held at ADeLA conferences. This article is intended to guide the reader through the current Gaia DR2, highlighting its properties, contents, but also giving some cautionary notes about caveats, retrieval and usage, closing with an outlook on the EDR3/DR3, planned for late 2020 and late 2021[5], as well as later releases.

2. GAIA DR2

Gaia's second catalogue was officially released on April 25, 2018 (Gaia Collaboration et al. 2018). It is a major step up from Gaia DR1, and in principle the first "real" Gaia dataset, featuring a number of 1.3 billion objects with full five parameter astrometry (positions, proper motions and parallaxes), which is close to the final mission goal. Even with its limitations, Gaia DR1 made a gigantic impact on kinematic astronomy; also hybrid catalogues, such as HSOY (Altmann et al. 2017)[6] have at least alleviated the lack of proper motions a bit, enabling

---

[1]Astronomisches Recheninstitut (Zentrum für Astronomie) der Universität Heidelberg, Mönchhofstr. 12-14, 69120 Heidelberg, Germany.

[2]SYRTE, Observatoire de Paris, PSL Research University, CNRS, Sorbonne Universits, UPMC Univ. Paris 06, LNE, 61 avenue de l'Observatoire, 75014 Paris, France.

[3]However no parallaxes.

[4]And the references therein.

[5]Postponed due to the impact of the CoViD19 pandemic.

[6]In this context, a hybrid catalogue is a proper motion catalogue using Gaia DR1 data as one (usually the 2nd) epoch, and other bround based catalogues, such as PPMXL (Roeser et al. 2010) as the other epoch. For the positions, proper weighting results in virtually only the Gaia data being relevant, but the proper motions have to rely equally on both epochs; therefore these hybrid catalogues are noticeably better
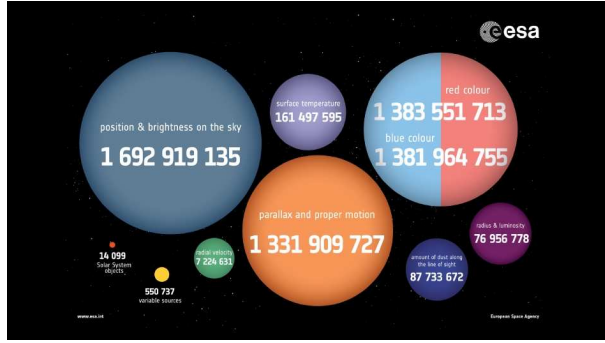
Fig. 1. The key numbers of Gaia DR2, i.e. the number of objects in total, of objects with 3-band photometry, full astrometry, those with radial velocities and various types of astrophysical quantities, and special types of objects, such as small solar system bodies or variable objects. Courtesy: ESA

studies mostly relying on relative rather than absolute proper motions.

Gaia DR2 completely changed the picture. Not only does it contain full high precision astrometry for more than 1.3 billion of a total of 1.7 billion stars (see Figure 1), i.e. $\sim$ 650 times more than TGAS, but also for the first time, three different magnitudes, and for smaller subsets, even radial velocities and astrophysical parameters, e.g. $T_{\rm eff}$ and $\log g$, metallicities, etc. The magnitudes in question are, besides the G-magnitude, already present in Gaia DR1, the $G_{\rm BP}$, a broad band blue magnitude, and $G_{\rm RP}$, a broad band red magnitude. The precision of the astrometry of Gaia DR2 is, depending on magnitude, between 30 and $\sim$ 1000 μas, and the photometric standard error in the region of mmag for most objects, even at G=20 mag, it is just over 0.01 mag. For more details about the performance of Gaia DR2 data, and some verification studies, please refer to Gaia Collaboration et al. (2018), and to the publications in https://www.cosmos.esa.int/web/gaia/\\dr2-papers or from the A&A special issue https://www.aanda.org/component/toc/\\?task=topic\&id=922.

Since the release of the dataset, there have been 3524 publications[7] which are at least partly based on Gaia DR2 data, for updates see https://ui.adsabs.harvard.edu/public-\\libraries/TMaUkmOYRYWxkdJxl619sw. Those

publications, and also those of the verification papers, which appeared at the same time as the release of Gaia DR2 (see above), contain many good examples which show the great potential of this catalogue. Here, in the limited space of this conference paper, I would like to focus on two very simple and yet instructive examples.

Firstly, in order to demonstrate the gigantic step in the quality of Gaia astrometry, in contrast to what we had before, Figure 2 shows vector point plots of the field surrounding the globular cluster M 4 (NGC 6121). This object is especially suited for such demonstration purposes, since it has a very large proper motion; thus field and object motions are well separated. The left panel shows, what is achievable with a very good purely ground-based catalogue[8] in this case the PPMXL (Roeser et al. 2010). The proper motions of the member stars can be distinguished in the lower left side of the diagram, semidetached from the field star proper motion distribution. The middle panel shows the same using HSOY data as an example for a hybrid catalogue, and an improvement is seen. While the distribution of the field stars stays more or less similar, the M 4 member motion distribution is much more concentrated and now fully detached from the field. However, when compared to Gaia DR2 (the right panel), the step up in astrometric quality is extreme - even with a much different colour scaling - the distribution of cluster member proper motions is extremely tight and the most prominent feature in the diagram. That the field distribution seems to be subdued when compared to the two other panels, is due to the (necessary) change of colour scaling. Only with Gaia DR2 it becomes clear, how small the dispersion of velocities of members of a globular cluster is, the previous catalogues only give a hint that it is tighter than that of the field. The field stars have more or less a similar dispersion in all three panels, due to their much higher intrinsic dispersion, given that these stars are in general unrelated to each other.

The second example is the rotation curve of the LMC. Soon after the release of Gaia DR1, van der Marel & Sahlmann (2016) presented their results for the astrometric rotation curve of the LMC, based on 19 TGAS stars, see Figure 3, left panel. This is indeed an impressive achievement. As part of the Gaia DR2 verification suite of papers published in a spe-

---

than ground-based catalogues, but still significantly inferior to the full Gaia astrometry, see Altmann et al. (2017). In fact, the TGAS part of Gaia DR1 (Michalik et al. 2015) is itself a kind of hybrid catalogue, since it relies on Hipparcos and Tycho2 (ESA 1997) as first epoch.

[7]As of April 14, 2020.

[8]For such well known objects, as M 4, dedicated photographic plates exist spanning a baseline of more than a century, allowing better small field astrometry, than with all sky catalogues, like Roeser et al. (2010) (mostly relative astrometry only) - here we are using this object as a show case only.
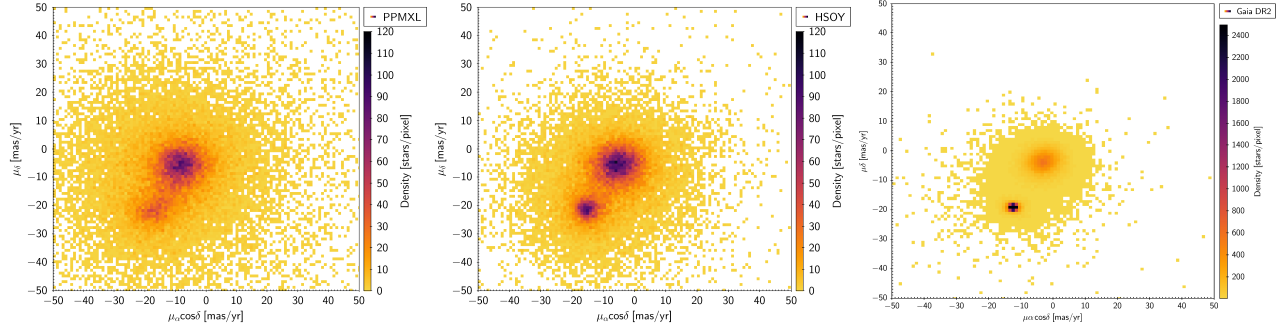
Fig. 2. Vector point plots of the proper motions within the field of M 4 (NGC 6121). The left panel shows purely ground based proper motions (PPMXL), the middle panel those of the hybrid HSOY catalogue, and the right panel those of Gaia DR2. The proper motions of the cluster members are located to the lower left of the centre of each diagram, those of the field stars are broadly centered in the centre of each panel. Please note the change in the colour scale in the right panel. The left and middle panel are taken from Altmann et al. (2017).
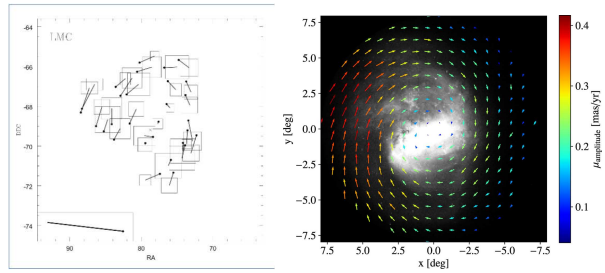

Fig. 3. The rotation curve of the LMC, as derived from Gaia DR1/TGAS data (left panel) by van der Marel & Sahlmann (2016) (adapted from their Fig. 1), and Gaia DR2 data, taken from Gaia Collaboration et al. (2018), one of the Gaia DR2 verification papers (their Figure 24). Note the difference in spatial orientation and thus direction of rotation between the two plots.

cial volume of A&A, Gaia Collaboration et al. (2018) have derived the proper motions of several globular clusters and dwarf spheroidal galaxies in the Galactic neighbourhood, as well as the proper motion and rotation of both Magellanic clouds. For this, they used all suitable member stars in observed, in the case of the LMC about 8 million, leading to a very complete transversal rotation field (the radial component is missing for the lack of radial velocities).

These two simple examples give an impression of the high quality (and quantity) of the Gaia DR2 data, and the enormous step ahead from Gaia DR1. For more cases, please refer to the verification articles in the special A&A volume, and the plethora of publications by the astronomic community at large.

However, also a few words of caution and warnings about caveats are in place, both in general and in more specific issues for Gaia DR2. No catalogue will ever be perfect, and while Gaia DR2 is the first "real" Gaia data publication, it is an early

one. Therefore the scan pattern is still quite apparent, and the number of visits still very inhomogeneous, even if this has significantly improved in respect to Gaia DR1. With future releases, these inhomogeneities are going to smooth out further, but will never entirely disappear - this is inherent to global surveys obeying a scan-law. Another important, yet often overlooked issue is the matter of source IDs. These will in part change from release to release, since more sources will be included, and a small number of objects may also disappear. Thus, it can not be generally assumed that the source ID for a given object in one release will have the same ID in another. Therefore, it is obligatory to always state the Gaia release in question, when referring to an object by its Gaia ID. Gaia DR2 itself still has a number of systematic effects, which will diminish in the upcoming releases. The parallaxes have a global offset of $-29\,\mu$as, and somewhat smaller large scale ones, the proper motions do also have slight systematic effects of this margin. Additionally there are some small scale effects as well, as can be clearly seen in the LMC region, which has some systematic deviations in a large number of similar parallaxes. All of these effects are described in detail in Lindegren et al. (2018), which covers the astrometric reduction and results and all issues involved of this release.

Another important issue is that of the data retrieval. The complete Gaia database is **huge**, several TB of data. Therefore, with our current technology it is simply not possible for everybody to retrieve the complete catalogue, and download it to the local hard drive. Thus the user of Gaia data (and other large data sources) will have to put some more thought into retrieval options. Fortunately, most astronomical analysis tools, such as Topcat (Taylor 2005) have implemented the required tools, starting

from simple cone searches to more elaborate queries, e.g. TAP (Table Access Protocol) queries. For this, there is also a database query language specially developed for astronomer's needs, called ADQL (Astronomical Data Query Language, Osuna et al. (2008)), featuring special astronomy related commands. Such tools and facilities, while having a steep learning curve, are of immense utility, since they allow to optimise even large data requests, and also enable the user to delegate some data selection and also computations to the data server rather than utilising the user's local computer. It is to be kept in mind that the bottle neck is the network and not so much the computing power at our disposal. While increasingly large datasets have been around for many years, the size of these is going to grow, and thus the way we handle them will be subject to perpetual change.

3. UPCOMING RELEASES

Gaia DR2 was released more than two years ago (as of the time of writing), and the Gaia consortium is working hard to release the third Gaia catalogue, which will be split into two parts. First, an early release (EDR3) which mostly contains everything Gaia DR2 contains, just using more data collected over a longer time, and thus better, both in precision and in accuracy, since the sky coverage of the data is more homogeneous, see discussion in the previous section. About a year later, the full Gaia DR3, including many more types of data for more and more complicated objects, will be released. Until recently the tentative release times for these two parts of Gaia DR3 were in the third quarter of 2020 and second half of 2021 respectively. However, the CoViD19 pandemic has had a toll on the operations of the Gaia consortium, and thus both parts will be delayed, by how much has not been determined yet, since the global situation was at the time of writing in a very volatile state, and things can change quite significantly in short time. For more current information, please refer to `https://www.cosmos.esa.int/web/gaia/release`, where also information on further releases can be found[9]. Some of the new features of the full Gaia DR3 will be object classifications for well behaved objects, including publication of the Bp/Rp and RVS spectra these are based on, better radial velocities of more objects, variable star epoch photometry, solar system objects, and more.

A further release covering the data taken during the nominal mission (July 2014 to July 2019) will become public in a few years (Gaia DR4). Depending on the length of the extended mission[10], a final catalogue, with the most sophisticated data treatment, reduction and analysis, will be released. However this will probably be in the late 2020's. It is to be stressed that the data processing and analysis not only encompasses more data each time, but also a broader scope of data, and more knowledge. For this reason the methods evolve with time, and involve more complicated inter-dependant processes. Therefore the time which needs to be invested will grow from release to release.

Two years after the publication of the first full catalogue from Gaia, the usage of this is in full swing. More is to come, and it is not an overstatement that this most exciting and complicated astrometric space mission will provide valuable data for astronomers for years, or even decades to come.

REFERENCES

Altmann, M. & Bouquillon, S. 2018, RMxAC, 50, 7

Altmann, M., Roeser, S., Demleitner, M., et al. 2017, A&A, 600, L4

Altmann, M. 2015, RMxAC, 46, 17

ESA 1997, ESA Special Publication, 1200

Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, A&A, 595, A1

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2016, A&A, 616, 1

Gaia Collaboration, Helmi, A., van Leeuwen, F., et al. 2018, A&A, 616, A12

Lindegren, L., Hernández, J., Bombrun, A., et al. 2018, A&A, 616, A2

Michalik, D., Lindegren, L., & Hobbs, D. 2015, A&A, 574, A115

Osuna, P., Ortiz, I., Lusted, J., et al. 2008, IVOA Recommendation 30 October 2008

Roeser, S., Demleitner, M., & Schilbach, E. 2010, AJ, 139, 2440

Taylor, M. B. 2005, Astronomical Data Analysis Software and Systems XIV, 29

van der Marel, R. P., & Sahlmann, J. 2016, ApJ, 832, L23

---

[9]As well as summaries of the contents of the previous releases.

[10]At the time of writing, Gaia was in its first 2-year term of extension, with a second 2-year term until end of 2022 being in the approval process.