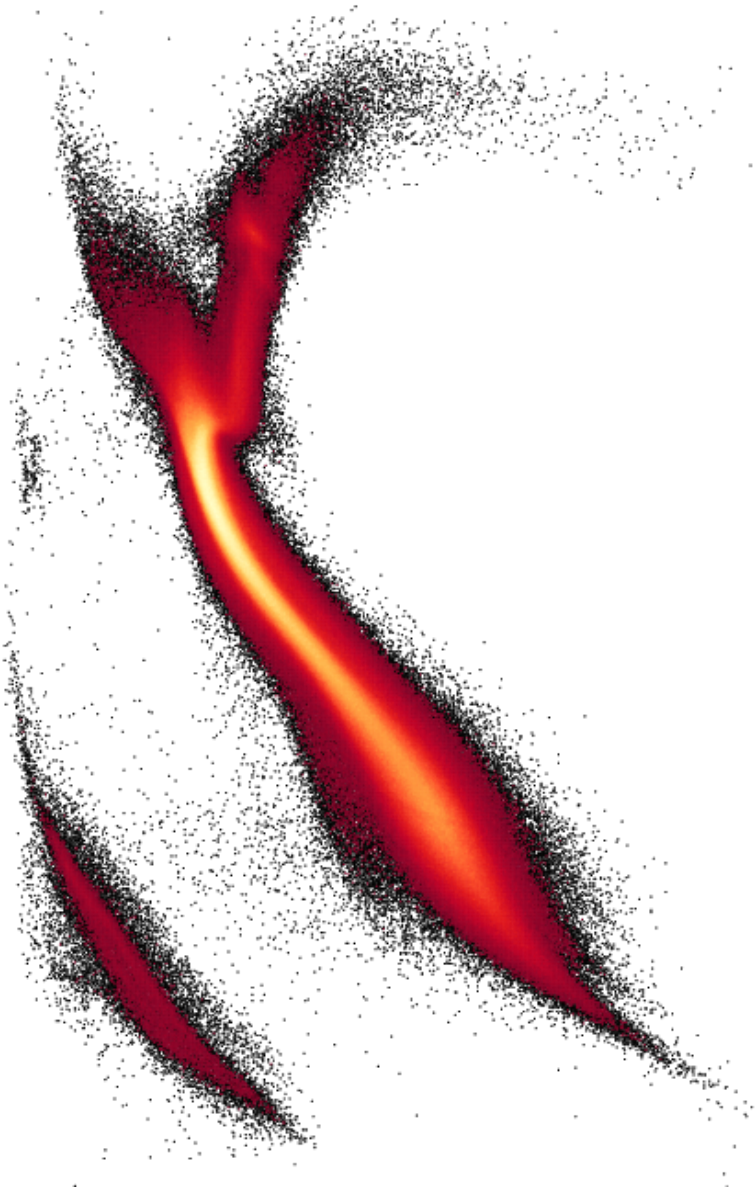
École Evry Schatzman 2023

Stellar physics with Gaia

Ile d'Oléron - October 1-6, 2023

Editors: C. Babusiaux, C. Reylé



Contents

Foreword

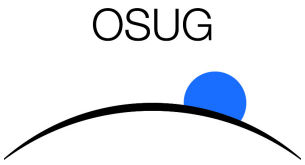
The Gaia mission is revolutionising our knowledge of the stellar content of the Milky Way by providing astrometric and photometric data for around two billion stars, as well as spectroscopic data for the brightest stars. In June 2022, the third catalogue of Gaia data (Gaia DR3) was released. Gaia DR3 provides for the first time spectra (spectrophotometry and spectroscopy), double star characterisations, classifications and astrophysical parameters based on the spectra, as well as new variable star types and more radial velocities. These data are so new and so rich for stellar and galactic physics that the idea of a dedicated Evry Schatzman School was born.

The 2023 edition of the Evry Schatzman School (EES 2023) of the CNRS/INSU French National Program for Stellar Physics (PNPS) was held in the island of Oléron (France) from 1st to 6th October. The aim of the school was to review the Gaia DR3 data, how to use them and how they are improving our knowledge of stellar physics. It aimed at enabling participants to master this complex content for their own scientific exploitation, both for stellar physics and for its applications to galactic studies. The school was structured around five main lectures: the introduction to the Gaia mission and how Gaia answers open questions in stellar physics (Orlagh Creevey), the usage of Gaia data to identify stellar populations (Alejandra Recio-Blanco), multiple systems with Gaia (Frédéric Arenou), Gaia variable stars (Laurent Eyer) and details of the H-R diagram with cold dwarfs (Céline Reylé) and hot stars (Yves Frémat). The afternoons were dedicated to hands-on sessions focusing on the Gaia archive usage, the usage of Gaia spectra (Carine Babusiaux), use case example on the study of star clusters (Orlagh Creevey), multiple systems (Frédéric Arenou) and variable stars (Laurent Eyer). A seminar on the application to galactic dynamics was also provided by Benoit Famaey. This book, the slides of the talks and the tutorials are available on the school website (https://ees2023.sciencesconf.org/). All the lecture notes in this book are also available on Arxiv and ADS.

We warmly thank our colleagues for the quality of the lectures, hands-on and discussions; the participants for their reactions, discussions and enthusiasm; the CAES La Vieille Perrotine; Marie-Hélène Sztefek for the administrative organisation; Nicolas Leclerc for his help during the hands-on session; Nadège Largarde for the organisation of the thesis killer; Nadège Meunier for the pictures that illustrate this book; our sponsors: Centre National de la Recherche Scientifique (CNRS), Programme National de Physique Stellaire (PNPS) and Cosmologie et Galaxies (PNPS), Centre National d'Études Spatiales (CNES), Observatoire des Sciences de l'Univers de Grenoble (OSUG).

The Scientific committe: Carine Babusiaux, Céline Reylé, Orlagh Creevey, Nadège Lagarde, Benoit Famaey, Yveline Lebreton

Sponsors of EES 2023



Participants

Frédéric Arenou (Lecturer), GEPI, Observatoire de Paris
Carine Babusiaux (Chair), IPAG, Université Grenoble Alpes
Marie Barbillon, Lagrange, Observatoire de la Côte d'Azur
Elysabeth Béguin, Lagrange, Observatoire de la Côte d'Azur
Sviatoslav Borisov, Department of astronomy of the University of Geneva
Garance Bras, LESIA, Observatoire de Paris
Valeria Cerqui, GEPI, Observatoire de Paris
Tianxiang Chen, Institut d'Astrophysique de Paris, Sorbonne Université
Orlagh Louise Creevey (SOC, Lecturer), Lagrange, Observatoire de la Côte d'Azur
Morgan Deal, Laboratoire Univers et Particules de Montpellier
Barnabé Deforet, UTINAM, Université de Franche-Comté
Lucie Degott, IAS, Université Paris-Saclay
Elisa Denis, Lagrange, Observatoire de la Côte d'Azur
Judit Donada Oliu, Departament de Física Quàntica i Astrofísica, Universitat de Barcelona
Laurent Eyer (Lecturer), Departement d'astronomie de l'Université de Genève
Benoit Famaey (SOC, Lecturer), Observatoire astronomique de Strasbourg
Antoine Fort, LESIA, Observatoire de Paris
Yves Frémat (Lecturer), Astronomy and Astrophysics Royal observatory of Belgium
Laurane Fréour, Department of astrophysics University of Vienna
Chloe Gutteridge, LESIA, Observatoire de Paris
Nadège Lagarde (SOC), Laboratoire d'Astrophysique de Bordeaux
Nicolas Leclerc, GEPI, Observatoire de Paris
Gautier Lecoutre, UTINAM, Université de Franche-Comté
Louis Manchon, LESIA, Observatoire de Paris
Alejandro Martín Escabia, IPAG, Université Grenoble Alpes
Thibault Merle, IAA, ULB/ROB
Nadège Meunier, IPAG, Université Grenoble Alpes
Eric Michel, LESIA, Observatoire de Paris
Achille Nouzarede, IPAG, Université Grenoble Alpes
Charly Pinçon, IAS, Université Paris-Saclay
Thomas Ravinet, UTINAM, Université de Franche-Comté
Alejandra Recio-Blanco (Lecturer), Lagrange, Observatoire de la Côte d'Azur
Daniel Reese, LESIA, Observatoire de Paris
Celine Reyle (SOC, Lecturer), UTINAM, Université de Franche-Comté
Michel Rieutord, IRAP, Université de Toulouse
Réza Samadi, LESIA, Observatoire de Paris
Caroline Soubiran, Laboratoire d'Astrophysique de Bordeaux
Wenchao Su, IRAP, Université de Toulouse
Mathieu Van der Swaelmen, INAF, Osservatorio Astrofisico di Arcetri

École Evry Schatzman 2023: Stellar physics with Gaia
C. Babusiaux, C. Reylé (eds)

STELLAR PHYSICS ACROSS THE HR DIAGRAM WITH GAIA

O. L. CREEVEY

Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, Bd de l'Observatoire, CS 34229, 06304 Nice cedex 4, France

e-mail: orlagh.creevey@oca.eu

Abstract. Gaia Data Release 3 (GDR3) contains a wealth of information to advance our knowledge of stellar physics. In these lecture notes we introduce the data products from GDR3 that can be exploited by the stellar physics community. Then we visit different regions of the HR diagram, discuss the open scientific questions, and describe how GDR3 can help advance this particular topic. Specific regions include hot OB and A type stars, FGK main sequence, giants, and variable sources, low mass stars, and ultra-cool dwarfs. Examples of scientific exploitation are also provided. These lecture notes are accompanied by a 3-hour lecture presentation and a 3-hour practical session that are publicly available on the website of the Ecole Evry Schatzman 2023: Stellar physics with Gaia, https://ees2023.sciencesconf.org/, see *Lectures* and *Hands-on Work*.

Keywords: Stars: fundamental parameters – Stars: variables – Stars: solar-type – Star: low-mass – Hertzsprung-Russel and C-M diagrams – Stars: emission-line – Galaxy: stellar content – Astronomical data bases – Catalogs – Surveys

1 Introduction to Gaia

The Gaia mission (Gaia Collaboration, Prusti et al. 2016) was launched just over 10 years ago on the 19 December 2013. After its journey to Lagrange point L2 and its commissioning stage, it began science operations in July 2014. Today (January 2024) it has been collecting scientific data for 9 and a half years. Its mission extension has been approved until approximately the end of 2024 to May 2025, when the micropropulsion gas will have reached its lower limit to enable science operations.

Gaia was designed to perform a full sky survey of about 1–2 billion of the brightest objects in visible light. It targets a G-band range spanning approximately 3 – 21 mag, similar to V-band but extending from 325 – 1050 nm (Montegriffo et al. 2023). These brightest targets represent an approximate 1% of the content of our Milky Way. Gaia's instruments enable the collection of three main data products: astrometry, spectrophotometry, and spectroscopy. These together allow the mission to fulfill its scientific objectives of understanding the formation and evolution of our Galaxy by measuring the accurate positions, motions, and intrinsic properties of its observed stars. Through these measurements, Gaia is also a space mission aimed at improving our knowledge of stellar physics. But Gaia's science does not just stop at stars and the Galaxy. As it measures all objects spanning 3 – 21 magnitude, it contributes significantly to the understanding of our own Solar System through the measurements of its asteroids and other minor bodies, to the understanding of sub-stellar objects such as brown dwarfs and exoplanets, and it also measures extra-galactic sources, such as quasars and other galaxies (Gaia Collaboration et al. 2023a,b,c; Delchambre et al. 2023). A series of two-page essays produced by M. Perryman provides an overview of the science legacy.

Gaia rotates about its own axis every 6 hours, and as it is inclined at about 45° relative the direction of the Sun, it also precesses. It also follows the orbit of the Earth around the Sun. This rotation, precession and orbital movements allow Gaia to scan the whole sky. However, not all positions on the sky are visited the same number of times, and there are imprints of the scanning law on the quality of the data. Over the nominal mission lifetime of 5 years, Gaia will have observed each source a median value of 70 times (70 transits), sometimes reaching as low as 5, but also reaching a high several hundred transits. Although the end of the nominal mission has passed

and it is now in its 10th year of observations, the Gaia Data Processing and Analysis Consortium (DPAC) has of today released data based only on 34 months of science operations. This is Gaia Data Release 3 (GDR3) which took place in June 2022 (Gaia Collaboration, Vallenari et al. 2023; Babusiaux et al. 2023). These are the data that are the subject of this school, lectures and practical work, see https://ees2023.sciencesconf.org/.

Gaia's measurements are time series of positions, G-band photometry, low resolution spectro-photometry from blue and red prisms (BP and RP, or together XP), and high resolution spectroscopy (R$\sim$11,000) from the Radial Velocity Spectrometer (RVS). During the processing by Gaia-DPAC, these are converted to (1) an astrometric solution (positions, proper motions, and parallax), derived from the time-series positions (2) mean and time-resolved G, BP, and RP integrated photometry, along with passbands and zeropoints, and (3) mean and time-resolved XP and RVS spectra. Another part of the DPAC processing uses these basic data products to provide astrophysical analysis; (1) analysis of multiple systems, e.g. solutions to two-body systems, (2) analysis of variable objects e.g. periods and amplitudes of pulsating stars, (3) astrophysical parametrization of all sources, e.g. class probabilities, and stellar parameters, (4) solar-system analysis, e.g. asteroid and orbit characterisation, (5) morphology and redshifts of extra-galactic sources, (6) science alerts, and (7) auxiliary products such as Gaia simulations and cross-match tables. In GDR3 all of the above-mentioned products from the astrophysical analysis was published along with 220 million mean XP spectra and 1 million RVS spectra. GDR3 also comprises the astrometric solution and mean photometry, as was published in Gaia Early Data Release 3 in December 2020 (Gaia Collaboration, Brown et al. 2021). All of these data products are made available through the Gaia Archive[1]. Examples of accessing these data are found in the practical sessions, and further information can be found on the Gaia Help pages[2].

In this lecture, I discuss a subset of the Gaia DR3, and focus on the stellar-related data products.

2 Gaia Observables

To fully understand how Gaia can advance our knowledge of stellar physics, it is important to understand what we consider the *basic* data products; the astrometry, spectrophotometry, and spectroscopy. I will introduce each of these in the following sections, while focussing on the products that are available in the Gaia archive, and notably in the table `gaiadr3.gaia_source`.

2.1 Astrometry

The most basic information that we can have about the content of the visible sky is a knowledge of the number of objects and their positions. The accuracy with which we can determine positions depends evidently on the brightness of a star but primarily on the spatial resolution of the instrument used to measure the positions. By eye we can estimate the positions of the brightest targets to roughly a few arcminutes, if we can also see a source with a known reference position. In fact the definition of 20/20 vision is the ability to resolve to one arcsecond. However, at night, conditions are not optimal, and the atmosphere induces aberrations, which increases this limit to a few arcminutes. This also means that if two stars are closer than a few arcminutes, then we are probably not capable of observing, or resolving, the individual sources. The use of binoculars can help go further allowing us to increase the resolving power by a factor typically about 10, and of course modern telescopes increase this number further.

One of the first catalogues of stars to be produced was by Hipparchos of Rhodes in approximately 130 BC. His catalogue comprised about 850 stars with a typical precision on the order of a degree (at that time there was no known reference position). Roughly 1700 years later, Tycho Brahe measured about 1000 stars with a roughly 1 arcminute precision. As we advanced in our knowledge of the sky, measurement techniques, and instruments, the numbers of stars and accuracy of their positions increased significantly to roughly one arcsecond until modern techniques and surveys appeared, see Fig. 1 left panel which illustrates the increasing precision of astrometric measurements over time, along with appearance of reference catalogues (Høg 2024).

The increased precision of positions also allowed the measurements of the relative positions of stars with respect to their *fixed* background. If we observe a star at two distinct moments, the star will appear to move with respect to the fixed background. How much it moves will depend on how far away the star is and the difference in angle from which we measure it. This apparent movement in the sky is called a parallax, and it

[1] https://gea.esac.esa.int/archive/
[2] https://www.cosmos.esa.int/web/gaia-users/archive

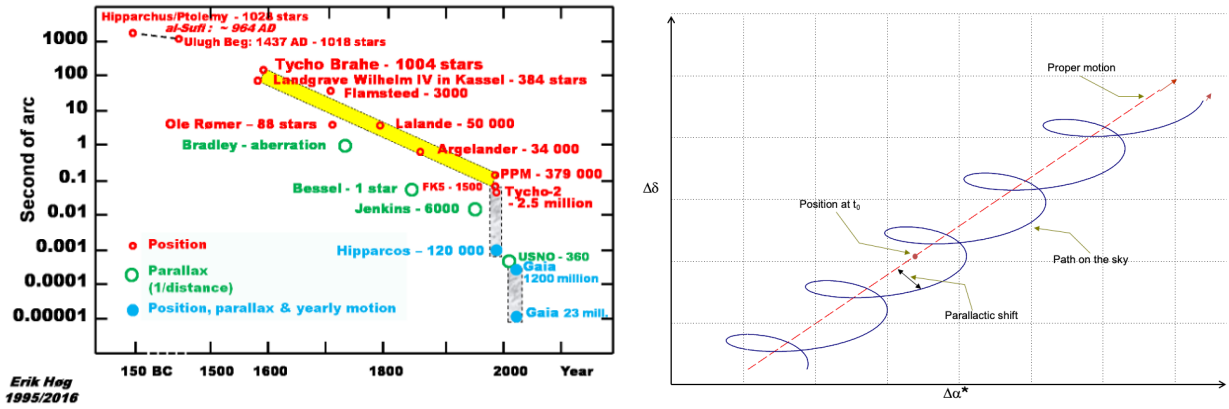


Fig. 1. *Left:* The precision in star positions in the sky for catalogues spanning the last 2000 years. Credit: Høg, E. 2024, arXiv:2402.10996. *Right:* The apparent motion of a star on the sky as we observe it from different positions and over time (blue). The red dashed line shows the proper motion of the star, and the black arrow indicates the parallactic shift as the observer's position moves around the Sun. Approximately five years of measurements is illustrated. Acknowledgement: Figure kindly provided by François Mignard.

allows one to estimate the star's distance by using trigonometry if we know the distance between the two points when the measurements were made. The use of the Earth's orbital motion around the Sun allows us to do just this, by measuring the position of the star relative to the background stars at a separation of 6 months when the Earth is exactly 2 astronomical units (AU) away from its first position. A stellar parallax was first measured by Bessel for 61 Cygni in 1838. He deduced a parallax of 314 ± 20 milli arcseconds (mas) which provided the first distance estimate of roughly 10 light years. Knowing the distance to objects allows one to estimate their true luminosities and radii, two of the fundamental parameters of stars which are a product of stellar structure and evolution.

The first catalogue to allow mas precision on the parallaxes of stars was the Hipparcos mission (ESA 1997). It measured the parallaxes of 120 000 stars up to a visible magnitude of roughly 10. This catalogue remained the reference catalogue until Gaia Data Release 2 in 2018 for most of these stars, except for the brightest where Gaia needs its full nominal mission data to better understand the systematics at these low magnitudes.

2.1.1 An astrometric solution

Given a set of position measurements of a star at specific time points, one can trace the apparent motion of a star with reference to fixed points in the background, see Fig. 1 right panel. For a single star, the parameters that are solved are (1) the mean position in right ascension (ra) and declination (dec) at a reference time denoted t_0, usually taken to be the central time of the span of the measurements, (2) the mean motion of the star on the projected sky background, referred to the proper motion in ra* and dec, and (3) the parallax, which is the radius of the circle that is traced per year of observation. These five parameters provide an astrometric solution for a single star. Due to telescope optics, the colour of the star will also play a role in the measured positions of the star, so that one must also consider this parameter. In Gaia DR3 the `pseudocolor` or `nueff_measured_in_astrometry` are these parameters. In the first case the colour is fitted (6-parameter solution) while in the second it is fixed (5-parameter solution). The number and quality of the measurements covering the apparent motion on the sky over the 34 months determines the uncertainties and correlations in the solved parameters. The purposes of these lecture notes is not to go in detail on how Gaia produces astrometric solutions, readers are therefore referred to e.g. Lindegren et al. (2021b), and Brown (2021) provides an annual review while describing the astrometric problem, the solution and some science highlights using Gaia data. Here we just focus on understanding the contents of the archive through a qualitative understanding of the data.

2.1.2 Archive parameters

The astrometric solution is what gives rise to each source identification. The main identifier in most of the archive tables is the `source_id` which is a set of between 16 and 20 numbers used to encode information concerning the position of the source in the sky, the processing center used to produce the solution, and a

running ID. For example, if one would like to obtain the level 12 HEALPix identification, one should divide the source ID by 34359738368. It is important to note that with each processing cycle, that is with each main Gaia Data Release, the source ID can change. If one would like to crossmatch the source ID with another catalogue, one should use the designation 'Gaia DR3 `source_id`', e.g. 'Gaia DR3 1872046574983497216'.

Each source ID has an associated ra and dec, along with proper motions and a parallax if the solution has at least five parameters, i.e. `astrometric_parameters_solved` = 31 or 95. If this number is 3, then it indicates that the astrometric solution has only fitted two parameters (ra, dec). Along with the main parameters, there are also several fields published pertaining to their uncertainties and correlations, e.g. `ra_parallax_corr`. Other parameters related to the astrometric solution concern the model parameters, e.g. `astrometric_parameters_solved` or `nu_eff_used_in_astrometry`. Additional information is made available concerning the quality of the model fit such as a chi-square, unit weight error. Examples of these parameters are `astrometric_gof_al`, `astrometric_chi_al`, or `astrometric_excess_noise`, where the `al` refers to the *along-scan* direction of the analysed field, and `ac` is perpendicular to this, otherwise known as *across-scan*. The latter parameter indicates how close the model is to the data, assuming a single star solution. A large value could therefore indicate a non-single source, or issues with convergence. Of particular interest is the `ruwe` or *re-normalised unit weight error* defined as $ruwe = \frac{\sqrt{chi/(goodobs-m)}}{f(G,G_{\rm BP}-G_{\rm RP})}$ where *chi* is the astrometric chi-square, *goodobs* is `astrometric_n_good_obs_al` and m is the effective number of model parameters. It is similar to a reduced chi-squared, but modified by a function (or renormalised) that depends on G and the source's colour. A value of 1 is an ideal value for a single star solution, and a value larger than e.g. 1.5 could indicate a binary. See the official online documentation for further details.

Additional information from the astrometric processing is also made available, such as the number of observations used in the solution, the number of peaks that are found in the images, or if the source is a `duplicated_source`. This latter could indicate cross-match or processing issues, or duplicity of a source, knowing that Gaia catalogue has a spatial resolution of 0.18 arcsec in GDR3 (compare to 0.4 arcsec in GDR2).

2.1.3 Astrometric biases

As detailed in Lindegren et al. (2021a), biases are present in the parallaxes of the astrometric solution. In particular for stars, these biases have been shown to be a function of magnitude, colour and position in the sky, see e.g. Fig. 19 from the above reference. A parallax zeropoint bias correction, $\pi_{\rm cor}$, has been proposed by the authors, to be used as judged relevant by a user of the archive. This bias takes the form of a correction where $\pi_{\rm cor} = \pi_{\rm GDR3} - Z_N$ where N is 5 or 6, referencing the number of parameters solved, and Z_N is a function of the three above mentioned parameters. A python function has been made available to the public by one of the authors, called `gaiadr3_zeropoint`, which returns the function Z_N. The code can be found on `https://gitlab.com/icc-ub/public/gaiadr3_zeropoint`.

2.2 Spectrophotometry

Two prism spectra are on-board Gaia which together provide low resolution spectra (resolution between 30 – 90, De Angeli et al. 2023) spanning 330 nm – 680 nm (BP) and 620 nm – 1050 nm (RP), see Fig. 2. The processing of these prism spectra provide two main data products (both mean and time series): (1) low resolution spectra, and (2) integrated photometry, $G_{\rm BP}$ and $G_{\rm RP}$. The G band photometry is also produced by the same processing unit by exploiting the astrometric CCDs. Details of the data reduction, processing and calibrations are given in Riello et al. (2021) for the photometry, De Angeli et al. (2023) for the spectral processing, and Montegriffo et al. (2023) for the external calibration of the spectra.

2.2.1 Use in DPAC

All subsystems within DPAC use a subset of these data products. Analysis of source variability (pulsating stars, eclipsing binaries, ...) is performed with the time series products, while source characterisation (source type, effective temperature, mass, ...) is performed with the mean data products. These spectra are produced from the individual time series and use a Hermite polynomial representation. Fig. 3 left panel illustrates a set of RP epoch spectra for an individual source, represented by the coloured dots, where the colour is indicative of its time stamp. The mean spectrum is overlaid, giving an accurate representation of this source's RP spectrum. The x-axis is pseudo-wavelength which can be converted into a physical wavelength by using a dispersion law. The shape of this spectrum is due to the instrument. These internally calibrated spectra can be converted to

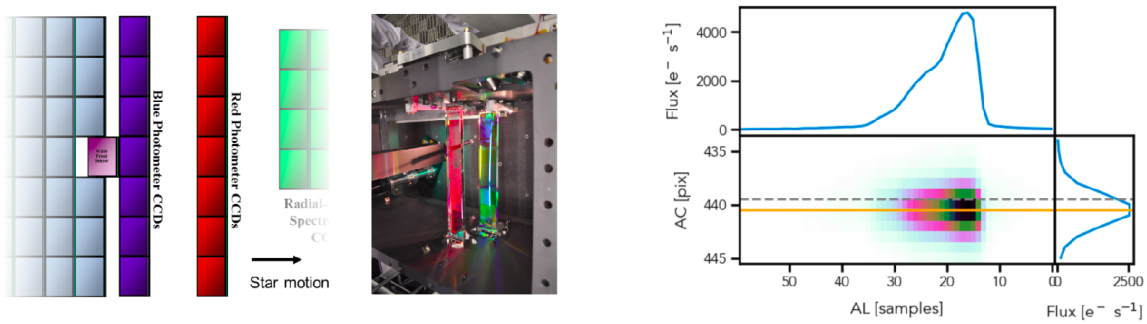


Fig. 2. Composite image provided by F. De Angeli illustrating the BP and RP prisms on board Gaia (center), the position of the prism CCDs in the focal plane (left) and an example of the BP/RP images analysed by the DPAC processing (right). Credit: De Angeli et al. 2023, A&A, 674, A2.

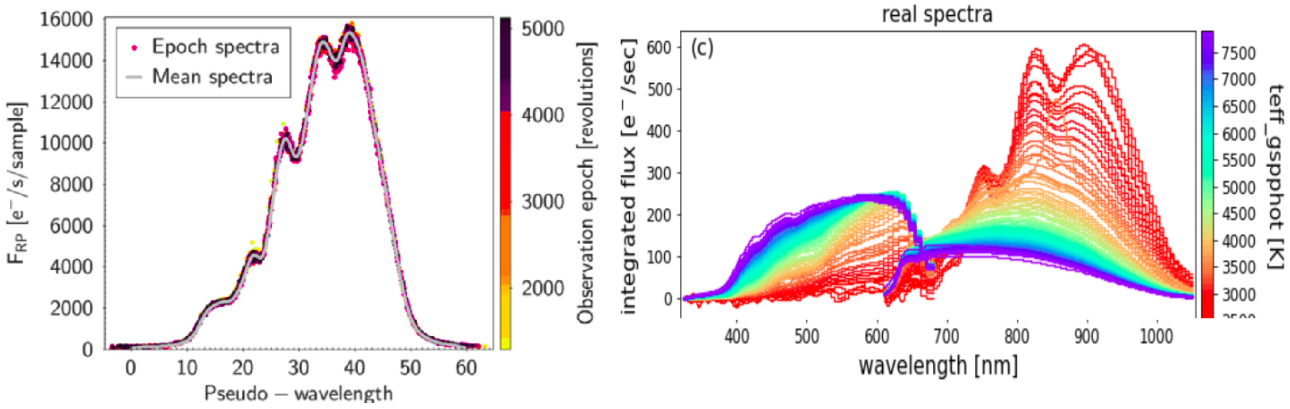


Fig. 3. *Left:* RP epoch spectra for one source with the mean spectrum overlaid. The colour-code indicates at which time the data were measured. The y-axis shows pseudo-wavelength which can be converted to wavelength using a dispersion law. Credit: De Angeli et al. 2023, A&A, 674, A2. *Right:* Examples of sampled mean spectra of various sources with low extinction for different $T_{\rm eff}$. Credit: Creevey et al. 2023, A&A, 674, A26.

a true spectral energy distribution having knowledge of the instrument model. The internally calibrated mean spectra are used for producing astrophysical parameters and for the analysis of variable sources. An example is given on the right panel of Fig. 3, where the BP and RP spectra of stars with low extinction and different effective temperatures $T_{\rm eff}$ are shown.

2.2.2 Archive parameters

220 million internally calibrated mean-spectra have been made available in GDR3. These are mostly for sources with $G < 17.6$ but many subsets of fainter sources have also been published. 100 million sampled mean spectra have also been published. These data products are found through the Gaia archive datalink. A python tool `GaiaXPy` to exploit these data has been made available at `https://gaia-dpci.github.io/GaiaXPy-website/`, and an example of its use is given in the practical course. In `gaiadr3.gaia_source` there are several fields related to the BP/RP processing, such as the number of photometric observations used for the processing of each prism e.g. `phot_rp_n_obs`, and the number of contaminated or blended transits e.g. `phot_bp_n_blended_transits`, and one may want to use these values to judge the quality of the spectra or an indication if the source may be contaminated by a nearby star.

Concerning the integrated photometry, derived colours, such as `bp_rp`, and several statistical parameters, such as the mean photometry (e.g. `phot_bp_mean_flux`) and error are made available in the catalogue. Passbands and zeropoints are also associated with the photometry and these are available at `https://www.cosmos.esa.int/web/gaia/edr3-passbands`. Another interesting parameter is the excess factor

which gives the ratio of the sum of the BP and RP integrated flux compared to the G integrated flux. A value deviating from the norm could indicate inconsistencies of the fluxes, in particular due to crowding issues, see Riello et al. (2021) their Sect. 9.4 for details. In GDR3 the precision in the photometry is 0.25 mmag at $G < 13$, 1 mmag at $G < 17$ and 5 mmag at $G < 20$. These values increase by a factor of 5 to 20 for $G_{\rm BP}$ and $G_{\rm RP}$.

Time series photometry for variable objects has also been published in GDR3, see lectures and practical work in this school by Eyer (2024). These sources are listed in the `gaiadr3.vari_summary` table, and G, $G_{\rm BP}$ and $G_{\rm RP}$ time series for these sources are available through datalink. In addition to the mean and time series photometry, a G-band bolometric correction tool is available under Gaia DR3 software tools and can be accessed directly here: `https://www.cosmos.esa.int/web/gaia/dr3-bolometric-correction-tool`, produced by the group providing astrophysical parameters, see Creevey et al. (2023).

2.3 Spectroscopy

A third instrument on-board Gaia is the Radial Velocity Spectrometer, RVS, a spectrograph with a resolution on the order of 11 000 spanning the 845 nm – 872 nm range and observing sources with $G < 16.2$. Fig. 4 left panel illustrates an example of a CCD image of a spectrum (top), along with the final data product that is produced after the spectral processing (middle), see Sartoretti et al. (2018) for precise details on this processing. As one can see, several prominant spectral lines can be seen, notably the Ca IR triplet at 849.8 nm, 854.2 nm and 866.2 nm. However, this is a very high signal-to-noise-ratio (SNR) spectrum, and many of the RVS spectra have much lower SNR, as illustrated in the bottom left panel. On the right panel, a spectral library is shown illustrating qualitatively the prominant features available for stars of different spectral types, or $T_{\rm eff}$. One can appreciate the difficulty of deriving physical quantities for a hot star of spectral type B or O given the lack of features available in the spectrum.

The RVS has been designed primarily to measure the radial velocity (RV) of each source, which when complemented with the astrometric solution, gives the full position and velocity motion of a star. For stars with $G < 12$ the RVs are measured on an epoch basis and the mean RV is derived, while for fainter targets the RVS epoch spectra are first combined and the RV is subsequently measured. The precise measurement of the RVs need *a priori* a precise characterisation of the star's atmospheric parameters, $T_{\rm eff}$, iron abundance, [Fe/H], and surface gravity $\log g$. It also requires knowledge of spectral broadening and the instrument's line spread function, which are determined during their processing. Therefore, an error induced in any of these parameters, could result in an erronous RV. Discussion and validation of the RVs are described in Katz et al. (2023). In the case of binarity, RVs of single (SB1) or double (SB2) lines can be measured. These time series of RVs, along with the time series of positions and photometry, are then treated by the non-single star subsystem to provide orbital parameters and, when possible, individual or minimal masses of one or both components. Lecture notes and practical work on non-single stars is discussed further in this school in Arenou (2024), see also the hands-on work with spectra.

2.3.1 Use in DPAC

The RVS spectra are used primarily to derive mean RVs of single stars (Katz et al. 2023) and time series of multiple systems and classical pulsating stars such as RR Lyrae and Cepheids (Clementini et al. 2023). The latter data products are then interpreted by the non-single star subsystem, as mentioned above, and the variability subsystem to produce pulsation properties. The mean RVS spectra are also used for astrophysical characterisation to produce atmospheric properties of sources (Creevey et al. 2023; Fouesneau et al. 2023; Recio-Blanco et al. 2023; Frémat 2024), and when possible, chemical abundances, see the lecture notes in this school by Recio-Blanco (2024).

2.3.2 Archive parameters

In `gaiadr3.gaia_source` a set of parameters produced by the RVS processing system are published, along with a sample of approximately 1 million RVS spectra, with SNR of between 5 and 300. The parameters comprise 34 million RVs and errors for $G_{\rm RVS} < 16$, where $G_{\rm RVS}$ comes directly from the flux in the spectra, see Sartoretti et al. (2023), and is also available as a published parameter called `grvs_mag`. Processing and model-fitting information is also provided, such as the number of observations used `rv_nb_transits`, number of visibility periods (a group of measurements taken in a short time span), or the method applied to process the RVs `rv_method_used`. Model-fitting information examples include the χ^2 p-value, goodness of fit, and template

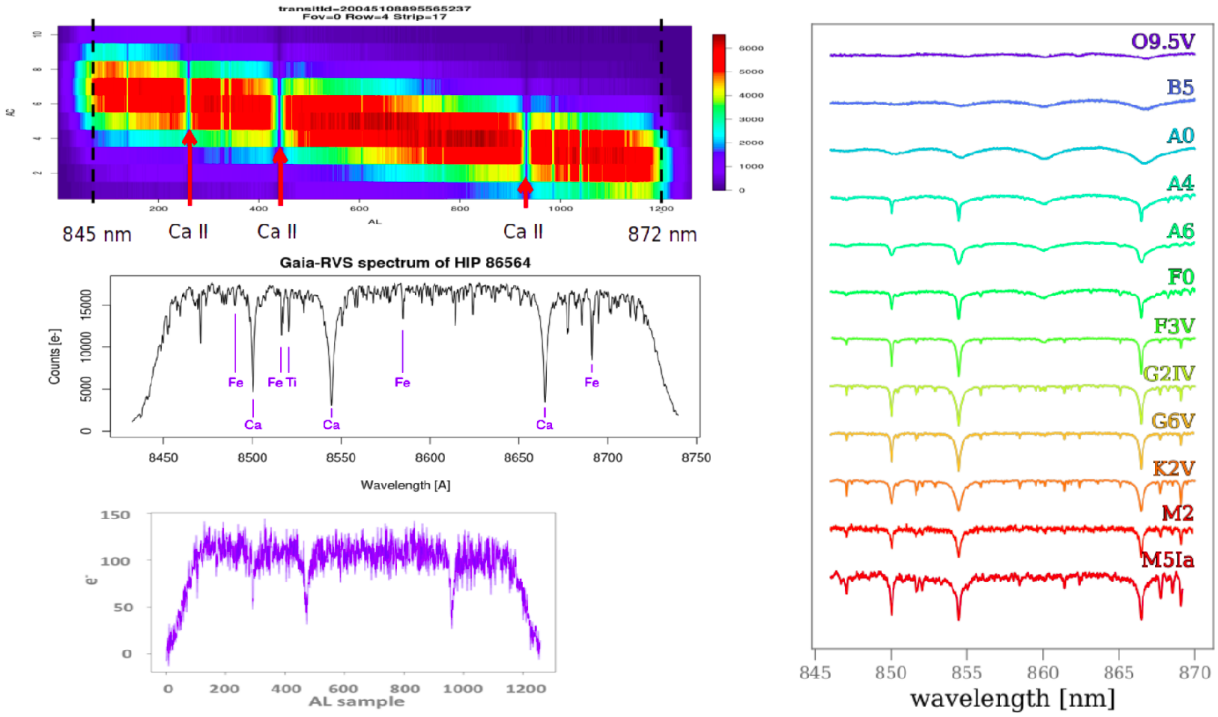


Fig. 4. *Left:* Top panel shows the RVS pixel spectra spanning 845 nm to 872 nm with the three Ca lines (the Calcium triplet) indicated by the red arrows. The middle panel shows an example of a 1D RVS spectrum derived from the processing for a high signal-to-noise spectrum, while the lower panel illustrates a noisier spectrum. Credits: Top panel: Sartoretti et al. 2018, A&A, 616, A6. Middle panel: ESA/Gaia/DPAC/Observatoire de Paris-Meudon/Olivier Marchal & David Katz, Lower panel: Figure produced by O. Marchal. *Right:* A spectral library showing RVS spectra for stars of spectral types 0 to M. Credit: Fouesneau et al. 2023, A&A, 674, A28.

atmospheric parameters used, e.g. `rv_chisq_pvalue`, `rv_template_teff`. In addition to these parameters an estimate of the broadening induced on the spectral lines due to many physical processes in the atmosphere such as rotation, convection, and winds, is also provided and summarised as the `vbroad` parameter. Frémat et al. (2023) shows that the broadening parameter is often a good estimate of the $v \sin i$, but how well it is a good representation depends on the $T_{\rm eff}$, and how accurately the $T_{\rm eff}$ can be measured, see Figs. 2 and 11 of that paper.

All parameters related to multiplicity, variability, or the astrophysical parameters are found in dedicated archive tables and will be discussed in more detail below and subsequent lectures.

2.4 Accessing multi-dimensional products

Multi-dimensional products are data products that can not be represented by a single value in an archive table, such as the RVS and XP spectra. Other products include Markov Chain Monte Carlo (MCMC) samples, see Sect. 3.5. These can be accessed after a query on the archive by clicking on the `datalink` button of the query results. They can also be accessed programmatically through scripts e.g. by using python. Concrete examples are provided in the practical session.

The single-object tab of the archive provides detailed information on individual sources, and allows a unique view of the object without having to download all of the products. Astrometric, photometric, radial velocity, and astrophysical parameters (as given in `gaiadr3.gaia_source`) are highlighted in a human-readable table form. The position of the star on the sky is shown in a skymap using ESASky. A visualisation of all available datalink and other time series data for this source are also shown. For users interested in knowing more about a particular source, this single-object tab is extremely useful.

3 Stellar physics across the HR diagram

Given the primary data products all described in Sect. 2, a wealth of astrophysical information on stars and other sources can be derived and used to improve our knowledge of stellar physics. Dedicated processing modules in DPAC provide a first homogenous analysis of the primary data products and these are made available to the community. Such examples are periods and amplitudes of pulsating stars found in the `gaiadr3.vari_*` tables, stellar parameters from the BP/RP and RVS spectra found in the `gaiadr3.astrophysical_parameters*` tables, or orbital information on multiple star systems found in the `gaiadr3.nss_*` tables, where '*' indicates the existance of multiple tables with the same root name. The rest of the lecture is dedicated to explaining this content of Gaia DR3 in the context of stellar physics. This will be done by addressing different regions of the HR diagram, where I will indicate the open questions in stellar physics and the available parameters in GDR3 that can be used to address these questions. I will also give a few examples of scientific exploitation of GDR3 in specific stellar regimes. Complementary dedicated lecture notes, presentations and practical work are provided in this school by Recio-Blanco (2024); Réyle (2024); Arenou (2024); Eyer (2024); Frémat (2024), which enter into more detail on specific topics. Before embarking on the visit through the HR diagram, I will begin with a brief introduction to the Astrophysical Parameters Inference System (Apsis, see Creevey et al. 2023) which is the subsystem responsible for the characterisation of a half billion stars ($T_{\rm eff}$, $\log g$, radius, extinction, ...).

3.1 Introduction

Astrophysical parameters (APs) in GDR3 are produced by the Apsis analysis system in DPAC (Creevey et al. 2023). Apsis exploits the astrometry, *mean* BP/RP and RVS spectra, and *mean* photometry along with stellar models and/or training data to derive a range of AP products. These products include probabilities and classifications (star, galaxy, spectral type,...), outlier object analysis through the use of self-organising maps, QSO and unresolved galaxy redshifts, interstellar extinction parameters such as A_0 (extintion parameter defined at a reference value of 547.7 nm) and a 2D total galactic extinction map, and stellar parametrization. Among the stellar parameters, Apsis estimates 470 million spectroscopic parameters from the BP/RP spectra and 6 million from the RVS spectra (e.g. $T_{\rm eff}$, [Fe/H]), 220 million evolutionary parameters (e.g. radius, age) and H-α equivalent widths, 2 million activity indices and about 1 million sources with many chemical abundances. These are produced by specific modules within the Apsis system that are designed to target a specific region of the HR diagram or use distinct data sets. Fig. 5 summarises the names of the Apsis modules on the left, the input data and dependencies for each module in the middle, while the figure on the right highlights the different regions of the HR diagram that are analysed by the stellar-based modules. In GDR3 the parameters from the stellar modules are found in the `gaiadr3.astrophysical_parameters` and `gaiadr3.astrophysical_parameters_supp` tables, with a very small subset copied to `gaiadr3.gaia_source`. The field names have their modules encoded in them, so the user can trace the methods and models used to derive the parameter, see Chapter 11 of the Gaia DR3 online documentation.

There are certain regions of the HR diagram that are discussed in detail in the lecture notes by Recio-Blanco (2024); Réyle (2024); Arenou (2024); Eyer (2024); Frémat (2024). These are pulsating stars and binaries which exploit the time series data, the hottest and coolest stars in the HR diagram, and chemical abundances. These will be mentioned in these lecture notes, but the reader will be referred to those lectures for details. Fig. 6 illustrates a HR diagram with the GSP-Phot parameters from Apsis, with the colour-code indicating relative number density. We have split the diagram into several regions and the lecture notes that follow will address each of these regions individually.

The Apsis module GSP-Phot (Andrae et al. 2023) analyses the mean BP/RP spectra of each source using four libraries of stellar atmospheres with a $T_{\rm eff}$ range between 2 500 and 50 000 K and fits the atmospheric parameters $T_{\rm eff}$, $\log g$, [Fe/H], and A_0 and distance assuming a single star. The assumptions laid out here explain the following few attributes in Fig. 6: (1) there are no white dwarfs because these models are not included, and therefore are assigned incorrect parameters, resulting in poor posterior distribution functions (pdf); (2) there are no dwarfs below 2 500 K, although we have highlighted the ultra-cool dwarfs because these are analysed by a different stellar module and these will be discussed further; (3) the binary sequence is not well represented, although one could use the diagram to estimate which sources are binaries; (4) some patchiness is seen and this is partially due to parts of the diagram having a relatively low number of stars, but also due to the boundaries in the model grids mostly seen at 6 000 K and 15 000 K. Concerning the white dwarfs, a user can exploit directly the photometry and astrometry to easily identify those sources, see e.g. Figs. 1 or 12 from Gaia Collaboration, Babusiaux et al. (2018).

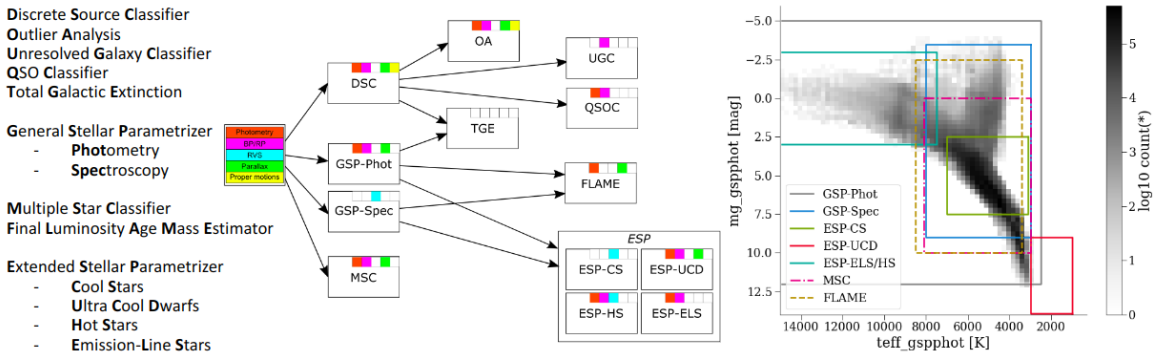


Fig. 5. The Apsis system comprises several modules that derive stellar and non-stellar parametrization and classification. Each module is listed on the left of the figure, and it is represented by a box in the functional diagram in the middle panel. The box also indicates which data are used for the processing and the arrows show the dependencies among modules. On the right panel a HR diagram is shown using the stellar parameters from Apsis. The different coloured boxes indicate the region of the diagram which are analysed by the different stellar-based modules. Credit: Creevey et al. 2023, A&A, 674, A26.

To complete the information about GSP-Phot, as it will be mentioned several times in the following sections, all sources are analysed using four libraries of stellar atmospheres. These are OB, A-star, PHOENIX, and MARCS. The results for all sources for all libraries are found in the `gaiadr3.astrophysical_parameters_supp` table. However, GSP-Phot also chooses the most appropriate library for each source based on convergence criteria. This best library for each source is given in the `gaiadr3.astrophysical_parameters` table and is indicated in the `libname_gspphot` field, and a subset of those parameters are also found in `gaiadr3.gaia_source`.

3.2 Hot stars

3.2.1 OB spectral types

O and B spectral-type stars are intermediate to massive stars. Due to their large masses and therefore very high central pressures, and thus temperatures, they consume their nuclear fuel at a very efficient rate, and thus evolve very rapidly. Due to the high central temperatures, they also go through several nuclear-burning phases and produce the heaviest elements. They are therefore key contributors to the enrichment of the interstellar medium through supernova. As they remain main sequence stars for such a short time they do not move far from their birth place, and this fact makes them excellent probes of the structure and dynamics of star forming regions and spiral arms, see e.g. Gaia Collaboration, Drimmel et al. (2023); Gaia Collaboration, Creevey et al. (2023).

In stellar physics OB stars contribute to our understanding of:

- mass-loss due to winds from the important contribution of radiation pressure,
- observations of their photospheres and stellar oscillations allow us to test opacities as produced by laboratory experiments,
- as these objects have many He lines, they can provide observational measures of He abundances and can help to constrain chemical enrichment laws, as well as constrain chemical transport mechanisms,
- the identification of the main sequence stars allows one to put important constraints on their ages (between 10 million – 200 million years), in particular if there is a companion or if the star is a member of a cluster,
- the more evolved OB stars become blue horizontal giant branch stars, are progenitors of supernova, and later become neutro'B'n stars or black holes, and are potential contributors to detectable gravitational waves.

In GDR3, one of the key products for OB stars is their identification, and then initial characterisation. Within Apsis, the ESP-ELS (extended stellar parametrizer – emission line stars) module provides firstly a

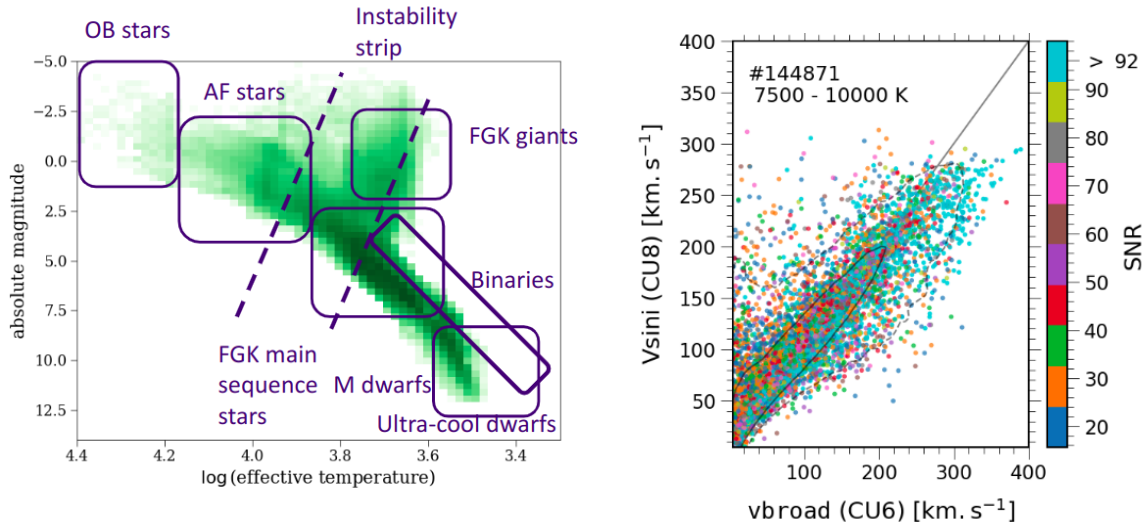


Fig. 6. *Left:* Different regions of the HR diagram presented in these lectures. *Right:* A comparison between two independent determinations of line broadening; x-axis `vbroad` and y-axis `vsini_esphs`. Credit: Ulla, A. et al., 2022, Gaia DR3 Online Documentation.

spectral typing of all sources up to $G < 17.65$, by classifying them into the broad categories of 'O','B', 'A', 'F', etc (in the archive the parameter is labelled `spectraltype_esphs`). Once the star has been identified, it is then tested if it is an emission line star, assigned a probability of being a particular emission-type, e.g. `classprob_espels_wcstar`, and then finally classified according to its emission as a Be star (beStar), Herbig Ae/Be star (HerbigStar), T Tauri star (TTauri), active M dwarf star (RedDwarfEmStar), Wolf-Rayet WC (wC) or WN (wN), or a planetary nebula (PlanetaryNebula) e.g. `classlabel_espels = wC`. The stars that have been identified as O, B, or A, are further analysed by the ESP-HS (hot star) module, which exploits the mean BP/RP and RVS spectra, to produce $T_{\rm eff}$, $\log g$, A_0, $E(G_{\rm BP}-G_{\rm RP})$, and $v \sin i$ if the RVS spectrum is available, with the assumption of solar metallicity. These parameters are also found in the `gaiadr3.astrophysical_parameters` table and are called `teff_esphs`, `logg_esphs` and `vsini_esphs`. If the RVS spectrum is not available, then $v \sin i$ is not derived. No specific treatment of rotation is included in the models that are used to produce the simulations of the low-resolution BP and RP spectra.

A second set of OB characterisation is provided by GSP-Phot, which includes its $T_{\rm eff}$, $\log g$, [Fe/H], distance, A_0 and $E(G_{\rm BP}-G_{\rm RP})$, by using the OB star library, and found in the `gaiadr3.astrophysical_parameters_supp` table. All sources have a solution with the OB library, so one should take care in selecting the sources. One could alternatively use directly the result in the `gaiadr3.astrophysical_parameters` table, where the `libname_gspphot` is 'OB' i.e. the preferred solution for this star is found to be 'OB' and not 'A', 'MARCS' or 'PHOENIX'. If results from both ESP-HS and GSP-Phot are available, the user should prioritise the former where both RVS and BP/RP spectra are used, because the models have been more finely tuned to the data focussing on hot stars. Further information on hot stars is given in the lecture by Frémat (2024).

Finally, just to be complete, although this is applicable to all sources, kinematic information such as proper motions, radial velocity and positions for these stars are given in `gaiadr3.gaia_source`.

3.2.2 AF spectral types

Spectral-type A and F stars have masses between approximately 1.4 – 2.1 $M_\odot$, and $T_{\rm eff}$ in the range of 7 600 – 10 000 K. Their spectra are dominated by H lines. Some famous examples are Vega and Altair, and a well known planet-host β Pic. They are in general rapid rotators but their rotational velocity changes rapidly with sub spectral type (or $T_{\rm eff}$) and more slowly with stellar evolution. Being intermediate mass stars, their typical main sequence lifetimes are on the order of about 1 billion years, and this makes them interesting probes for stellar ages, in particular in clusters. Many of these stars are also of variable type known as δ Scuti or γ Dor, or hybrid-mode, and show large amplitude pulsations that can usually be detected from ground-based observatories. For this reason, these are some of the pulsating stars that have been studied for a very long time.

However many questions regarding their pulsation and damping mechanisms still exist.

In stellar physics open questions relating to AF stars concern:

- the impact of rotation on the stellar structure and its evolution,
- the understanding of magnetic braking, the loss of a star's angular momentum through magnetic fields,
- physical processes in the interior of stars which causes pulsation instability,
- the high precision determination of $T_{\rm eff}$ from the analysis of the highly-temperature sensitive H lines.

GDR3 contributes to these open questions in many aspects. Firstly, the time series for a total of about 10 million variable stars, of which a large fraction are A–F type, are available in the archive. This variability comprises pulsation periods, eclipsing binaries and rotational modulation allowing the determination of rotation period. Secondly, the variability catalogue already provides many of these fundamental variability statistics. Some specific examples in the archive are the `gaiadr3.vari_ms_oscillator` table which provides frequencies, amplitudes, and phases of the largest detected amplitudes, or the `gaiadr3.vari_rotational_modulation` which provides statistics and periods concerning the rotational information due to e.g. spots on stars, or the `gaiadr3.vari_short_timescale` table which gives timescales and periods, along with other statistics of sources with short timescale variability in the photometry.

Gaia Collaboration, De Ridder et al. (2023) presents a specific exploitation of the variability tables while focussing on pulsations in main sequence stars with M $\geq$ 1.3 $M_{\odot}$. They concluded that the most populated variability class is the δ Scuti variable class, they identified over 10 000 g-mode pulsators including many γ-Dor stars (see their Fig. 6), an analysis of the empirical period-luminosity relation of these stars revealed different regimes which depend on the oscillation period (see their Fig. 12), and they further concluded that stellar rotation attenuates the amplitude of the dominant oscillation mode of these stars. All of these results come from the analysis of the data available in GDR3.

The astrophysical parameters for A and F stars have been derived by three different Apsis modules. The first is GSP-Phot, which has been mentioned above, and parameters from this module are available for all sources down to $G < 19$. Secondly, some astrophysical parameters have been derived for the cooler A and F stars by GSP-Spec (upper limit for reliable `_gspspec` is approximately 7 500 K), but rotational effects are not taken into account in GDR3 in their analysis of the RVS spectra. The impact of this parameter on the quality of the data can be quite high, and one should use the GSP-Spec flags to quantify this impact, see the online documentation for details. For the cooler F stars, rotation has a smaller impact on the spectral lines, and these parameters are therefore mostly reliable (see flags). Finally the ESP-HS module, mentioned in the above section, analyses the A stars by combining, when possible, the BP/RP and the RVS spectra. As explained earlier, the $v \sin i$ is derived along with the atmospheric parameters, and a comparison between the `vbroad` parameter in `gaiadr3.gaia_source` that is derived by the RVS processing and the `vsini_esphs` is shown in Fig. 5 right panel for stars with $T_{\rm eff}$ between 7 500 and 10 000 K. The one-to-one line is shown to guide the eye, and one can note in general a good agreement between both determinations, although a scatter is seen. The different values arise from the very different methodologies used to determine the parameter. For details, one should consult the Gaia Archive documentation, and the published papers where the processing of these parameters is described Creevey et al. (2023); Frémat et al. (2023).

3.3 Low mass main sequence stars

Low-mass main sequence stars are of spectral type K, M, L and T. These are typically faint objects because of their small size and thus low luminosity, and for this reason they are also difficult to detect and characterise. Below about 0.3 $M_{\odot}$, these stars are fully convective. Open questions in stellar physics in this low mass regime concern the transition region between fully and partially convective regimes in stellar interiors and the impact on energy transport in the fully convective regime. The transition region at the stellar – sub-stellar border is also an area of active research, and this contributes to the understanding of the initial mass function. As these objects also evolve very slowly – on the order of 25 to over 100 billion years, many objects are still in their pre-main sequence phase, and indicators of youth or activity contribute to developing an evolution (pre-main) sequence. GDR3 plays an important role for the understanding of these low main sequence objects firstly from their precise photometry and parallaxes, which allows their detection. Many low-mass companions in binaries have also been detected using their proper motions, and these binaries often have a more massive component

which can contribute to determining the low mass component's mass and / or age, e.g. (Sarro et al. 2023; Gaia Collaboration, Creevey et al. 2023).

In addition to the astrometric and photometric information available in `gaiadr3.gaia_source`, the Apsis system also provides a characterisation of stars in the low mass region. GSP-Spec provides reliable atmospheric parameters of stars down to a $T_{\rm eff}$ of about 3 600 K, while GSP-Phot also provides these parameters down to 2 500 K, which is the lower $T_{\rm eff}$ limit of the MARCS grid. In addition to these atmospheric parameters, FLAME also provides some evolutionary parameters (luminosity, radius, mass, age) for these same sources, although masses and ages are limited to stars with masses above 0.5 $M_{\odot}$.

An additional Apsis module ESP-UCD (*Extended Stellar Parametrizer – Ultra Cool Dwarfs*) exploits the RP spectra, along with the astrometry and photometry for identification, in order to derive a $T_{\rm eff}$ for the very low mass sources, focussing on the range of 500 K – 2 700 K. The module uses a Gaussian Process regression model that is trained on a set of Gaia RP spectra. The choice to use this empirical training was due to the shortcomings in the BT Settl spectra, which inhibited a direct use of simulated data. This latter fact also highlights the important contribution that Gaia DR3 potentially brings to providing constraints on the model spectra of such low mass objects. In GDR3 the `gaiadr3.astrophysical_parameters` table one can find the $T_{\rm eff}$ for around 90 000 ultra cool dwarfs, along with a quality flag (0, 1, or 2) with 0 being the most reliable. Fig. 7 shows a histogram of the UCD $T_{\rm eff}$, and the colour code indicates the flag (blue = 0, green = 1, grey = 2). Note the log scale for interpretation of the numbers. The inset shows the distribution of these parameters by G and parallax.

Further analysis of these UCDs is presented in Gaia Collaboration, Creevey et al. (2023), where the sources with flags = 0 or 1 are coupled with infra-red data from 2MASS (Cutri et al. 2003), and luminosities and radii are derived. These luminosities and radii are found in the `gaiadr3.gold_sample_ucd` table, and one can combine the data from both tables (or more) for further exploration and analysis of these low mass objects. Fig. 7 right panel illustrates the relationship found between luminosity and $T_{\rm eff}$. Only error bars for sources with $T_{\rm eff} < 1\,900$ K are shown, and the colour-code shows the logarithm of the χ^2 fit. One can clearly see a change in the slope of the luminosity – $T_{\rm eff}$ (and radius – $T_{\rm eff}$) diagram around 1 900 K and again at 1 500 K which could hint at specific interior physical processes taking place which change the output luminosity (and radius).

Further scientific exploitation of these objects can also be done by examining these objects that are companions to more massive stars, identified through their common motions. An analysis of their more massive counterpart, such as their ages, can provide important constraints on the nature of these low mass objects and the limits of the stellar / sub-stellar regime, see e.g. Section 5 in Gaia Collaboration, Creevey et al. (2023). The readers are also referred to the lecture notes by Réyle (2024) who presents in much more detail the open questions and available data on ultra-cool dwarfs.

3.4 Exotic stars

There are many different exotic objects in all regions of the HR diagram, that are not indicated in Fig. 5. Among these exotic stars, one can mention many emission-line stars such as P Cygni stars with winds, accretion inflows in T Tauri stars, rotating discs around Be stars, explosions such as novae and type II supernovae, transfer of material between stars e.g. in symbiotics, and carbon stars, among many others. Firstly the identification of these objects is of great importance and GDR3 plays an important role here. Further classification of the objects, and an analysis of their mean and time domain observed properties provides important constraints towards a more concrete understanding of the objects.

As mentioned in Sect. 3.2 the ESP-ELS module provides class probabilities and suggests a class labelling of emission-line stars. One can revisit the class probabilities to re-label these objects. There are a total of 57 511 class probabilities and labels of ELS in GDR3 found in the `gaiadr3.astrophysical_parameters` table.

A spectral typing of 218 million sources is also provided. Apart from the standard O, B, A, F, G, K, and M spectral classes, an additional class of 386 936 candidate carbon stars 'CSTAR' has been also identified (`spectraltype_esphs = 'CSTAR'`). Carbon stars are typically asymptotic giant stars with carbon-enriched atmospheres due to either the transfer of matter in binary systems or to the pollution by nuclear He fusion products from the interior to the exterior layers of stars. An analysis of the spectral indices derived from the BP/RP spectra is provided in Gaia Collaboration, Creevey et al. (2023) where in addition to the class labels, detailed analysis of the band head strengths of the C_2 lines at 482.3 nm and 527.1 nm, and CN lines at 773.3 nm and 895.0 nm of the BP/RP spectra is performed. An additional sample of 15 740 stars with the strongest CN features is provided in the `gaiadr3.gold_sample_carbon_stars` table.

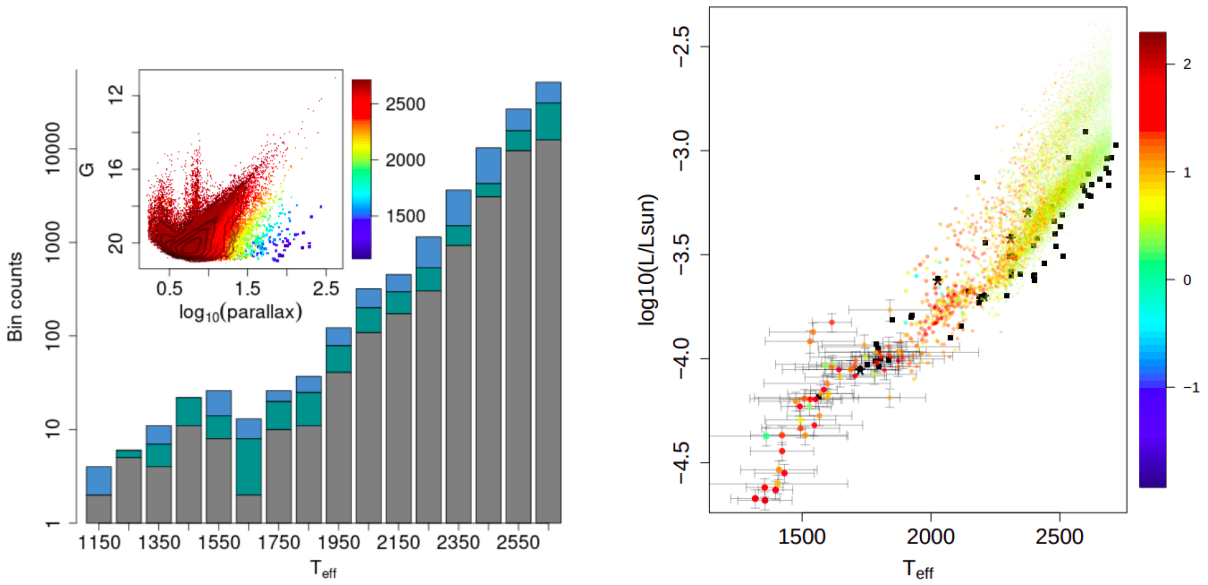


Fig. 7. *Left:* Distribution of the $T_{\rm eff}$ of ultra cool dwarfs in GDR3 produced by the Apsis/ESP-UCD module. Credit: Creevey et al. 2023, A&A, 674, A26. *Right:* Luminosities and $T_{\rm eff}$ for the UCD sources with flag = 0 or 1. Error bars are shown only for the sources with $T_{\rm eff} < 1\,900$ K. Credit: Gaia Collaboration, Creevey et al. 2023, A&A, 674, A39.

An analysis of the H-alpha line in the BP/RP spectra is also provided in GDR3 for 235 million objects. This line is also used for the identification of the ELS: Be stars, Herbig Ae/Be stars, T Tauri stars, active M dwarf stars, Wolf Rayet WC stars and WN stars, and planetary nebula, see the parameter `ew_espels_halpha` and related products in the `gaiadr3.astrophysical_parameters` table.

In addition to these characterisations, light curves are provided for those stars considered variable in GDR3. A simple query on the archive such as the following

```
select count(*)
from gaiadr3.astrophysical_parameters as ap
inner join gaiadr3.vari_summary as vs on vs.source_id = ap.source_id
where classlabel_espels is not NULL
```

indicates that there are light curves of 14 240 sources classified as ELS and also classified in the variability table, and therefore have epoch photometry. Exploration of these light curves and mean spectra are left to the users.

3.5 FGK stars

This section focusses on FGK main sequence, sub-giant and giant stars, which are the most numerous spectral types in GDR3. They have typical lifetimes on the order of 1 - 15 billion years, so they are ideal sources to trace the star formation history and evolution of our Galaxy. The determination of their stellar parameters allows one to explore stellar populations and knowledge of their chemical abundances allows one to trace the chemical evolution of the Galaxy, and explore processes of chemical transport in stellar interiors and atmospheres.

Focussing specifically on stellar physics, the following open questions can be addressed by GDR3: how are chemical elements in stellar interiors transported? How does magnetism in the interiors and atmospheres of stars govern the stellar structure and evolution? How is energy transported from the interior to the surface of stars and through their atmospheres? And what is the impact of interface regions such as radiative and convective zones, interiors and atmospheres, and atmospheres and the interstellar medium? What is the impact of binarity and activity on the evolution of stars? By focussing on systems where numerous observational constraints can be imposed, such as clusters, binaries, and seismic targets, how well can we determine stellar ages? What is the impact of different micro- and macroscopic processes on the stellar structure, evolution and global parameters and how do these processes influence the knowledge of their ages? By accurately characterising stars and their ages, what can we learn about evolutionary processes in stars (e.g. rotation), how do exoplanetary systems form and evolve? And how does our Milky Way form and evolve? Is our own Sun a typical star and how does it

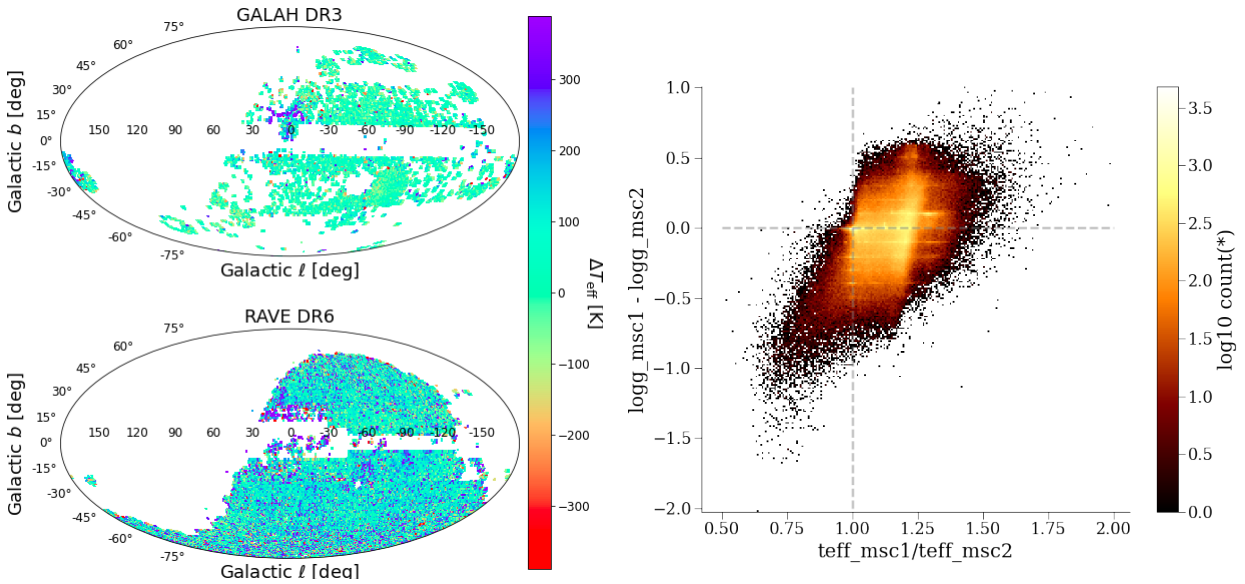


Fig. 8. *Left:* Differences in $T_{\rm eff}$ from GSP-Phot compared to two catalogues from the literature as a function of sky position. Credit: Andrae et al. 2023, A&A, 674, A27. *Right:* Distribution of the differences in $\log g$ and $T_{\rm eff}$ of the two components derived by MSC, with the colour code indicating number of sources. Credit: Creevey et al. 2023, A&A, 674, A26.

compare to other sun-like stars in terms of magnetism, rotation, activity, and chemical composition? And can we place our own Solar System in the context of other planetary systems (ages, masses, radii of host star and its planets)? All of these questions can be addressed, at least partially, using GDR3. In the following sections we describe the relevant content of the archive that can allow one to address these questions.

3.5.1 Atmospheric parameters from BP/RP spectra

BP and RP spectra for about 200 million sources are available in DR3 for analysis. As part of the Apsis system, two modules provide estimates of atmospheric parameters and extinction, by exploiting the mean BP/RP spectra, astrometry and photometry. The first is GSP-Phot, described earlier, which analyses FGK stars for $G < 19$ using MARCS (2 500 K – 8 000 K) and PHOENIX (4 000 K – 10 000 K) atmospheric models coupled to evolutionary tracks. $T_{\rm eff}$, $\log g$ [Fe/H], A_0, $E(G_{\rm BP} - G_{\rm RP})$, and distance are derived, along with uncertainties and monte carlo markov chains. The latter are datalink products and can be downloaded after quering of sources. The GSP-Phot products are found in the `gaiadr3.astrophysical_parameters_supp` table for all models, with the best library result in the `gaiadr3.astrophysical_parameters` table and a subset of the parameters are copied to `gaiadr3.gaia_source`. One should prioritise the use of the latter two tables for selecting the FGK type stars, and one can then choose to use the homogenous set of MARCS or PHOENIX found in the `supp` table. To complement the data in GDR3, a proposed calibration of the metallicities is made available due to a systematic bias observed in the results. The calibration is made available as a software tool available on the archive webpages, and can be found directly here on github: https://github.com/mpi-astronomy/gdr3apcal. In addition to the atmospheric parameters, the radius and absolute magnitude are also derived by GSP-Phot.

A comparison of $T_{\rm eff}$ with external catalogues is shown on the left panel of Fig. 8 for GALAH DR3 (Buder et al. 2021) and RAVE DR6 (Steinmetz et al. 2020). Biases in the $T_{\rm eff}$ can be seen for regions close to the galactic plane, and this is a result of the high correlation between $T_{\rm eff}$ and A_0 in the BP/RP spectra. Making a refined selection on the sources, such as that proposed by Gaia Collaboration, Creevey et al. (2023), see below, where astrometric and photometric data are of higher quality, results in much smaller biases when comparing to external catalogues, see e.g. Fig. 9 from that paper.

A second module also exploits the BP and RP spectra in a similar manner to GSP-Phot, but it assumes that the source is a composition of two individual (non-resolved) components. It is called the *Multiple Stellar Classifier* MSC, and provides for each source two $T_{\rm eff}$, two $\log g$, and a unique distance and A_0, assuming a solar metallicity prior. The module is trained on empirical data covering the FGK spectral ranges. Results for

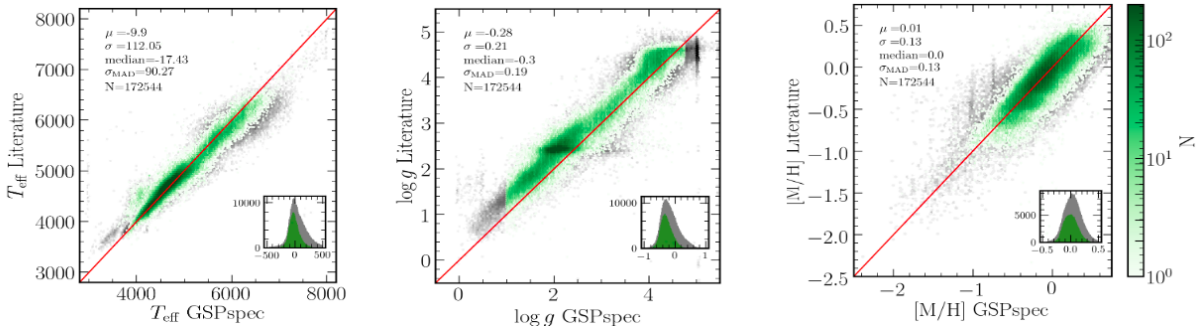


Fig. 9. Comparison to the literature of $T_{\rm eff}$, $\log g$, and [Fe/H] derived by GSP-Spec from the RVS spectra for high quality flagged sources. Credit: Recio-Blanco et al. 2023, A&A, 674, A29.

all sources to $G < 18.25$ are provided, and one should use the log-posterior parameter to evaluate how reliable the results are, i.e. how likely the source is to be composed of two non-resolved components. These parameters are also found in the `gaiadr3.astrophysical_parameters` table as `teff_msc1`, `logg_msc2`, etc. Fig. 8 right panel illustrates the range of parameter differences that are explored by MSC. Components are generally within 1 dex of $\log g$ of each other, while $T_{\rm eff}$ ratios vary typically between 0.75 and 1.50.

3.5.2 Atmospheric parameters from RVS spectra

The atmospheric parameters $T_{\rm eff}$, $\log g$, [Fe/H], and [α/Fe] are derived from the RVS spectra by the GSP-Spec module (Recio-Blanco et al. 2023) for approximately 6 million sources with $G < 15$ focussing on the 3 500 – 8 000 K range, see the accompanying presentation in this school from Recio-Blanco (2024). Two algorithms process the data and the results based on the MatisseGauguin algorithm are found in the `gaiadr3.astrophysical_parameters` table, e.g. `teff_gspspec`, while the results based on the ANN method are found in the `gaiadr3.astrophysical_parameters_supp` table, e.g. `alphafe_gspspec_ann`. In addition to the atmospheric parameters, a flag containing 42 characters is proposed, with each character being a number 0, 1, 2, or 9 indicating the quality of a source based on a particular parameter, see the online documentation for details. The first 11 characters of the flag are related directly to the quality of the four atmospheric parameters and consider the sources of errors from the input data, e.g. a large `rv_error`, and the quality of the parametrisation itself. The other characters refer to the chemical abundances. A comparison of the $T_{\rm eff}$, $\log g$, and [Fe/H] with the literature for sources with flags 0 or 1 are shown in Fig. 9. A small bias is seen in $\log g$ and a calibration is proposed for both $\log g$, [M/H], and the individual abundances. The description of the calibration is given in Recio-Blanco et al. (2023) and can also be found directly on the archive software tools pages. The reader is referred to the accompanying presentation by Recio-Blanco (2024) for more details including scientific exploitation of the parametrization and the chemical abundances in the context of galactic archaeology.

3.5.3 Evolutionary parameters

Mass and evolutionary parameters are derived by the *Final Luminosity Age Mass Estimator* FLAME. These comprise radius, luminosity, gravitational redshift, masses, ages and evolutionary stage, along with two auxiliary products, a bolometric correction and a flag. These parameters are produced by using both the GSP-Phot and GSP-Spec atmospheric parameters as input and therefore two sets of evolutionary parameters are available in GDR3. The first ones are based on the GSP-Phot parameters and the results for 270 million sources with $G < 18.25$ are found in the `gaiadr3.astrophysical_parameters` table, e.g. `mass_flame`. The second set of results for 6 million sources, based on the GSP-Spec parameters, are found in the `gaiadr3.astrophysical_parameters_supp` table, e.g. `age_flame_spec`.

Luminosity, $L_\star$, is derived from the G, parallax (sometimes `distance_gspphot` is used, see `flags_flame`), A_G, and a bolometric correction based on the input atmospheric parameters. The bolometric correction is made available to the user and can be accessed from the Gaia archive software tools or found directly here: `https://gitlab.oca.eu/ordenovic/gaiadr3_bcg`. Then, using $T_{\rm eff}$ and $L_\star$, the radius is derived. From radius and $\log g$, the gravitational redshift is then calculated. Finally, using $T_{\rm eff}$, $L_\star$, and [Fe/H], the mass, age

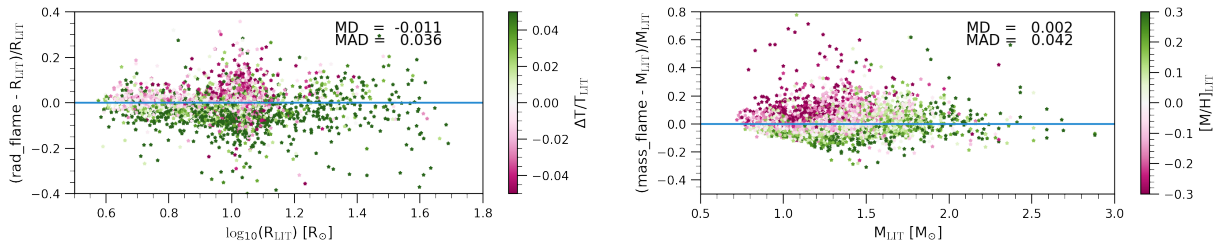


Fig. 10. *Left:* Comparison of `radius_flame` with asteroseismic radii from Pinsonneault et al. (2018) for giant stars. Colour-code indicates the relative difference in $T_{\rm eff}$ estimates assumed in the analysis, which partially explains the dispersion that is found. *Right:* Comparison of `mass_flame` with mass estimates from Casagrande et al. (2011). Colour-code indicates the metallicity from the literature. Credit: Fouesneau et al. 2023, A&A, 674, A28.

and evolutionary status are inferred using the BASTI stellar evolution models (Hidalgo et al. 2018) with a prior of solar metallicity.

A comparison with the literature for mass and radius is shown in Fig. 10. Left and right panels compare giant star radii and main sequence star masses from Pinsonneault et al. (2018) and Casagrande et al. (2011), respectively. The colour-code indicates on the left panel the relative difference in assumed $T_{\rm eff}$ for the inference, and on the right panel, the literature metallicity, and these can explain some contributions to the scatter that is seen. These figures illustrate that the results are in general agreement with what is expected.

The impact of the solar metallicity prior has been studied in Creevey & Lebreton (2022), which can be found in the Gaia DPAC Public Documents, and empirical corrections to the mass are proposed given the user's input metallicity, which take the form of

$$M_{\rm cor} = M_{\rm FLAME} - f(M_{\rm FLAME}, [M/H]) \quad (1)$$

where $f(M_{\rm FLAME}, [{\rm M/H}]) = \sum_{\rm i=0}^{1} a_{\rm i} x^{\rm i}$ is a linear function to apply to the published mass and the user's [Fe/H] to produce a corrected mass. The coefficients are made available in Table 6 of Creevey & Lebreton (2022), and Fig. 11 left panel illustrates results from simulations showing the residuals, in the sense of $M_{\rm true} - M_{\rm FLAME}$, when we apply (blue) or not (colours) the proposed corrections for main sequence stars for a metallicity of [Fe/H] = –0.70. The right panel shows a comparison of input evolution index assuming [Fe/H] = –0.70 versus the recovered evolution index assuming a solar metallicity prior for main sequence ($<$ 420), subgiant (420 – 490), and giant ($>$ 490) stars. While a one to one correspondance is not clear, FLAME still identifies whether the star is main sequence, sub-giant, or giant, even with an incorrect metallicity assumption. Corrections to the evolutionary status are also proposed in Creevey & Lebreton (2022) Table 7.

3.5.4 Stellar activity

Within Apsis, the *Extended stellar parametrizer – cool stars*, ESP-CS, analyses the cores of the Ca II IR Triplet lines in the RVS spectra for evidence of stellar activity by comparing the observations with models defining a purely photospheric spectrum. The excess in the core, given in nm, is related to activity in the chromosphere of the star or mass accretion, with a value on the order of 0.03 – 0.05 nm separating the two regimes, see Lanzafame et al. (2023). In GDR3, about 2 million estimates of `activityindex_espcs` are derived for stars with $G < 15$ and $T_{\rm eff}$ between 3 000 K and 7 000 K. An uncertainty and a flag `activityindex_espcs_input` indicating the source of the atmospheric parameters (GSP-Phot or GSP-Spec) are also published. Readers are referred to Creevey et al. (2023); Lanzafame et al. (2023) for full details of the processing and interpretation of the index.

3.5.5 Metallicities from variability

Metallicities and absorption are possible to derive for RR Lyrae stars using pulsation periods and phase differences using multi-colour photometry. This is described in detail in Clementini et al. (2023). In Gaia DR3 metallicities for 133 557 RR Lyrae are available in the `gaiadr3.vari_rrlyrae` table. A comparison with a subset of metallicities from the GSP-Spec pipeline is shown in Fig. 23 of that paper, which provides an independent validation of the parameters.

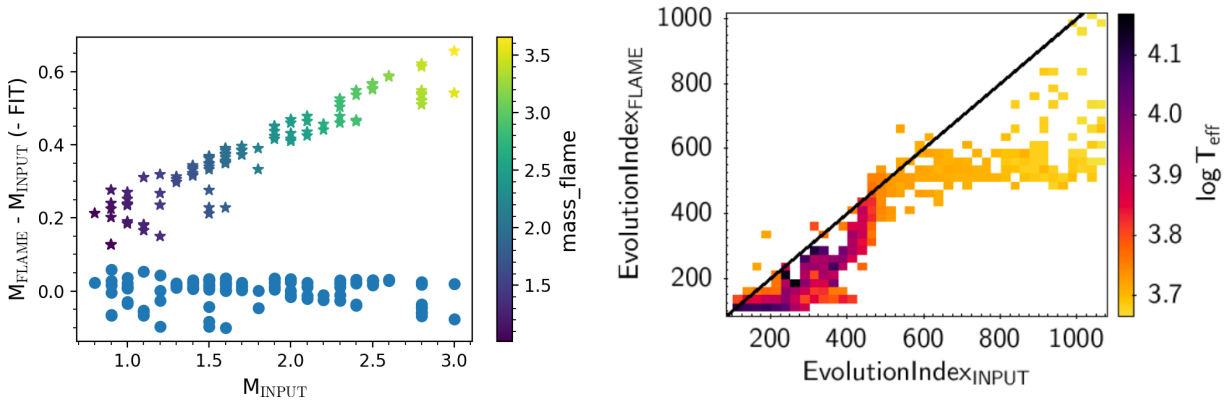


Fig. 11. *Left:* Residuals of mass from FLAME simulations for [Fe/H] = –0.70, indicating the performance of FLAME using a solar metallicity prior for main sequence stars. The coloured stars indicate the residuals without applying a correction, while the blue dots show the residuals after applying the empirical corrections. *Right:* Comparison of input evolution index assuming [Fe/H] = –0.70 with the output evolution index assuming a solar prior. Main sequence, subgiants, and giants are delimited at the values of 420 and 490. Credit: Creevey & Lebreton, 2022, Public DPAC Documents.

3.5.6 Examples of scientific exploitation

Several examples of exploitation of the astrophysical parameters for FGK stars are described in Gaia Collaboration, Creevey et al. (2023). Here we describe some of these to give some concrete uses of GDR3.

FGKM gold sample A gold sample of about 3.3 million stars is proposed based on very high precision astrometric and photometric data, along with the highly reliable atmospheric and evolutionary parameters from Apsis. These data can be found in the archive in the table `gaiadr3.gold_sample_fgkm_stars`. The distribution of these sources on the sky is shown on the left panel of Fig. 12. This is of interest to examine because it is mostly dominated by the parallax precision and signal-to-noise ratio which is a result of the number of observations for each source and their brightness. One can immediately interpret that there is not a homogenous coverage in terms of quality of data across the full sky, and one must invoke some quality indicators for scientific exploration, e.g. uncertainties, flags, auxiliary information, such as `ruwe`. This work makes a very highly selective cut on many parameters in terms of quality because the aim was to produce a relatively small sample with the best data quality. The resulting parameters for the final sample is shown in the HR diagram on the right panel. Plotting a similar HR diagram, but colour-coding by the stellar age, reveals a gradual expected gradient which indicates the high precision of these ages, see their Fig. 12.

Exoplanetary stellar properties Cross-matching the golden sample of FGKM stars with known exoplanet host stars using `http://exoplanets.org/` dated March 2022, results in a sample of about 600 planets. About 100 of these have measured radial velocities and light curve transits from the literature, allowing the derivation of masses and star-planet relative radii. We used the stellar masses and radii from the gold sample in order to derive absolute radii, masses, and ages of the 96 exoplanetary systems. The planet properties are made available through CDS at `J/A+A/674/A39` (Gaia Collaboration, Creevey et al. 2022), and Fig. 13 illutrates planet mass – radius diagrams for Earth-like (left) and Jupiter-like (right) planets, along with several models of planet interiors from Guillot & Gautier (2015) and Zeng et al. (2016).

Solar analogues Our Sun is the main reference point for understanding the structure and evolution of stars and their atmospheres. However as it can not be observed with the same instruments as those used for other stars, it is not possible to put it on the same scale with absolute accuracy. We therefore rely on solar analogues which are stars that resemble the Sun in terms of specific parameters, which depend on the astrophysical question in hand. For example, in order to derive accurate $T_{\rm eff}$ and [Fe/H] of other stars, we would aim to choose stars with the same $\log g$, $T_{\rm eff}$, and [Fe/H] of the Sun, while if we aim to study age-chemical abundance relations, then we would select stars that also have similar masses to the Sun.

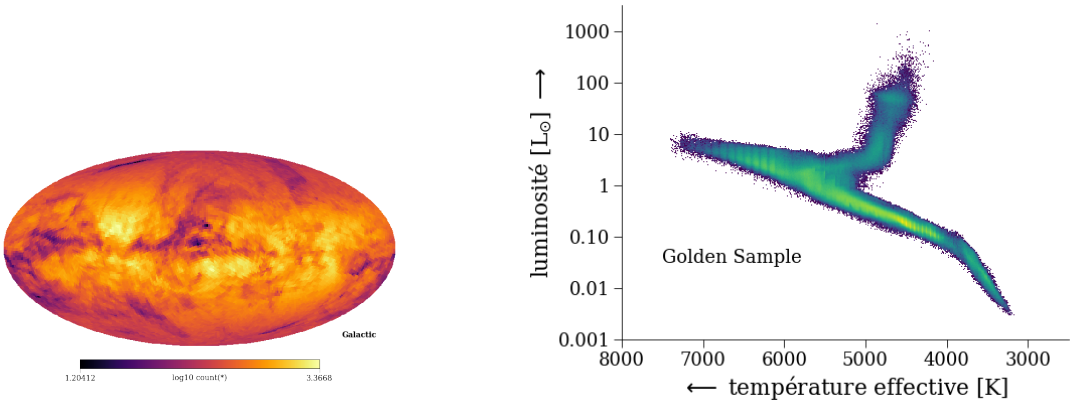


Fig. 12. *Left:* Distribution of sources on the sky for the FGKM gold sample comprising approximately 4 million stars. *Right:* Resulting HR diagram using the gold sample with colour-code indicating number density of stars. Credits: adapted from Gaia Collaboration, Creevey et al. 2023, A&A, 674, A39.

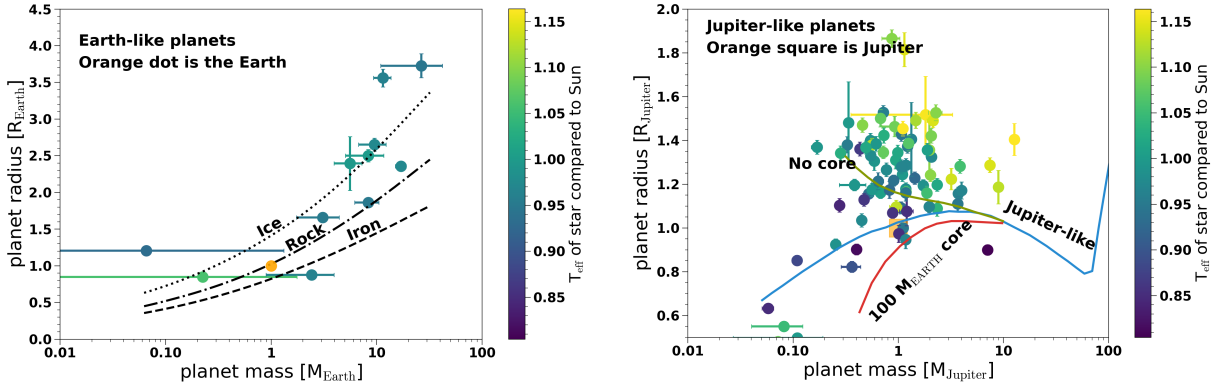


Fig. 13. Planet mass radius diagrams for Earth-like (left) and Jupiter-like (right) planets. The solar system planets are shown in orange, and several theoretical model predictions are overlaid indicating the stellar bulk composition and density. Credit: adapted from Gaia Collaboration, Creevey et al. 2023, A&A, 674, A39.

Finding solar analogues has not been a very easy task, as it relies on obtaining many different types of measurements of several candidate analogues, and observing, reducing and analysing these observations. With Gaia DR3 providing a homogenous catalogue of stellar parameters for the full sky, this job becomes a lot easier.

A selection of 5 863 solar analogues was made based on matching the Sun's atmospheric parameters, and mass and radius, to the parameters provided by GSP-Spec and FLAME, along with high precision parallaxes. Of these, 1046 have published RVS spectra. In Fig. 14 we show the the median RVS spectra of the 1046 solar analogues and the *scatter* includes respectively 68% and 90% of these analogues. The similarity of the spectra attests to the correct parametrization in Gaia DR3. The full sample of 5 863 analogues can be found on the archive table `gaiadr3.gold_sample_solar_analogues`.

Performing the same research using the GSP-Phot and FLAME parameters results in a total of 234 779 candidates, 7884 of which have RVS spectra and all of which have XP spectra. We note, however, that as XP spectra are affected by interstellar extinction, one must consider the scientific case before selecting the sources.

By using the `gaiadr3.gold_sample_solar_analogues` tables and choosing those sources with very low extinction, by selecting on `azero_gspphot` < 0.001, we obtain 682 solar analogue sources for which we can derive the absolute magnitude and the intrinsic colours of the Sun, see Gaia Collaboration, Creevey et al. (2023). Using the photometric data directly from the Gaia archive, we obtain

$$(G_{BP} - G_{RP})_{\odot} = (0.818 \pm 0.029) \text{ mag} \quad (2)$$

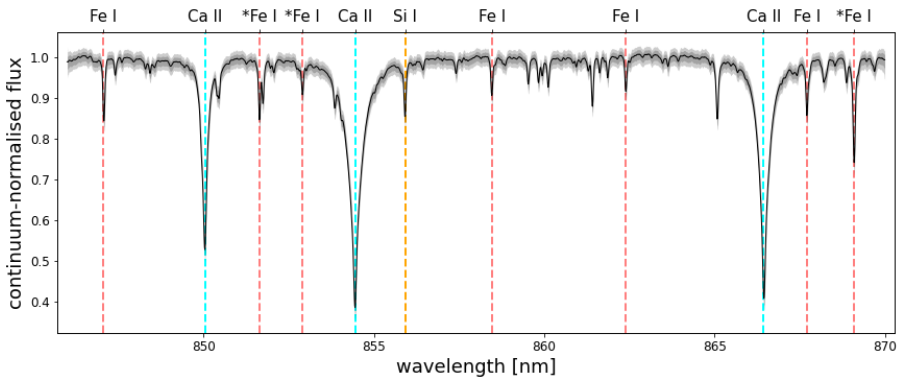


Fig. 14. The Gaia DR3 RVS spectra of 1046 solar analogues. The inner and outer grey shaded regions contains 68% and 90% of the sample. Credit: ESA/Gaia/DPAC adapted from Gaia Collaboration, Creevey et al. 2023, A&A, 674, A39.

$$(G_{BP} - G)_\odot = (0.324 \pm 0.016) \text{ mag} \tag{3}$$

$$(G - G_{RP})_\odot = (0.494 \pm 0.020) \text{ mag} \tag{4}$$

By combining the resulting sources with the data from Gaia Collaboration, Montegriffo et al. (2023) using the `gaiadr3.synthetic_photometry_gspc` table of synthetic magnitudes derived from the XP spectra, we can further calculate many other solar colours, e.g. $(r-i)_{\rm SDSS,\odot} = (0.134\pm 0.012)$ mag, $(i-z)_{\rm SDSS,\odot} = (0.33\pm 0.011)$ mag, or $(B-V)_\odot = (0.64 \pm 0.03)$ mag. The example query to extract the data from the three tables is given as follows

```
select gs.source_id, b_jkc_mag, v_jkc_mag, i_jkc_mag, r_jkc_mag,
g_sdss_mag, r_sdss_mag, i_sdss_mag, z_sdss_mag, phot_g_mean_mag,
phot_bp_mean_mag, phot_rp_mean_mag, parallax
from gaiadr3.gold_sample_solar_analogues as gs
inner join gaiadr3.synthetic_photometry_gspc as sp on sp.source_id = gs.source_id
inner join gaiadr3.gaia_source as gg on gg.source_id = gs.source_id
where azero_gspphot < 0.001.
```

The mean absolute magnitude is $M_{G,\odot} = (4.614 \pm 0.179)$ mag in agreement with the analysis in Creevey et al. (2023) where $M_{G,\odot} = 4.66$ mag is adopted for the FLAME module, and $M_{G,\odot} = 4.67$ mag which was obtained by Casagrande & VandenBerg (2018) using Gaia DR2 data.

The future of our Sun By selecting sources with 1 solar mass and solar metallicity without any constraints on age, we can construct an evolution sequence of a star like the Sun. This is illustrated in Fig. 15 where we have selected 5323 stars with the given requirements and plot them on a HR diagram in orange over a background of stars from the FGKM gold sample. In each panel from left to right we choose one star at a particular age and highlight its position, size, and age (the size is not to scale) according to Gaia DR3. Each highlighted star therefore shows what the Sun would look like at different points in its future, starting with its current age of 4.58 billion years until it reaches the base of red giant branch where it swells up at the age of 11.75 billion years. Further detailed studies of these individual stars can be performed to study the fate of our Sun by refining the age estimates, fundamental, structural, and atmospheric properties, and chemical compositions. Of particular interest would be the identification of planetary systems around these stars to study the evolution and habitability of future solar systems.

3.6 Variability and multiplicity

Variable stars are those that change their brightness over time, either regularly, semi-regularly, or irregularly. The changes in brightness can be due to the intrinsic nature of stars (true variable stars), or due to the environment, such as an interaction with another star, or a planet passing in front of the star, or any binary or multiple system. Variability and multiplicity can also be detected through the astrometric measurements, e.g.

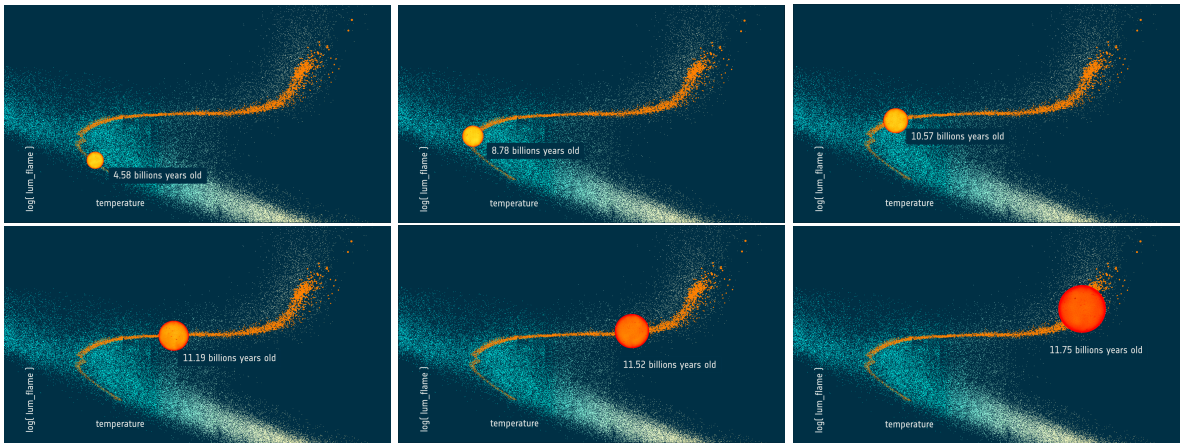


Fig. 15. HR diagrams showing some of the Golden Sample of FGKM stars. Overlaid in orange are a selection of stars of one solar mass and solar metallicity from that sample which traces out an evolution sequence. Each panel highlights one star along the evolution sequence, with its size (diameter) represented by the size of the symbol, although not to scale, and the age of star indicated in the panel. This represents most of the lifecycle of our own Sun. Credit: ESA/Gaia/DPAC, adapted from the movie produced by L. Rohrbasser, K. Nienartowicz, O.L. Creevey, L. Eyer, C. Reyle.

photocenter displacements due to huge convective cells in asymptotic giant stars which give rise to a "scattered" position and thus large uncertainties on the astrometric parameters (Chiavassa et al. 2022), or regular repetitive 'circular' patterns due to two or more gravitationally bound objects whose positions are constantly changing, but over a regular time period, see e.g. this video illustrating the sky projected motion of a binary star. Both variability and multiplicity can also be detected in radial velocity measurements.

Two coordination units within Gaia DPAC are dedicated precisely to the study of variability and multiplicity, and they are the specific subjects of other lectures in this school, so details will be left to Eyer (2024); Arenou (2024). Here we just summarise that all variable objects are found in the Gaia archive in the tables beginning with `vari`, and the `gaiadr3.vari_summary` table summarises the statistics of the time series of all variable sources along with flags indicating in which variable table the source is found, e.g. source 4071605906863950848 is in both `gaiadr3.vari_short_timescale` and `gaiadr3.vari_eclipsing_binary`. The user will then find further details about that source in those tables. While all of the multiple sources are found in the *non-single star* tables, which are those that begin with `gaiadr3.nss`. Four such tables are available, e.g. the `gaiadr3.nss_two_body_orbit` provides a parallax solution for the two components of binary systems which considers the astrometric motion of both stars in the system. The astrometry for these sources should, in general, be preferred over the standard astrometric solution given for a single star in `gaiadr3.gaia_source`. Users are referred to Eyer (2024) and Arenou (2024) for lectures and examples in this school on both of these topics.

4 Conclusions

Gaia DR3 contains a wealth of new data products that can be exploited by the stellar physics community. Starting with the basic data products described in Sect. 2 — (mean and time-series) photometry, high- and low- resolution spectra, and astrometry — the user can derive many physical parameters of stars and their environments. A homogenous analysis of the full Gaia data set is provided in Gaia DR3, based on the time-series and mean data products and this is described in Sect. 3. A general description of the Apsis chain is given, which is responsible for the parametrization of stars and other sources based on the mean data products, while details of the variability and multiplicity analysis is described in Arenou (2024) and Eyer (2024). We then visited different regions of the HR diagram, discussed the open questions, and highlighted how Gaia DR3 can help address these specific questions in stellar physics. Concrete examples of scientific exploitation of GDR3 for the specific regions across the HR diagram was also illustrated. Example queries to access the data are available

in the accompanying tutorials and slides, available at the website of the Ecole Evry Schatzman 2023[3].

The author would like to warmly thank Anthony Brown, Francesca De Angeli, and David Katz for providing supporting material which was used to prepare part of this lecture, and the organisors of the *Ecole Evry Schatzman 2023*, in particular, Carine Babusiaux. The author would also like to thank Gaia Coordination Unit 8 for their hard work over many years that resulted in this wonderful catalogue of astrophysical parameters, and to everyone involved in making Gaia possible, from operations, to ground-processing, archive preparation and validation. This work has made use of data from the European Space Agency (ESA) mission *Gaia* (https://www.cosmos.esa.int/gaia), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, https://www.cosmos.esa.int/web/gaia/dpac/consortium). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement.

References

Andrae, R., Fouesneau, M., Sordo, R., et al. 2023, A&A, 674, A27
Arenou, F. 2024, EES 2023, These proceedings
Babusiaux, C., Fabricius, C., Khanna, S., et al. 2023, A&A, 674, A32
Brown, A. G. A. 2021, ARA&A, 59, 59
Buder, S., Sharma, S., Kos, J., et al. 2021, MNRAS, 506, 150
Casagrande, L., Schönrich, R., Asplund, M., et al. 2011, A&A, 530, A138
Casagrande, L. & VandenBerg, D. A. 2018, MNRAS, 479, L102
Chiavassa, A., Kudritzki, R., Davies, B., Freytag, B., & de Mink, S. E. 2022, A&A, 661, L1
Clementini, G., Ripepi, V., Garofalo, A., et al. 2023, A&A, 674, A18
Creevey, O. L. & Lebreton, Y. 2022, DPAC Public Documents
Creevey, O. L., Sordo, R., Pailler, F., et al. 2023, A&A, 674, A26
Cutri, R. M., Skrutskie, M. F., van Dyk, S., et al. 2003, VizieR Online Data Catalog, II/246
De Angeli, F., Weiler, M., Montegriffo, P., et al. 2023, A&A, 674, A2
Delchambre, L., Bailer-Jones, C. A. L., Bellas-Velidis, I., et al. 2023, A&A, 674, A31
ESA, E. 1997, 1200
Eyer, L. 2024, EES 2023, These proceedings
Fouesneau, M., Frémat, Y., Andrae, R., et al. 2023, A&A, 674, A28
Frémat, Y. 2024, EES 2023, These proceedings
Frémat, Y., Royer, F., Marchal, O., et al. 2023, A&A, 674, A8
Gaia Collaboration, Bailer-Jones, C. A. L., Teyssier, D., et al. 2023a, A&A, 674, A41
Gaia Collaboration, Galluccio, L., Delbo, M., et al. 2023b, A&A, 674, A35
Gaia Collaboration, Schultheis, M., Zhao, H., et al. 2023c, A&A, 674, A40
Gaia Collaboration, Babusiaux, C., van Leeuwen, F., Barstow, M. A., et al. 2018, A&A, 616, A10
Gaia Collaboration, Brown, A. G. A., Vallenari, A., Prusti, T., et al. 2021, A&A, 649, A1
Gaia Collaboration, Creevey, O. L., Sarro, L. M., Lobel, A., et al. 2023, A&A, 674, A39
Gaia Collaboration, Creevey, O. L., Sarro, L. M., Lobel, A., et al. 2022, VizieR Online Data Catalog, J/A+A/674/A39
Gaia Collaboration, De Ridder, J., Ripepi, V., Aerts, C., et al. 2023, A&A, 674, A36
Gaia Collaboration, Drimmel, R., Romero-Gómez, M., Chemin, L., et al. 2023, A&A, 674, A37
Gaia Collaboration, Montegriffo, P., Bellazzini, M., De Angeli, F., et al. 2023, A&A, 674, A33
Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., Brown, A. G. A., et al. 2016, A&A, 595, A1
Gaia Collaboration, Vallenari, A., Brown, A. G. A., Prusti, T., et al. 2023, A&A, 674, A1
Guillot, T. & Gautier, D. 2015, in Treatise on Geophysics, ed. G. Schubert, 529–557
Hidalgo, S. L., Pietrinferni, A., Cassisi, S., et al. 2018, ApJ, 856, 125
Høg, E. 2024, arXiv e-prints, arXiv:2402.10996
Katz, D., Sartoretti, P., Guerrier, A., et al. 2023, A&A, 674, A5
Lanzafame, A. C., Brugaletta, E., Frémat, Y., et al. 2023, A&A, 674, A30
Lindegren, L., Bastian, U., Biermann, M., et al. 2021a, A&A, 649, A4
Lindegren, L., Klioner, S. A., Hernández, J., et al. 2021b, A&A, 649, A2

[3]https://ees2023.sciencesconf.org/

Montegriffo, P., De Angeli, F., Andrae, R., et al. 2023, A&A, 674, A3
Pinsonneault, M. H., Elsworth, Y. P., Tayar, J., et al. 2018, ApJS, 239, 32
Recio-Blanco, A. 2024, EES 2023, These proceedings
Recio-Blanco, A., de Laverny, P., Palicio, P. A., et al. 2023, A&A, 674, A29
Réyle, C. 2024, EES 2023, These proceedings
Riello, M., De Angeli, F., Evans, D. W., et al. 2021, A&A, 649, A3
Sarro, L. M., Berihuete, A., Smart, R. L., et al. 2023, A&A, 669, A139
Sartoretti, P., Katz, D., Cropper, M., et al. 2018, A&A, 616, A6
Sartoretti, P., Marchal, O., Babusiaux, C., et al. 2023, A&A, 674, A6
Steinmetz, M., Guiglion, G., McMillan, P. J., et al. 2020, AJ, 160, 83
Zeng, L., Sasselov, D. D., & Jacobsen, S. B. 2016, ApJ, 819, 127

École Evry Schatzman 2023: Stellar physics with Gaia
C. Babusiaux, C. Reylé (eds)

GAIA STELLAR POPULATIONS

A. RECIO-BLANCO

Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, Bd de l'Observatoire, CS 34229, 06304 Nice Cedex 4, France

e-mail: alejandra.recio-blanco@oca.eu

The slides presented during the school are available at https://ees2023.sciencesconf.org. They present:

École Evry Schatzman 2023: Stellar physics with Gaia
C. Babusiaux, C. Reylé (eds)

STELLAR MULTIPLICITY SEEN FROM GAIA

F. ARENOU

GEPI, Observatoire de Paris, Université PSL, CNRS, 5 Place Jules Janssen, 92 190 Meudon, France
e-mail: Frederic.Arenou@obspm.fr

The slides presented during the school are available at https://ees2023.sciencesconf.org. They present:

The tutorial associated to this lecture is available as a python notebook at https://ees2023.sciencesconf.org.

École Evry Schatzman 2023: Stellar physics with Gaia
C. Babusiaux, C. Reylé (eds)

THE VARIABLE SOURCES IN THE GAIA ARCHIVE

L. EYER

Département d'Astronomie de l'Université de Genève, Chemin Pegasi 51, 1290 Versoix, Swizterland
e-mail: laurent.eyer@unige.ch

Abstract. At the core of the Gaia mission is a multi-epoch survey consisting of astrometric, photometric, spectrophotometric, and spectroscopic measurements. The astrometric time series provides parallaxes and proper motions, along with information on astrometric binary systems. The photometric time series offers a means to investigate the variability of the sources. Due to their whole-sky, multi-epoch nature, multiple instruments, their magnitude range covering 21 magnitudes, and their remarkable photometric precision, these data allow us to describe the variability of celestial phenomena in an unprecedented manner. For the third Gaia Data Release (DR3), the data collection spanned 34 months, with a median number of field-of-view measurements in the G band of about 44, reaching up to 270. At publication time, DR3 delivered the largest collection of variable sources with an associated classification across the entire sky. All these sources have their G, G_{BP}, G_{RP} epoch data published and accessible in the Gaia ESA archive. In summary, there are 10.5 million variable sources, including 9.5 million variable stars and 1 million QSOs. Additionally, 2.5 million galaxies were identified thanks to spurious variability caused by the non-axisymmetric nature of galaxies and the way Gaia collects data. Moreover, all the epoch data and time series of nearly 1.3 million sources in a pencil beam around the Andromeda galaxy are published, regardless of their status (constant or variable); This dataset is known as the Gaia Andromeda Photometric Survey. We also introduce the citizen science project, GaiaVari to classify variable stars, the Focused Product Release delivered on October 10, 2023. In the future, DR4 will cover 66 months, and we hope DR5 will have accumulated 10.5 years of data.

Keywords: stars: variables: general - Galaxy: stellar content - catalogues - stars: oscillations - binaries: eclipsing - starspots

1 Introduction

In astronomy, from its beginnings to now, the primary means of learning about the Universe and its constituents has been through the photons. We can find the following text[1] by Evry Schatzman in his book "Astrophysique" Schatzman (1963):

> *The fundamental basis of astrophysics lies in the radiation that reaches us from celestial objects. Whether in the visible or invisible spectrum, it is always a form of radiation that we measure.*

Only from the late 20th century were alternative methods used to learn about our Universe, first as exceptions to the sentence above, such as the cosmic rays (Hess 1912), then with important implications, such as the neutrinos (from the Sun, Davis et al. 1968), and the gravitational waves in 2015 (Abbott et al. 2016).

The classical subdivision in observational astronomy about the three main observation branches are:

- Astrometry: Knowledge of celestial objects obtained from their positions, motions, and shapes. Probably this discipline is as old as humanity.
- Photometry: Knowledge of celestial objects obtained from the measurement of their integrated light over a range of wavelengths. Historically, photometry went hand in hand with astrometry as they are the most basic perceptions of the human eye. The magnitude scale, as it is well known, is inherited from the ancient greeks (and formalized by Norman Pogson, cf. Pogson 1856).

[1] Translated from French.

- Spectroscopy: Knowledge of celestial objects obtained from their spectra. This technique emerged in the 19th century, with Joseph von Fraunhofer discovering dark lines in the spectra of the Sun and other stars, and Gustav Kirchhoff and Robert Bunsen correctly interpreting these dark lines as being caused by specific atomic elements (for a brief historical account see Appenzeller 2013).

It should be remarked that we gain knowledge not only from the source but also from what is between the source and the observer, e.g. the interstellar medium. Observations from the ground can also allow us to determine the properties of the atmosphere and its evolution, such as in Burki et al. (1995).

Each of these three main observation branches benefits from multi-epoch observations. The most knowledge is obtained when these different multi-epoch domains are combined. Such merging is fruitful for finding and calibrating standard candles, identifying stellar populations, probing the invisible with microlensing events (e.g. merging the photometry and the astrometry as in Sahu et al. 2022) and determining astrophysical parameters thanks to binaries (Eyer et al. 2015), to pulsating stars (Baade-Wesselink method, Carney et al. 1992), etc.

It should be noted that on the side of photometric surveys, there has been a booming period for these past 30 years with OGLE (Udalski 2003), HAT (Bakos et al. 2004), Kepler (Borucki et al. 2010), and ZTF (Bellm et al. 2019), to name just a few. The future is also bright with the LSST (Ivezić et al. 2019) and PLATO (Rauer et al. 2014), again to name just a few. Spectroscopy is entering into the era of systematic multi-epoch global surveys, obviously with Gaia, but also with 4MOST (de Jong et al. 2019), and in the future with WST (Mainieri et al. 2024).

Now, if we turn to the Gaia cornerstone mission, the "tour de force" of Gaia is to assemble multi-epoch data for the entire sky on these three main observation branches: astrometry, photometry, and spectroscopy.

In this short text, we will mostly restrict the presentation to what has been put in the Gaia archive for the variable sources.

2 The Gaia mission: An introduction

Let us review key aspects of the Gaia mission: Gaia is a cornerstone mission of the European Space Agency science program (Gaia Collaboration, Prusti et al. 2016), tasked with performing a systematic survey of all objects brighter than G=20.7, attaining a survey of more than 2 billion celestial sources. The measurements collected by Gaia consist of position (astrometry), brightness, and colours (photometry and spectrophotometry), along with radial velocity (spectroscopy). It should be noted that the survey for radial velocities is restricted to brighter sources (about $G = 15$). Gaia was launched on December 19, 2013 by a Soyuz rocket. The CCD camera aboard Gaia is the largest ever deployed in space, reaching nearly one billion pixels. Over its ten-year mission, each source, on average, will be observed 140 times in each of the 9 CCDs[2] in the G band, as well as in each of the BP and RP CCDs. The number of measurements for the radial velocity spectrometer instrument will be comparatively fewer, averaging approximately 80 times. The results of the Gaia mission, that the Data Processing and Analysis Consortium processes, are made accessible to the public through Data Releases. DR4 and DR5 are anticipated to be delivered in 2026 and no later than 2032, respectively.

We will focus on this article on the DR3 results, which compiles 34 months of observations and nearly one trillion CCD measurements over the whole celestial sphere.

Several factors contribute to Gaia's exceptional status as a mission:

- Unprecedented Astrometry: Gaia provides unparalleled astrometric data, i.e. positions, parallaxes, and proper motions for more than a billion stars over the whole celestial sphere. The astrometric precision is available on the Gaia webpage (`https://www.cosmos.esa.int/web/gaia`[3]): For Gaia DR4, that is, for the nominal mission, the parallax uncertainty is 22 μas at magnitude 15. This precision can be compared to the 20-25 μas estimates from pre-launch assessments in 2006 for a G2 star at magnitude 15 (see e.g. Eyer 2006). It's important to note that the errors in position and proper motion need to be multiplied by 0.8 and 0.5, respectively. The astrometry is also able to detect astrometric binaries, exoplanets, microlensing effects, etc…

- Three Instruments on a Single Platform: astrometry, [spectro-]photometry, spectroscopy with the derivation of radial velocities. This results in a unprecedented uniformity of data across the entire celestial sphere.

[2] for each of the 7 raws at the exception of raw 5, for which the number is 8.
[3] Note: This page has undergone changes in URL and error estimates following the better understanding of the performance.

- Multi-Epoch Measurements: The mission conducts multi-epoch measurements of the entire sky, providing a comprehensive view of variable celestial objects, see also Section 3.
- 10-Year Time Baseline: With a 10-year mission duration (if all goes as planned), Gaia offers high frequency precision, particularly beneficial for periodic objects.
- Quasi-Simultaneous Measurements: Gaia achieves quasi-simultaneous measurements. Other multi-band photometric surveys do not have this simultaneity and this complicates somewhat the analysis. For more details, see Section 3.
- Time-Domain Selection Function: Because of its predefined scanning law, Gaia allows us to determine the time-domain selection function, enabling the understanding of observational biases.
- Extensive Number of Measurements: Gaia's data volume is significant, with nearly 1 trillion CCD measurements for DR3 (i.e. 34 months).
- High Dynamical Range: Gaia covers a high dynamical range, capturing data from the brightest sources ($G \approx 1.7$ in DR3) to magnitudes as faint as $G \approx 21/22$.
- Space-based Operation: Gaia's position in space provides stability and allows us to access the entire celestial sphere from a single platform, which is not achievable by any single ground-based optical telescope. However, operating in space presents its own challenges, such as the high cost of space missions[4], the impossibility of repairs, micrometeoroid hits, and solar flares/the radiation damages. Indeed Gaia's location at L2, 1.5 million kilometers from Earth in the direction opposite to the Sun, leaves it unshielded by Earth's magnetic field.
- Cyclic Improvements: The Gaia consortium implements cyclic improvements, performing systematic data analyses, with not only more data at each cycle, but with improved calibrations, and enhanced outlier detection.

Each of these elements makes Gaia extremely fruitful for variable and binary stars. These are among the objects that benefit the most from Gaia. Due to the semi-regular Gaia sampling, strictly periodic objects can be studied in detail despite the gaps in the scanning law. Also transient sources are and will be classified in the data releases (Eyer et al. 2023; Rimoldini et al. 2023) and are also detected in real time by the Gaia Science Alert System (Hodgkin et al. 2021). Indeed, since the beginning of its science operations, Gaia has continuously (with few interruptions) provided alerts, identifying potentially time-sensitive events. If not addressed promptly, these events could lead to significant scientific loss. To date (June 2024), the Science Alerts team has released over 25,000 of these alerts. About 27% of the alerts are classified, with approximately 60% of these classified alerts identified as supernovae.

3 The Gaia time sampling

The time sampling[5] of Gaia is very particular and semi-regular. It is optimized to achieve astrometric precision as uniform as possible across the entire celestial sphere. This sampling is known as the Nominal Scanning Law (NSL): Gaia continuously sweeps the sky, with the spacecraft's rotation axis precessing on a cone with a 45-degree opening angle to the Sun. A full precession cycle on the cone is completed every 63 days. The spacecraft rotates on its axis every 6 hours, and the 106.5-degree angle between Gaia's two fields-of-view (preceding and following fields-of-view) results in a time interval of 1 hour and 46 minutes between them (the exemplary angle of 253.5-degree between the following and preceding fields-of-view takes thus 4 hours 14 minutes).

Figure 1 shows the sky plot for the number of observations (the field-of-view transit). We took the field `matched_transits`, i.e. the total number of field-of-view transits matched to this source. This number is higher or equal to the `num_selected_g_fov` of `gaiadr3.vari_summary` table, because it does not take into account the number of observations that are removed.

The second plot, Fig. 2, presents the number of CCD observations as a function of ecliptic latitude. On a field crossing the CCD observations are separated by 4.85 seconds. The Annex A of Eyer et al. (2017) provides

[4] As Prof. Bohdan Paczynski humorously noted, "One dollar in space is worth less than one dollar on Earth".
[5] referred to as the cadence in LSST/Vera C. Rubin observatory.

detailed properties of the scanning law for 5 years, including various representations, histograms of the lags, and spectral windows for different regions of the sky.

Because the scanning law is well-defined, it is possible to study the selection function. Eyer & Mignard (2005) conducted a study on periodic signals based on an earlier design of the spacecraft and a 5-year scanning law. They concluded that the ability to detect the correct period of a periodic signal strongly depends on the ecliptic latitude. For DR3, significant differences in detection exist even at fixed ecliptic latitudes due to the scanning law creating regions with a low number of observations; see Fig. 1.

There are two exceptions to the Nominal Scanning Law:

1. At the beginning of the mission, a 28-day period was planned to assist with photometric calibrations. During this time, the same stars and regions of the sky were scanned nearly continuously. The axis of rotation of the spacecraft was kept in the ecliptic plane (at a 45-degree angle to the Sun) without precession, allowing the ecliptic poles to be scanned regularly during these 28 days, with repeated sequences of intervals 1h46m, 4h14m, 1h46m, etc. As the spacecraft orbits around the Sun, stars at lower ecliptic latitudes (in absolute value) are observed for shorter sequences.

2. The second exception occurred when, for 12 months[6], a modified NSL was introduced to improve the astrometric solution. As a positive side effect it reduces aliasing peaks. In this mode, the precession of the spacecraft's rotation axis was reversed. However, the spacecraft's rotation direction remained unchanged due to the constraints of the Time Delay Integration (TDI) mode. Note that the reverse motion of the rotation axis on the cone increased the consumption of cold gas.

Although the scanning law determines which stars are observed, understanding the temporary absence of an observed source is challenging. Several factors could explain a missing observation: the star might have truly disappeared (the most interesting case) by becoming very faint due to an eclipse or occultation, it could be just below the detection threshold of the less precise sky mapper and thus not observed, the star can be perturbed by the other field-of-view (e.g. bright star/extended object) which is superposed on the same focal plane, the data might have been deleted due to excessive data volume onboard, or technical problems might have occurred.

4 The Gaia G-band photometric precision

Fig. 3 shows the uncertainty in the mean G-band photometry as a function of the G magnitude. There are two plots: the results from DR1 and DR3. DR2 is not displayed to avoid too many figures and it falls between the two. The plot is based on a selection of stars with absolute ecliptic latitudes below 60 degrees to avoid the perturbation of ecliptic pole scanning. The uncertainty is estimated from the flux standard deviation divided by the square root of the number of measurements (after some clipping). Naturally, this uncertainty decreases as the number of measurements increases. In case of variable sources, the standard deviation is obviously larger. Although more measurements may reveal more instrumental calibration features, this is not the case here. On the contrary, some features visible in DR1 are less prominent or have disappeared in DR3 demonstrating the improvements of the calibrations. It should be noted that in the Gaia archive, uncertainties are provided in flux rather than in magnitude. Therefore, these uncertainties need to be converted into magnitude uncertainties for Fig. 3. We used the formula $\text{Uncertainty(magnitude)} = 1.086 \times \text{Uncertainty(flux)}/\text{flux}$, where the term 1.086 is coming from $2.5/\ln(10)$.

We will not enter into the debate on using flux versus magnitudes to evaluate the brightness of a celestial source here, opening Pandora's box. The main point is to avoid being too dogmatic about these issues and to use the approach that makes the most sense for the problem at hand. As a side note, negative fluxes were unfortunately excluded in the mean calculation process (Riello et al. 2021; Evans et al. 2023). Negative fluxes can occur due to statistical variations or overestimation of the background for example. This exclusion introduces a bias in the magnitude estimation at the faint end: the sources are brighter than they are in reality. Starting from DR4, negative fluxes will be included.

5 The Gaia data releases and the variability

The data release process is iterative and advances with each step in both quantity and quality, this is particularly evident for the variability processing and analysis. So let us have a quick look at the DR1 and DR2 data releases.

[6] 6 months will be included in DR4, the rest in DR5.

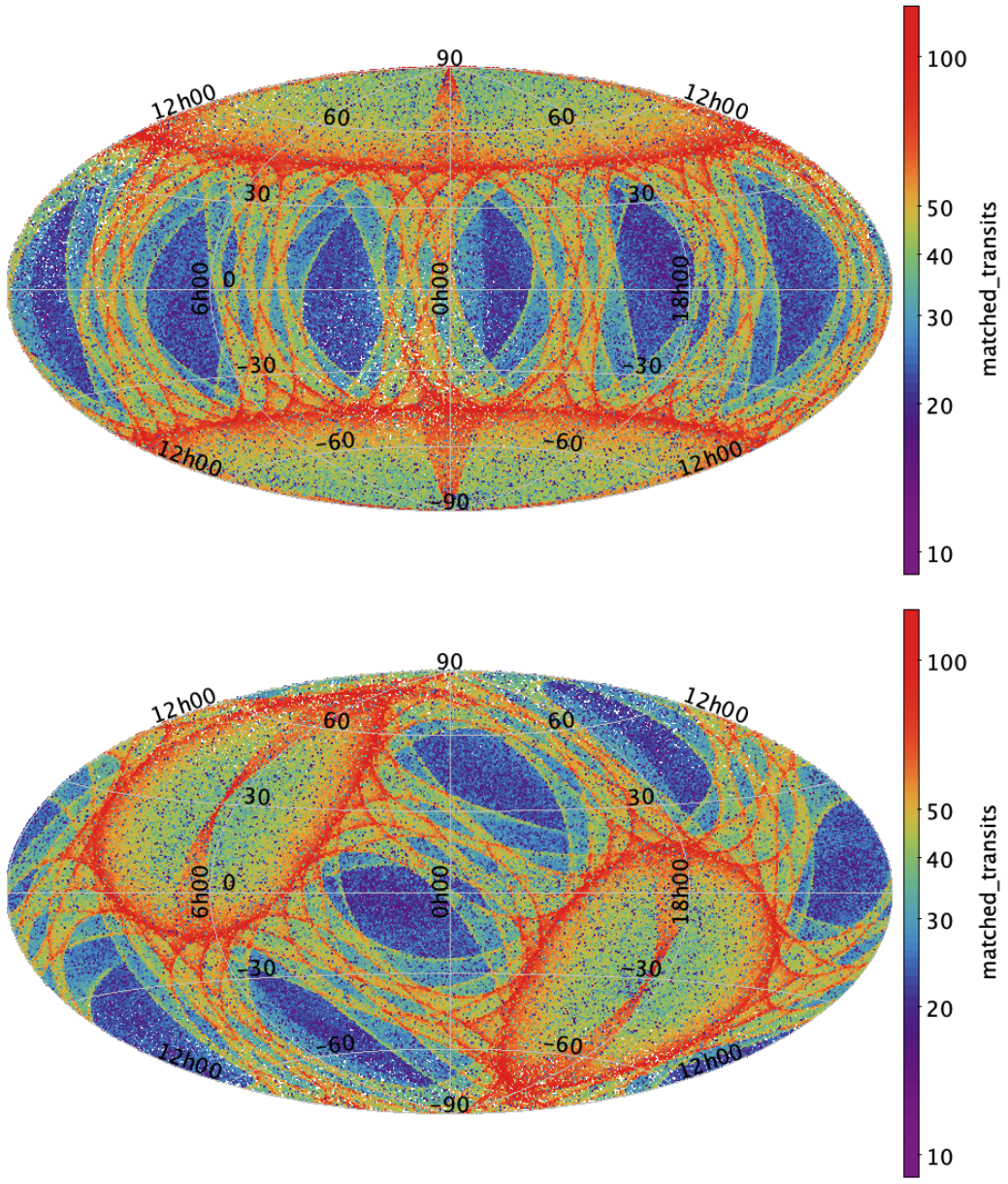


Fig. 1. Number of fields-of-view data per source in ecliptic coordinates (top) and Galactic coordinates (bottom) for DR3 in Aitoff projections. We see the complex sampling which is clearly not completed yet, with blue regions, i.e. low number of transtis, around the ecliptic plane. We see that one of these low number of measurements is in the direction of the lower bulge at positive Galactic longitudes.

A global pipeline was set up for the variability processing and analysis of Gaia data. Initially, it was applied to the cycle 1 data, the details of this processing can be found in Eyer et al. (2017). The DR1 publication also covers the data properties of cycle 1, along with certain aspects of the 5-year sampling obtained through simulations. Based on 14 months of data, the DR1 variable star catalogue showcased a selection of 3,194 variable stars. This data release consisted of two types of variability, namely Cepheids and RR Lyrae stars in the direction of the Large Magellanic Cloud (Clementini et al. 2016). The Large Magellanic Cloud, situated near the South Ecliptic pole, benefited from the regular sampling during the 28 days of the Ecliptic Pole Scanning Law. In addition, a performance verification paper (Gaia Collaboration, Clementini et al. 2017) presented the TGAS parallaxes for Cepheids and RR Lyrae stars, demonstrating that already the improvements with respect to Hipparcos were

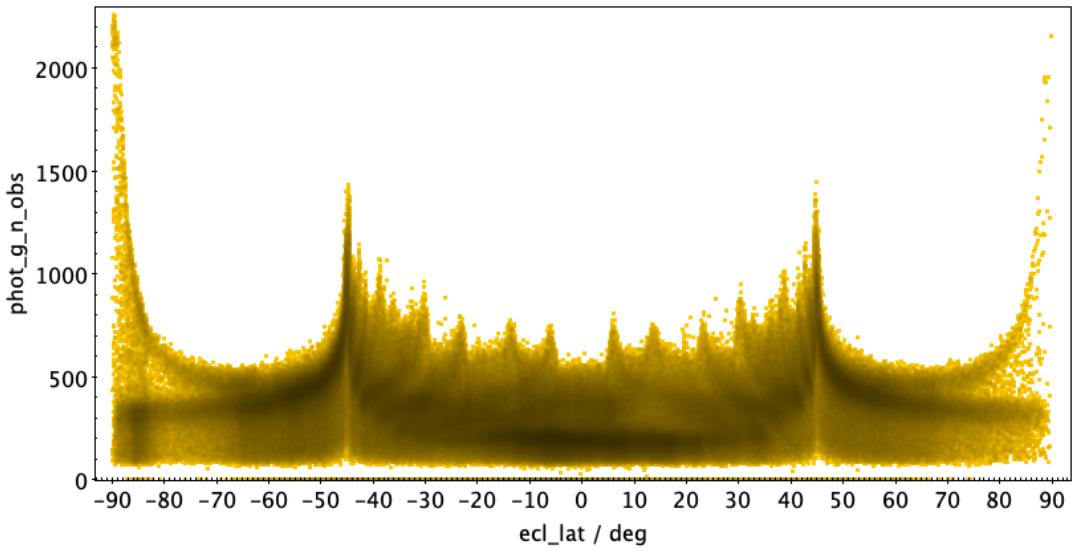


Fig. 2. Number of CCD observations per source as a function of ecliptic latitude (β) for DR3. A large number of observations are visible at $\pm 45°$ due to the Nominal Scanning Law. Additionally, stars at the ecliptic poles (both ends of the plot) have an exceptionally high number of observations, exceeding 2000, due to the Ecliptic Pole Scanning Law, which was in effect for the first 28 days of the mission.

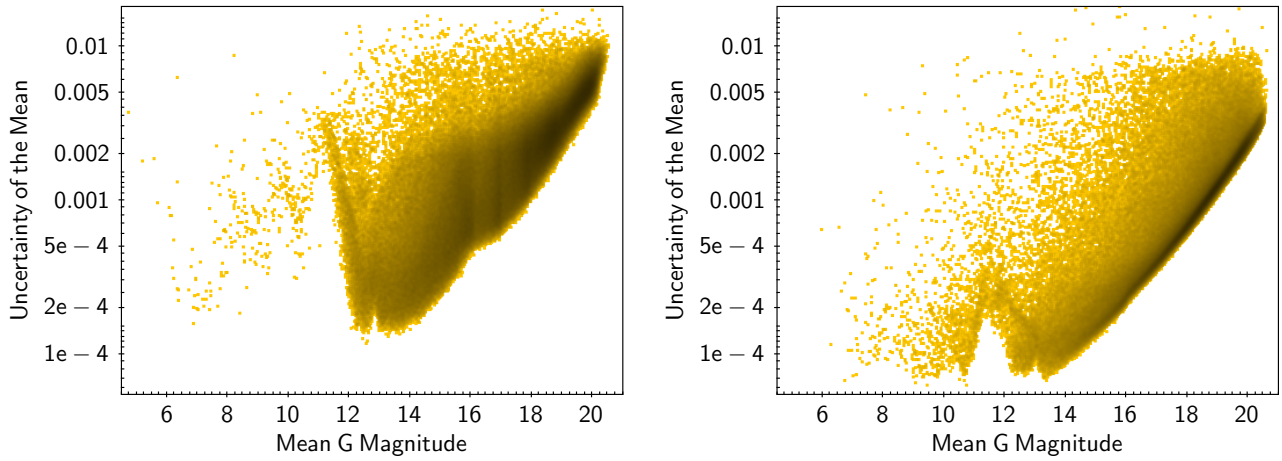


Fig. 3. Uncertainty on the mean versus the mean G magnitude. Two plots are displayed side by side, for the comparison of DR1 (left) and DR3 (right) photometry. We see that not only has the noise gone down thanks to the square root of the number of measurements but also that the calibration has significantly improved. The various changes in windowing schemes or gates are better taken into account in DR3 than in DR1.

impressive.

The DR2 based on 22 months of operation data reached 550,737 variable stars (Holl et al. 2018). The DR2 contained the following variability types: RR Lyrae stars, Cepheids, long period variables, δ Scuti/SX Phoenicis stars, stars with rotation modulation and stars with short time scale variability. A performance verification paper (Gaia Collaboration, Eyer et al. 2019) presented the properties of variable sources across the Hertzsprung-Russel Diagram. As a remarkable result, we can follow the path of stars, their motion in the Hertzsprung-Russel Diagram as their magnitude and colour change[7].

More generally, information and details for all data releases can be found in various articles (many are listed in this section for details on variability processing and analysis) or in the Gaia documentation associated with

[7] https://www.cosmos.esa.int/web/gaia/gaiadr2_cu7

the data releases on the archive website, https://gea.esac.esa.int/archive/documentation/GDR3/ for DR3. This document, available online and also in PDF format, contains more than 1,500 pages.

It is also important to consult the "known issues" page. For each data release, there is a section on the Data tab of https://www.cosmos.esa.int/web/gaia/data. Under these sections, you can find the known issues for each data release. For example, the known issues for DR3 can be found at the following address: https://www.cosmos.esa.int/web/gaia/dr3-known-issues.

A note on the archive names: you have in the ESA Gaia archive (https://gea.esac.esa.int/archive/) the different data sets associated with the different data releases, `gaiadr1.table_name`, `gaiadr2.table_name`, `gaiaedr3.table_name`, `gaiadr3.table_name`, `gaiafpr.table_name` which relates the Gaia Data Release 1 (DR1), Gaia Data Release 2 (DR2), Gaia Early Data Release 3 (EDR3), Gaia Data Release 3 (DR3), Focused Product Release (FPR) respectively. For a given source, the source_id can change. A last remark on the nomenclature related to variability: for the tables names associated to the variability processing and analysis, starting from the second data release, these table names start with "`vari_table_name`".

In the following sections, we will focus more on the DR3 as it has the most numerous outputs and diversity in variability types.

5.1 The DR3 results

For the third data release of Gaia, the variability analysis is described in Eyer et al. (2023). In summary, as of the data release in June 2021, Gaia provided the largest catalogue ever of classified variable sources across the entire sky. In DR3, we have 10.5 million variable sources, comprising 9.5 million variable stars and 1 million AGN. Additionally, thanks to a spurious perturbation in the photometry, we identified 2.5 million galaxies.

A general overview of the DR3 result, as mentioned above, can be found in Eyer et al. (2023). In the Gaia archive https://gea.esac.esa.int/archive/, the different associated tables related to the variability processing and analysis can be seen in Fig. 4.

Fig. 4. Tables available related to the variability processing and analysis in the Gaia archive. We opened the `vari_summary` table with just the first three lines.

Variability types	Variability types
(1) ACV|CP|MCP|ROAM|ROAP|SXARI: α2 CVn, (Magnetic) Chemical Peculiar, Rapidly Oscillating Am/Ap[8] SX Arietis stars	(13) MICROLENSING: Star with microlensing event
(2) ACYG: Alpha Cygni-type variable	(14) RCB: R Coronae Borealis stars
(3) AGN: Active Galactic Nuclei (including Quasars)	(15) RR: RR Lyrae stars
(4) BCEP: Beta Cephei type variable	(16) RS: RS Canum Venaticorum type variable
(5) BE|GCAS|SDOR|WR: Subset of eruptive variable types: B-type emission line star, Gamma Cassiopeiae, S Doradus, and Wolf-Rayet	(17) S: Set of stars with short timescale variability
(6) CEP: Cepheid variable types: delta Cepheid, anomalous Cepheid, and type-II Cepheid	(18) SDB: Subdwarf B stars
(7) CV: Cataclysmic variable	(19) SPB: Slowly Pulsating B-star variable
(8) DSCT|GDOR|SXPHE: delta Scuti, gamma Doradus, and SX Phoenicis	(20) SN: Supernovae
(9) ECL: Eclipsing Binaries of types: Beta Persei (Algol), Beta Lyrae, and W Ursae Majoris	(21) SOLAR_LIKE: Stars with solar-like variability due to flares, spots, and rotational modulation
(10) ELL: Ellipsoidal variable	(22) SYST: Symbiotic variable star
(11) EP: exoplanetary transits	(23) WD: Variable White Dwarf of types: ZZ Ceti (DAV, ZZA), V777 Her (DBV, ZZB), and GW Vir (DOV, ZZO)
(12) LPV: Long Period Variable stars of types: omicron Ceti (Mira), OGLE Small Amplitude Red Giants, and semiregular	(24) YSO: Young Stellar Object

Table 1. The 24 variability types considered in DR3. Please note that in some cases, a denomination encompasses several different variability types.

A vast and systematic compilation of the literature was done in Gavras et al. (2023); it combines the sources of 152 catalogues crossmatched with the Gaia data and contains 4.9 million variable sources. The catalogue is available at the CDS (Centre de Données astronomiques de Strasbourg, `https://cds.unistra.fr`, Gavras et al. (2022)).

Part of this compilation was used as a training set for the supervised classification (Rimoldini et al. 2023). For the classification, the methods employed are Random Forest and eXtreme Gradient Boosting models in a mixture of multi-class and binary classifiers. There are 24 variability (grouped) classes. Some of these are combined types. Therefore, when counting the subclasses, we reach a total of 35 variability types. The variability classes list can be found in Table 1.

The table that contains the variability type class from the Machine Learning is `gaiadr3.vari_classifier_result`.

A validation of the classification for the young stellar object class was done by Marton et al. (2023).

After the classification, Specific Object Studies are done for various topics. They act as a validation of the classification, and also these activities compute further parameters that describe the class. It is important to note that there are two sources for a given variability type: (1) The Machine Learning classification and (2) the classification of Specific Object Studies. There may be differences between these two classifications.

The Specific Object Studies covers the following topics (listed in the alphabetical order, as in the Gaia archive). Here we just give the number of sources and also some selected highlights:

- **AGN:** `gaiadr3.vari_agn` Carnerero et al. (2023). There are 872,228 sources in this table, this number is obtained with the ADQL query (just shown for this example):

  ```
  SELECT COUNT(*)
  FROM gaiadr3.vari_agn
  ```

  We can compare this number from the number of AGN from classification, we have 1,035,207 sources, see following ADQL command:

[8] here probably the small oscillations are not detected.

```
SELECT COUNT(*)
FROM gaiadr3.vari_classifier_result
WHERE best_class_name = 'AGN'
```

A surprising result is that even with the gaps of the Gaia time series, the time delay for a source, the lens system DESJ0501-4118, was derived.

- **Cepheids:** `gaiadr3.vari_cepheid` Ripepi et al. (2023). There are 15,021 sources. A Fourier series model of the light curve is fitted and the parameters are published in this table. The estimation of the mean should be taken for this table and not from `gaia_source` table. A subclassification is also given. Radial velocities time series are published for 798 sources. It is the largest catalogue of Cepheids having radial velocities.
- **Compact companion**: `gaiadr3.vari_compact_companion` Gomel et al. (2023). There are 6,306 sources. If the variability is caused by ellipsoidal deformation, a mass ratio can be determined from the Fourier parameters of the light curve. This mass ratio helps identify potential candidates for white dwarfs, neutron stars, and black holes. Other methods to detect black holes are based on astrometry and radial velocities (El-Badry et al. 2023; Chakrabarti et al. 2023; El-Badry et al. 2023; Gaia Collaboration, Panuzzo et al. 2024).
- **Eclipsing binaries:** `gaiadr3.vari_eclipsing_binary` Mowlavi et al. (2023). There are 2,184,477 sources. The largest catalogue of eclipsing binaries ever published on the whole sky. In this table, we adopt a geometrical modelling of the light curve and extract some relevant statistics. A sub-sample is then studied with a physical modelling, see Pourbaix et al. (2022). The results are published in the non-single star table `gaiadr3.nss_two_body_orbit`. It contains 86,918 systems with the flag nss_solution_type EclipsingBinary and 155 for EclipsingSpectro.
- **Long Period Variables:** `gaiadr3.vari_long_period_variable` Lebzelter et al. (2023). There are 1,720,588 sources. Thanks to the RP spectrophotometry, 546,468 stars were classified as carbon rich candidates.
- **Microlensing:** `gaiadr3.vari_microlensing` Wyrzykowski et al. (2023). This is the first whole sky microlensing event list, which contains 363 sources.
- **Main Sequence Oscillators:** `gaiadr3.vari_ms_oscillator` Gaia Collaboration, De Ridder et al. (2023). There are 54,476 sources. See Section 5.1.1.
- **Planetary transits:** `gaiadr3.vari_planetary_transit` Panahi et al. (2022b). The table contains 214 sources. Please note that the table `gaiadr3.vari_planetary_transit_13june2022` is incorrect. For details, please see Section 5.1.3.
- **Rotation modulation:** `gaiadr3.vari_rotation_modulation` Distefano et al. (2023). There are 474,026 sources. The Period Amplitude diagram shows different regimes.
- **RR Lyrae stars:** `gaiadr3.vari_rrlyrae` is taking care of the RR Lyrae variables and is described in Clementini et al. (2023). There are 271,779 sources. RR Lyrae stars are well known standard candles. There are several subtypes of them: the Bailey's ab and c types corresponding to fundamental mode pulsation and first overtone, respectively. The ab type light curves have a very recognisable shape with a sharp rise and slower declined with a period typically of half a day, the c types have a more sinusoidal light curve. An example of the quite astonishing results is the metallicity map derived from the RR Lyrae light curves in Figure 5. There are 1,100 sources with radial velocities time series which are in the table of RR Lyrae stars.
- **Short Time scales:** `gaiadr3.vari_short_timescale` is described in Eyer et al. (2023). There are 471,679 sources. This procedure did not go through removal of sources and reveals also instrumental and/or calibration artefacts.

For the completeness of the archive description, three additional tables are listed under the variability folder: one is a summary of the variability analysis, and the other two are associated with the radial velocities of Cepheids and RR Lyrae stars:

- `gaiadr3.vari_summary` table provides a global overview. It helps determine which table contains a given source (Eyer et al. 2023), and provides statistical attributes of the time series.
- `gaiadr3.vari_epoch_radial_velocity`: the table contains the time series of the radial velocities.
- `gaiadr3.vari_rad_vel_statistics`: the table contains statistics of the radial velocities.

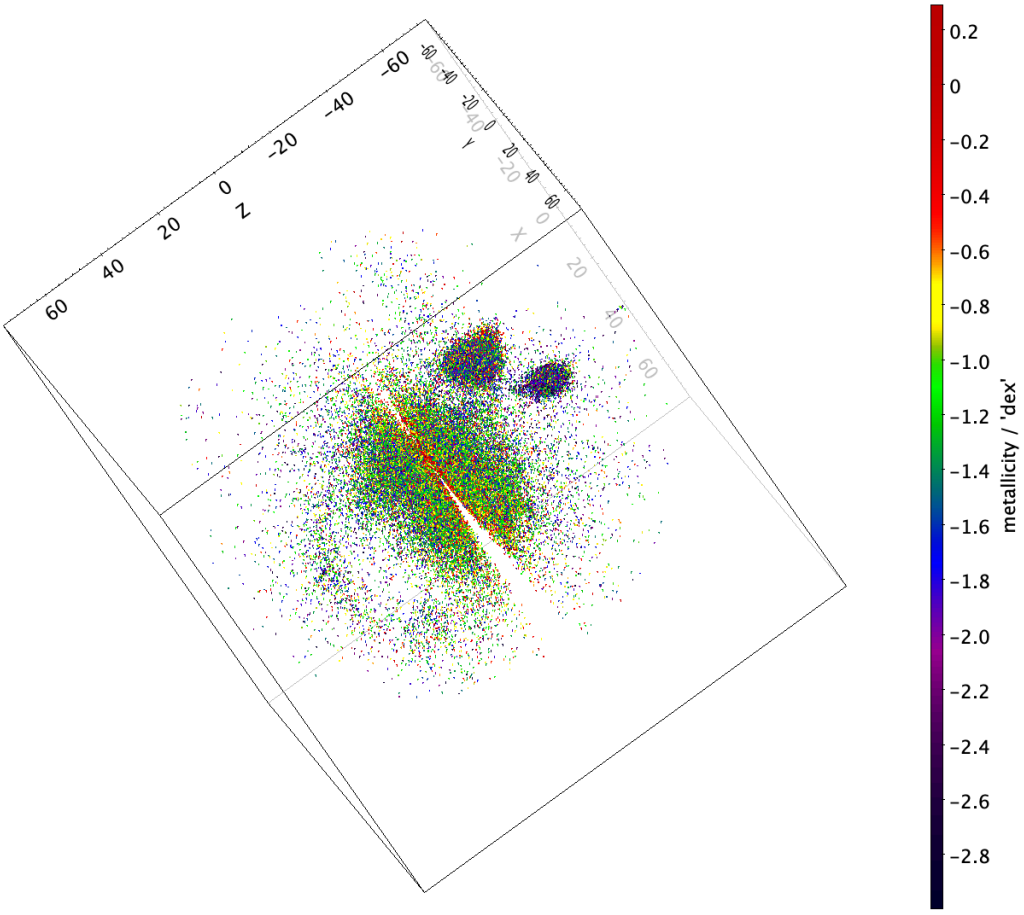


Fig. 5. Distribution of RR Lyrae stars in a 3D projection using the distance modulus. The LMC and SMC can be seen behind the Milky Way RR Lyrae stars. The metallicity near the Galactic plane is higher, lower in the halo, then lower in the LMC and then even lower in the SMC, as expected. We also see the Sagittarius stream. We can guess the presence of other structures in the halo.

5.1.1 The Performance Verification Paper on variability

A performance verification paper (PVP) authored by Gaia Collaboration, De Ridder et al. (2023) investigate the pulsation properties in main sequence OBAF-type stars. Among its noteworthy findings are: (1) the establishment of a period-luminosity relation for δ Scuti stars. (2) It is further shown that stellar rotation seems to reduce the amplitude of variability. (3) Discrepancies arise when comparing the observed locations of these variables with the predictions of theoretical models. Such inconsistencies may be attributed to inaccuracies in determining astrophysical parameters using Gaia instruments, and/or it is conceivable that rotational effects play a significant role, potentially altering the instability regions within the Hertzsprung-Russell diagram. (4) A quite remarkable result (see Appendix B of the work by Gaia Collaboration, De Ridder et al. (2023)) which demonstrates Gaia performance: with as few as 40 measurements, Gaia detects two accurate frequencies in certain cases when compared with the reference observations from the Kepler mission. As Gaia extends its

measurement duration to a decade, the precision of frequency determination, which scales inversely with the observation time basis, is anticipated to reach exceptional levels of accuracy.

There is also a publication on the spurious effect of spurious photometric variability introduced by the angle by which the satellite scans a (non axi-symmetric) source (Holl et al. 2023a).

5.1.2 The Gaia Andromeda Photometric Survey

To offer a teaser of DR4 and DR5, which will include epoch photometry/time series for **all sources**, we provide in the Gaia archive a pencil beam with a 5.5-degree opening angle centred toward the Andromeda galaxy, ensuring photometry availability for all sources within this region (Evans et al. 2023). This selection was made to provide a sufficiently broad sample of celestial sources reaching nearly 1.3 million sources; in addition to the Andromeda galaxy stars, the foreground stars, with significant parallaxes, allow us to provide a well-populated Hertzsprung-Russell Diagram. This dataset contains all the G, G_{BP}, G_{RP} time series in the Gaia archive. The Evans et al. (2023) article also presents a renormalisation of the photometric uncertainties (as these uncertainties are usually underestimated). We also show a way to combine the three photometric bands with a Principal Component Analysis method to enhance the detection and the characterisation of the variable sources, an approach similar to what was done with SDSS data (Süveges et al. 2012).

5.1.3 The exoplanets from Gaia with the transit method

Although the time span of DR3 was limited to 34 months, the quality of the photometry is sufficient to discover exoplanets by the transit method (Panahi et al. 2022b). Currently, there are two exoplanets detected by the transit method and confirmed by radial velocity measurements, Gaia-1b and Gaia-2b. The list of candidates is provided in the Gaia table `vari_planetary_transit` and contains 214 sources: 173 systems are known, and 41 are new (among them the two above already cited); the list of these 41 candidates can be found in Eyer et al. (2023). There is further work with the exoplanets with the transit methods, there has been a special agreement with the TESS mission (Ricker et al. 2015) to verify the TESS candidates by Gaia. Again, with only 34 months, we are able (1) to show that about 5% of the TESS candidates are background eclipsing binaries and also (2) to confirm about 5% of the cases where the transit can be identified (Panahi et al. 2022a).

As a side note, it is noteworthy to mention that: (1) astrometric exoplanet candidates were proposed by Holl et al. (2023b). At the time of writing, one of these candidates, named Gaia-3b, was confirmed thanks to ground-based radial velocity measurements (Sozzetti et al. 2023). (2) Exoplanets have also been discovered through microlensing. Gaia21dnc (Ban et al., in Prep) and Gaia22dkv (Wu et al. 2023) are the first two exoplanets detected through follow-ups of microlensing events alerted by the Gaia Science alert system (Hodgkin et al. 2021).

5.2 The focused product release: Time domain radial velocities for Long Period Variable stars

After the third data release, on October 10, 2023, there was a focused product release by the Gaia consortium, consisting of several products:

- Astrometry and photometry from engineering images taken in the ω Centauri region.
- The first results of quasars' environment analysis for gravitational lenses search.
- Extended radial velocity epoch data for Long Period Variables.
- Diffuse Interstellar Bands from aggregated RVS spectra.
- Updated astrometry for Solar System objects.

Here we present the Long Period Variables (LPV) (Gaia Collaboration, Trabucchi et al. 2023). The results of this focused product release are based on the DR3 processing both in photometry and radial velocities. A reminder of DR3 on Long Period Variables: It contains 1.7 million LPV. A selection was performed of stars fainter than $G = 6$ and brighter than $G = 14$. The total number of sources is 9,614 stars with radial velocities times series. Though this number may look small, it is a record holder: it is the largest sample ever published of Long Period Variable stars with radial velocity time series. It should be noted that the knowledge of radial velocities allows us to distinguish between pulsation (so true LPVs) and ellipsoidal variability.

The tables in the Gaia archive are located in the directory Gaia Focused Product Release and the subdirectory Variability. There are three tables:

- `gaiafpr.vari_epoch_radial_velocity`: Contains the radial velocity time series.
- `gaiafpr.vari_long_period_variable`: Similar to the DR3 table of the same name, but with added radial velocity data (frequencies and amplitudes are recomputed in both the radial velocity and G photometric band time series).
- `gaiafpr.vari_rad_vel_statistics`: Contains statistics on the radial velocities.

5.3 Citizen Science Projet: GaiaVari

GaiaVari is a citizen science project available on the Zooniverse platform that invites the community to classify variable stars from the public Gaia DR3 results based on the following various visual elements:

- The light curve, representing the G band magnitude measurements as a function of time (colour-coded with time).
- The folded curve, displaying the G band magnitude measurements as a function of phase computed from the observed time and an estimated period (the colour-coded according to the light curve, so according to the time, not the phase).
- The position in the Hertzsprung-Russell Diagram (the absolute G magnitude as a function of $G_{BP} - G_{RP}$) and the motion of the variable star within this diagram. There are no corrections from the extinction/reddening.
- The position in the Milky Way (the distance is taken as the inverse of the parallax).

The project is limited to stars with sufficiently good parallax measurements, allowing for the determination of a distance so that it can be placed within the Milky Way. The sample to classify is not exclusively focused on periodic objects, although the majority of objects chosen fall into this category. An example of the representation is given in Figure 6 for an RR Lyrae star.

Up to now, there have been one beta campaign and two official campaigns of GaiaVari. The first one classified the source into the possible following variability types: Eclipsing binaries, Cepheids, RR Lyrae stars and long-period variable stars. In the second campaign, the following variability types were added to the previous list: δ Scuti stars and stars with ellipsoidal variability. In both campaigns, a "None of the above" option was also introduced.

After making a classification choice, and only then, there is the possibility to leave comments. By choosing this option, one can also discover the Gaia sourceID and the classification from the automated supervised machine learning method (Rimoldini et al. 2023). There is also a link towards ESASky (Merín et al. 2015), and then one can obtain information from databases like Simbad at "Centre de Données astronomiques de Strasbourg" (CDS) (Wenger et al. 2000), etc.

Some citizen scientists systematically classify RR Lyrae stars and Cepheids into subgroups, such as RRab, RRc types, or Type-I, Type-II Cepheids. Such sub-classifications can be compared with the Gaia classification (Clementini et al. 2023; Ripepi et al. 2023) or the literature.

The community easily identifies the Blazhko effect, although, at times, it is challenging to distinguish Blazhko from an RRd type (double-mode RR Lyrae stars). Additionally, systematic differences in classification between what is published in DR3 and the classifications by some highly educated citizen scientists, particularly in RR Lyrae and Cepheids and their sub-classification, are noteworthy. It will be very interesting to compare these differences.

Due to demand we also created a permanent educational campaign for teachers and schools that can use it as an introduction resource to variable stars and HR diagram understanding.

We plan to conduct additional campaigns. At a point, the plan is to include radial velocity information for a more comprehensive dataset offered by Gaia.

Furthermore, there are plans to write articles on the results of GaiaVari from the first two campaigns,

GaiaVari is funded by ESA and driven by Sednai Sarl with a collaboration with the University of Geneva, ESA and ScienceNow, see https://www.zooniverse.org/projects/gaia-zooniverse/gaia-vari.

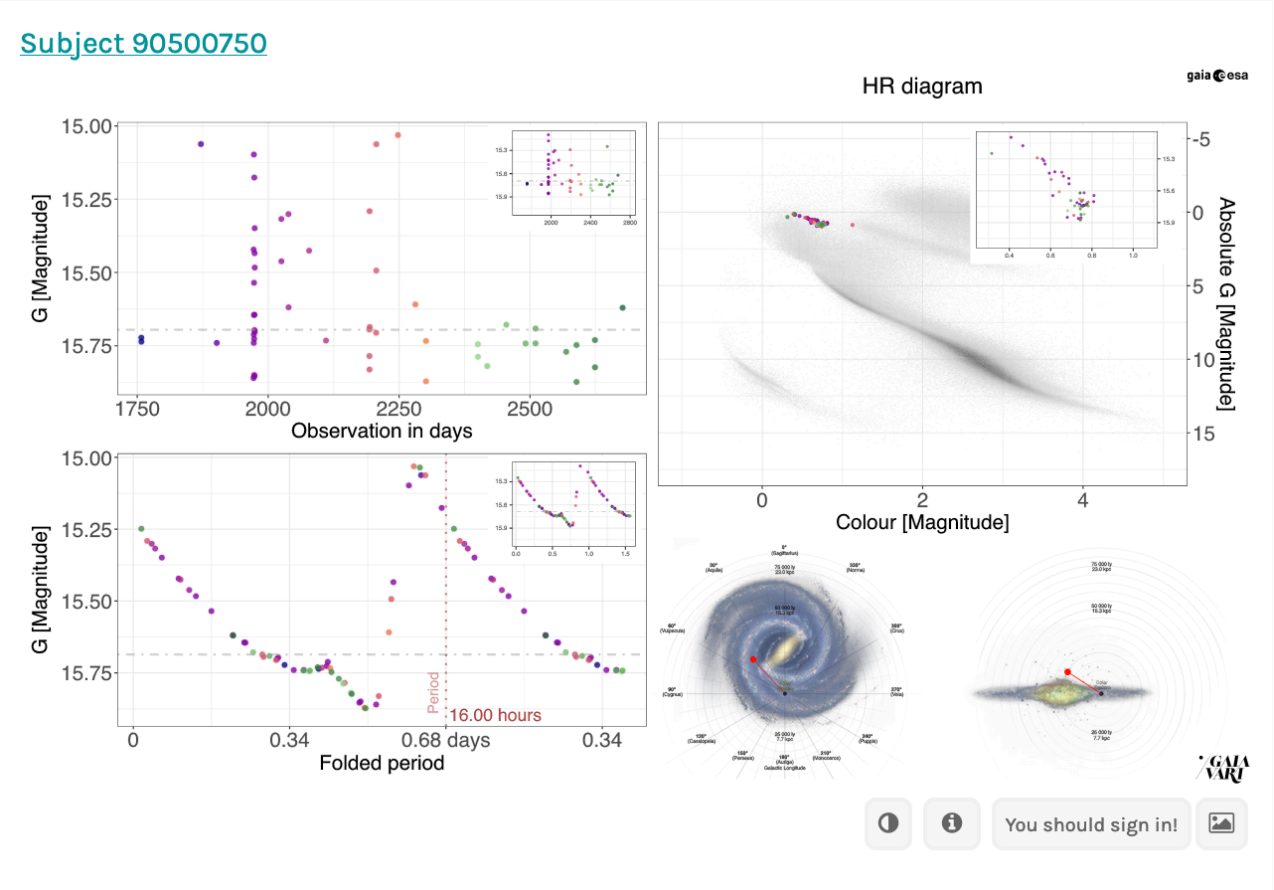


Fig. 6. RR Lyrae star of Bailey's type ab in GaiaVari. Upper left panel: the time series: Lower left panel: the folded curve, with the period fo 16 hours. Upper right panel: the motion of the star in the HR diagram. Lower right panel: the position in the Milky Way.

6 Conclusions

The quantity and quality of Gaia's data are impressive. Among the many scientific subjects that Gaia is touching, one is particularly positively impacted: time domain astronomy. In this context, we mostly presented the analysis of photometrically variable sources of DR3, covering 34 months of data with its multitude of results. Even with the relatively short DR3 time span, numerous records were broken in the number of variable sources, and the unprecedented description of them sets a new standard. Looking ahead, Gaia's data will nearly double its number of measurements in DR4 and, again, nearly double it for DR5, reaching about 10.5 years of science operational data (if all goes as planned). The Gaia variability pipeline for processing and analysing data in DR4 and DR5 will receive additional variability classes, features and improvements to the current ones, ensuring not only an increase in quantity but also an enhancement in quality. This holds the promise of even more interesting results for both the fourth and fifth Gaia data releases. Undoubtedly, Gaia's data alone provide remarkable insights into variable sources. However, we can anticipate exciting developments, such as combining Gaia data with other photometric (e.g. Vera Rubin observatory/LSST (Ivezić et al. 2019), ZTF (Bellm et al. 2019), PLATO (Rauer et al. 2014), etc.) or spectroscopic surveys (e.g. WST, Mainieri et al. (2024)).

I am grateful to the organisers of this School, in particular to Carine Babusiaux. I also appreciated the discussions with PhD students and postdocs. I would like to thank Mathieu Van der Swaelmen, whose question on the zero point parallax correction prompted a change in the Gaia consolidated catalogue for DR4. I thank Nicholas Chornay and Krzysztof Nienartowicz for reading and commenting on this manuscript.

I am thankful to the Geneva Gaia team, the Gaia Variability Processing and Analysis international team (CU7), and DPAC consortium. The success of such a mission lies in the competence and dedication of these Consortium members. Such an article would not exist without their contributions.

During the writing of this document, I had to face the declining health and the passing of my mother, Andrée Eyer Pétermann. Throughout my life, she was a constant source of support for my high aspirations and my work. I therefore dedicate this text to her.

I met Evry Schatzman only once when I was a young Ph.D. student. I was deeply moved by his diligence and interest in the young researcher I was, working on Hipparcos variable stars at the time. This is why I am quite certain that Evry Schatzman would have been most enthusiastic about the Gaia mission and its results.

The texts and figures are original; several software were used to improve the readability and the quality of the text, such as ChatGPT, Grammarly and Google Translate. This work made use of TOPCAT software (Taylor 2005). This work has made use of data from the European Space Agency (ESA) mission *Gaia* (https://www.cosmos.esa.int/gaia), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, https://www.cosmos.esa.int/web/gaia/dpac/consortium). Funding for the DPAC has been provided by national institutions, in particular, the institutions participating in the *Gaia* Multilateral Agreement.

References

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016, Phys. Rev. Lett., 116, 061102

Appenzeller, I. 2013, Introduction to Astronomical Spectroscopy

Bakos, G., Noyes, R. W., Kovács, G., et al. 2004, PASP, 116, 266

Bellm, E. C., Kulkarni, S. R., Graham, M. J., et al. 2019, PASP, 131, 018002

Borucki, W. J., Koch, D., Basri, G., et al. 2010, Science, 327, 977

Burki, G., Rufener, F., Burnet, M., et al. 1995, A&AS, 112, 383

Carnerero, M. I., Raiteri, C. M., Rimoldini, L., et al. 2023, A&A, 674, A24

Carney, B. W., Storm, J., & Jones, R. V. 1992, ApJ, 386, 663

Chakrabarti, S., Simon, J. D., Craig, P. A., et al. 2023, AJ, 166, 6

Clementini, G., Ripepi, V., Garofalo, A., et al. 2023, A&A, 674, A18

Clementini, G., Ripepi, V., Leccia, S., et al. 2016, A&A, 595, A133

Davis, R., Harmer, D. S., & Hoffman, K. C. 1968, Phys. Rev. Lett., 20, 1205

de Jong, R. S., Agertz, O., Berbel, A. A., et al. 2019, The Messenger, 175, 3

Distefano, E., Lanzafame, A. C., Brugaletta, E., et al. 2023, A&A, 674, A20

El-Badry, K., Rix, H.-W., Quataert, E., et al. 2023, MNRAS, 518, 1057

Evans, D. W., Eyer, L., Busso, G., et al. 2023, A&A, 674, A4

Eyer, L. 2006, Mem. Soc. Astron. Italiana, 77, 549

Eyer, L., Audard, M., Holl, B., et al. 2023, A&A, 674, A13

Eyer, L. & Mignard, F. 2005, MNRAS, 361, 1136

Eyer, L., Mowlavi, N., Evans, D. W., et al. 2017, arXiv e-prints, arXiv:1702.03295

Eyer, L., Rimoldini, L., Holl, B., et al. 2015, in Astronomical Society of the Pacific Conference Series, Vol. 496, Living Together: Planets, Host Stars and Binaries, ed. S. M. Rucinski, G. Torres, & M. Zejda, 121

Gaia Collaboration, Clementini, G., Eyer, L., Ripepi, V., et al. 2017, A&A, 605, A79

Gaia Collaboration, De Ridder, J., Ripepi, V., Aerts, C., et al. 2023, A&A, 674, A36

Gaia Collaboration, Eyer, L., Rimoldini, L., Audard, M., et al. 2019, A&A, 623, A110

Gaia Collaboration, Panuzzo, P., Mazeh, T., Arenou, F., et al. 2024, A&A, 686, L2

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

Gaia Collaboration, Trabucchi, M., Mowlavi, N., Lebzelter, T., et al. 2023, A&A, 680, A36

Gavras, P., Rimoldini, L., Nienartowicz, K., et al. 2023, A&A, 674, A22

Gavras, P., Rimoldini, L., Nienartowicz, K., et al. 2022, VizieR Online Data Catalog, J/A+A/674/A22

Gomel, R., Mazeh, T., Faigler, S., et al. 2023, A&A, 674, A19

Hess, V. F. 1912, Phys. Z., 13, 1084

Hodgkin, S. T., Harrison, D. L., Breedt, E., et al. 2021, A&A, 652, A76

Holl, B., Audard, M., Nienartowicz, K., et al. 2018, A&A, 618, A30

Holl, B., Fabricius, C., Portell, J., et al. 2023a, A&A, 674, A25

Holl, B., Sozzetti, A., Sahlmann, J., et al. 2023b, A&A, 674, A10

Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, ApJ, 873, 111

Lebzelter, T., Mowlavi, N., Lecoeur-Taibi, I., et al. 2023, A&A, 674, A15

Mainieri, V., Anderson, R. I., Brinchmann, J., et al. 2024, arXiv e-prints, arXiv:2403.05398

Marton, G., Ábrahám, P., Rimoldini, L., et al. 2023, A&A, 674, A21

Merín, B., Salgado, J., Giordano, F., et al. 2015, arXiv e-prints, arXiv:1512.00842

Mowlavi, N., Holl, B., Lecoeur-Taïbi, I., et al. 2023, A&A, 674, A16

Panahi, A., Mazeh, T., Zucker, S., et al. 2022a, A&A, 667, A14

Panahi, A., Zucker, S., Clementini, G., et al. 2022b, A&A, 663, A101

Pogson, N. 1856, MNRAS, 17, 12

Pourbaix, D., Arenou, F., Gavras, P., et al. 2022, Gaia DR3 documentation Chapter 7: Non-single stars, Gaia DR3 documentation, European Space Agency; Gaia Data Processing and Analysis Consortium. Online at https://gea.esac.esa.int/archive/documentation/GDR3/index.htm, id. 7

Rauer, H., Catala, C., Aerts, C., et al. 2014, Experimental Astronomy, 38, 249

Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003

Riello, M., De Angeli, F., Evans, D. W., et al. 2021, A&A, 649, A3

Rimoldini, L., Holl, B., Gavras, P., et al. 2023, A&A, 674, A14

Ripepi, V., Clementini, G., Molinaro, R., et al. 2023, A&A, 674, A17

Sahu, K. C., Anderson, J., Casertano, S., et al. 2022, The Astrophysical Journal, 933, 83

Schatzman, E. 1963, Astrophysique (Masson et Cie)

Sozzetti, A., Pinamonti, M., Damasso, M., et al. 2023, A&A, 677, L15

Süveges, M., Sesar, B., Váradi, M., et al. 2012, MNRAS, 424, 2528

Taylor, M. B. 2005, in Astronomical Society of the Pacific Conference Series, Vol. 347, Astronomical Data Analysis Software and Systems XIV, ed. P. Shopbell, M. Britton, & R. Ebert, 29

Udalski, A. 2003, Acta Astron., 53, 291

Wenger, M., Ochsenbein, F., Egret, D., et al. 2000, A&AS, 143, 9

Wu, Z., Dong, S., Yi, T., et al. 2023, arXiv e-prints, arXiv:2309.03944

Wyrzykowski, Ł., Kruszyńska, K., Rybicki, K. A., et al. 2023, A&A, 674, A23

École Evry Schatzman 2023: Stellar physics with Gaia
C. Babusiaux, C. Reylé (eds)

ULTRACOOL DWARFS IN GAIA

C. REYLÉ

Université de Franche-Comté, Institut UTINAM, CNRS UMR6213, OSU THETA Franche-Comté-Bourgogne, Observatoire de Besançon, BP 1615, 25010 Besançon Cedex, France

e-mail: celine.reyle@obs-besancon.fr

Abstract. The Gaia optical observations are not the most suitable spectral domain for studying the low-mass, faintest and reddest part of the main sequence. Nevertheless, the large number of objects observed with an unprecedented precision, including trigonometric parallax, makes it an unrivaled dataset for studying the stellar-substellar boundary. In this paper, I review the contribution of the successive catalogues offered by the Gaia mission to study the low-mass, ultra-cool objetcs. I also present further characterisations and scientific exploitations of the Gaia sample.

Keywords: stars: low-mass – brown dwarfs – Galaxy: stellar content – surveys – catalogues

1 Introduction

Ultra-cool dwarfs (hereafter UCDs) are defined as M7 and later type objects by Kirkpatrick et al. (1997). Their temperature is below about 2700 K (e.g. Rajpurohit et al. 2013). They are the link between stars and brown dwarfs as they can be either the least massive stars or brown dwarfs, spanning the stellar-substellar masses. The UCD census is still incomplete even within 25 pc of the Sun: Bardalez Gagliuffi et al. (2019) estimated that 62% of M7-L5 are catalogued in this volume.

In what follows, I review the contribution of the Gaia space mission (Gaia Collaboration, Prusti et al. 2016) on the study of UCDs. I first describe the UCD regime, in between the lowest mass stars and brown dwarfs, and give hints on the motivation to detect and characterize them. I illustrate the powerful utility of Gaia, thanks to its huge, homogeneous and precise dataset to reveal the stellar-substellar boundary. There is no doubt that Gaia next data releases will complete the picture.

2 Low-mass stars and browns dwarfs

A vast amount of energy is released via nuclear fusion occurring in the core of a star. The fusion reaction is accompanied by a net loss of mass, resulting in a release of energy. To fuse hydrogen, the core must have temperatures larger than 3 millions K. On the Main Sequence, the thermal pressure from fusion keeps a star from gravitational collapse and the star is in thermal and hydrostatic equilibrium. The stellar matter follows classical statistical physics: classical nearly perfect gas equation of state and quasi-static equilibrium condition. The radius of a star is proportional to its mass (see e.g. Figure 15 in Boyajian et al. 2012, and blue symbols in Figure 1). The less mass a star has, the more it needs to contract to heat the core, and the smaller it will be on the Main Sequence. M-dwarfs are the lowest mass stars, ten times smaller and ten times less massive than the Sun.

The question on how small can a star be and form by gravitational collapse of a giant molecular cloud is still open. The formation scenario, from filamentary clouds to pre-stellar cores to stars, has been refined in particular thanks to the wide-field *Herschel* photometric survey of nearby star-forming cloud complexes (André et al. 2010). However the gravitational collapse of a giant molecular cloud into very low mass objects requires a very high density (about 10^7 cm^{-3} to form a 0.07 M$_\odot$ object). Alternative mechanisms to form such small objects are still under debate: embryo ejection (Bate et al. 2002; Goodwin & Whitworth 2004; Reipurth & Mikkola

2015), photo-erosion of cores near massive stars (Whitworth & Zinnecker 2004), protostellar disc fragmentation (Vorobyov & Basu 2010; Attwood et al. 2009; Stamatellos & Herczeg 2015), gravoturbulent fragmentation (Padoan & Nordlund 2004; Hennebelle & Chabrier 2008; Bonnell et al. 2008; Lomax et al. 2016), or compression by turbulent flows in molecular clouds (Stamer & Inutsuka 2019).

The contraction of the forming object makes its core temperature and density increase. If its mass is not sufficient, the temperature will not reach the hydrogen burning threshold and the object will go towards the electron degeneracy limit. The collapse is stopped by electron degeneracy pressure: this substellar object is named a brown dwarf. This theoretical picture has been proposed in 1963 by Kumar (1963); Hayashi & Nakano (1963). Kumar (1963) estimated a hydrogen burning minimum mass of 0.07 $M_\odot$ (73 M_{Jup}) for objects with solar composition and 0.09 $M_\odot$ (94 M_{Jup}) for low metallicity objects. Brown dwarfs are the densest hydrogen-rich objects known (see Figure 1 from Hatzes & Rauer 2015). The macroscopic properties of the substellar matter are ruled by different physics and follow a different equation of state than in stars (e.g. Saumon et al. 1995; Chabrier et al. 2023).

Because brown dwarfs do not undergo stable hydrogen fusion, they cool down over time, progressively passing through later spectral types as they age. They are low-luminosity, very red objects, and difficult to detect. Strategies for finding these coolest objects are to search for them in young clusters (young brown dwarfs are brighter), to search for them as companions to other objects, to search for red objects in large scale surveys, eventually among large proper motion objects (being an indication of possible nearby, intrinsically faint, objects). First discoveries came out in the late 90s, thanks to the advance of the near-infrared technology: Teide 1 in the Pleiades (Rebolo et al. 1995), the cool brown dwarf Gliese 229B as a companion of the low-mass star Gliese 229A (Nakajima et al. 1995), Kelu 1 and three other free-floating brown dwarfs (Ruiz et al. 1997; Delfosse et al. 1997), and many discoveries since then thanks to the large scale surveys (DENIS, 2MASS, CFHTLS, SDSS, SIMP, UKIDSS, WISE, PanSTARRS,...).

In parallel to these numerous discoveries, evolutionary models were developed by several groups, assuming cloudless or cloudy atmospheres, placing the stellar-substellar limit between 70 and 79 M_{Jup} (Chabrier & Baraffe 2000; Burrows et al. 2001, 2011; Saumon & Marley 2008; Baraffe et al. 2015; Marley et al. 2021). The empirical mass limit of hydrogen fusion has also been determined by Dupuy & Liu (2017) from the dynamical mass of 31 binaries. They estimated the stellar-substellar boundary at 70±4 M_{Jup}. In general, brown dwarf properties are bracketed between those of low-mass stars and massive giant planets. 13 M_{Jup} is the mass limit to still allow nuclear fusion, that of deuterium (e.g. Burrows 1999). The IAU uses this mass as the limit to define a brown dwarf and an exoplanet, no matter how they formed (see Lecavelier des Etangs & Lissauer 2022).

The minimun size of a brown dwarf is about the size of Jupiter. In a more massive brown dwarf, gravitational force is higher and causes a larger fraction of the brown dwarf to become degenerate, causing it to have a smaller radius. The mass to radius relation shows a local minimum in the most massive brown dwarfs. The reversal of the mass-radius relation at the hydrogen burning limit is predicted by models (Figure 1). At a given mass, theoretical isochrones predict that older objects have smaller radii. Brown dwarfs and low mass stars observed by transit in binary systems, with age estimate from the primary star, are efficient to test the age-radius effect. Generally they validate model radii (Carmichael et al. 2021; Grieves et al. 2021).

Figure 2 (left panel) shows the evolution of objects with mass between 10 and 100 M_{Jup} in the luminosity vs age diagram. The tracks are obtained from Baraffe et al. (2015) models. Once a star reaches the Main Sequence, its luminosity stabilizes, contrary to substellar objects. As a consequence, very low-mass stars, brown dwarfs, and planetary mass objects can have the same brightness (and effective temperature). As an illustration, a 2 Myr planetary mass object, a 50 Myr brown dwarf, or a 300 Myr star, have the same luminosity ($\sim 10^{-3}$ $L_\odot$). This is also clearly illustrated from observations of a sample of substellar companions with well-constrained ages and spectroscopically-derived classifications, in Figures 6 and 8 by Bowler (2016).

The evolution in effective temperature is also shown in Figure 2 (right panel). The horizontal lines at 2200 K, 1400 K, and 550 K roughly represent the transition between M and L dwarfs, L and T dwarfs, T and Y dwarfs, respectively. Spectral types of stars on the Main Sequence, O B F G K M, is a temperature (and also mass) related sequence. The spectral sequence for brown dwarfs, M L T Y, is not only a temperature sequence but also an evolutionary sequence. A massive, old, brown dwarf can have the same spectral type of a less massive, young, brown dwarf. The youngest brown dwarfs are M-types. Most of the brown dwarfs then evolve from M, L, and T-types. Only the less massive ones go to Y-types. The edge of the hydrogen burning main sequence is an L-dwarf.

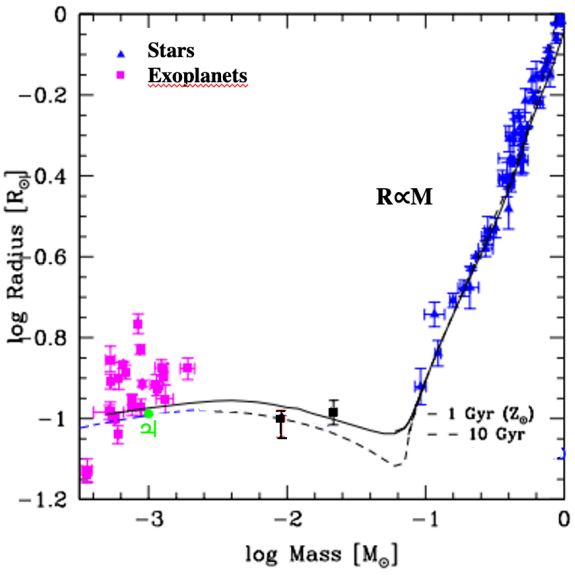


Fig. 1. Theoretical mass to radius relation computed from models at solar composition at 1 Gyr (solid line) and 10 Gyr (dashed line), superimposed to observationally-determined values (symbols). Jupiter is shown in green. Adapted from Chabrier et al. (2009).

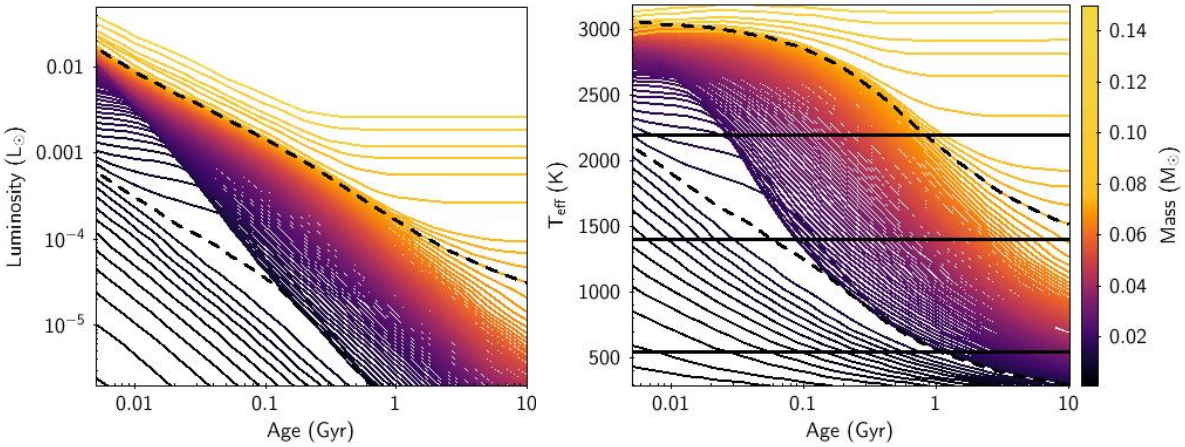


Fig. 2. Evolutionary tracks (left panel: luminosity, right panel: effective temperature) from Baraffe et al. (2015) models, for objects with planetary to stellar mass. The dotted lines are the 73 and 13 M_{Jup} tracks. The horizontal lines (right panel) show the M/L, L/T, and T/Y transitions, from top to bottom.

3 Why are low-mass stars and brown dwarfs interesting to study?

Low-mass stars are the most numerous in the Galaxy. In the 10 pc volume-limited sample compiled by Reylé et al. (2021, 2022), 61% of the stars are M-dwarfs, half of them with a spectral type from M3 to M5. The Gaia Catalogue of Nearby Stars (GCNS, Gaia Collaboration, Smart et al. 2021) is volume-limited to 100 pc, compiled from the Gaia early third data release (EDR3, Gaia Collaboration, Brown et al. 2021). It contains at least 92% of stars of stellar type M9 within 100 pc of the Sun. The distribution of the stars peaks at the partly to fully convective stars boundary (Figure 3). The M-dwarfs ($8.25 \leq M_G \leq 15.5$) account for 64% of the full sample.

Low-mass stars and brown dwarfs evolve slowly and have a lifetime much larger than the age of the universe. Thus they span all ages and can be found in all populations, from young stellar associations to the halo (see section 5). They cover a huge range of astrophysical properties. Their absolute magnitudes span 20 magnitudes and 5 magnitudes in colors! This large variety of characteristics is the consequence of the complex physical processes acting in their cool atmospheres.

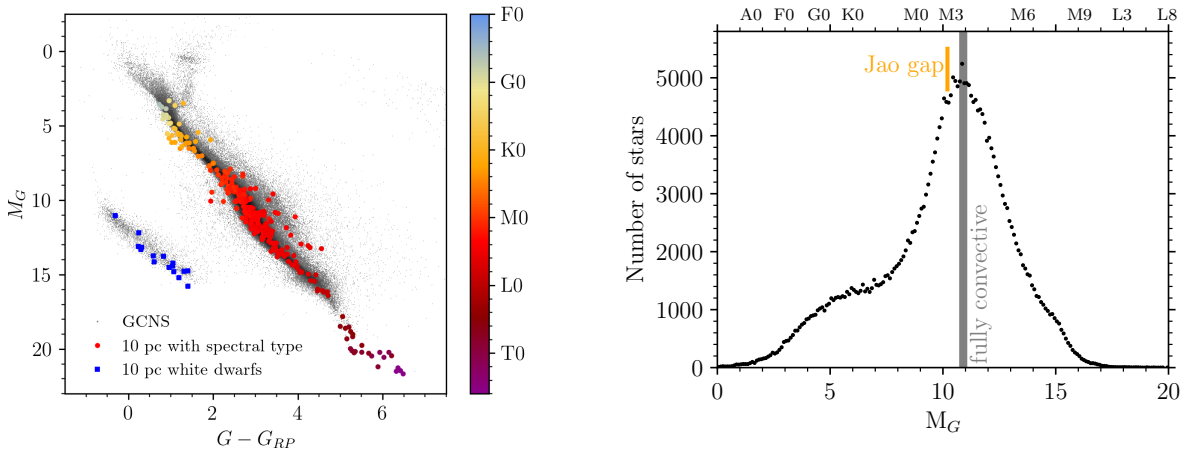


Fig. 3. Left panel: color absolute magnitude diagram of the GCNS (grey dots) superimposed with the 10 pc sample, colored by spectral type. Right panel: number of main sequence stars, in 0.1 magnitude bins, as a function of absolute magnitude, obtained from the GCNS. The high quality and precision of the sample allow to detect the Jao gap (Jao et al. 2018), a structure in the main sequence related to structural instabilities.

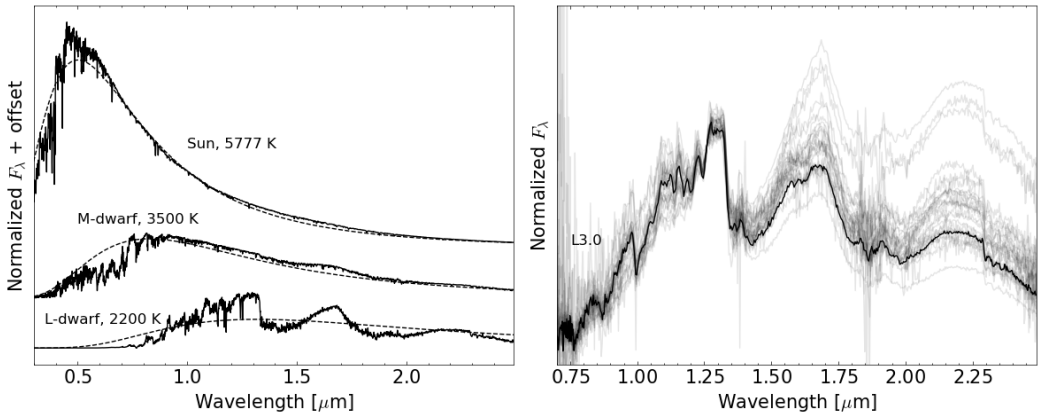


Fig. 4. Left: stellar model spectra and black body emission at the same temperature for the Sun, a M-dwarf, and a L-dwarf. Right: Spectra of various L-dwarfs, all classified as L3. The spectra come from the SpeX Prism Spectral Libraries, maintained by Adam Burgasser (Burgasser & Splat Development Team 2017). The black spectrum is the standard 2MASSW J1506544+132106 (Burgasser 2007). Courtesy Thomas Ravinet.

They remain elusive due to their faint magnitude. Modeling their complex, cool atmosphere is still a challenge. The strong absorptions due to the formation of molecules make their spectrum to deviate from the black body spectrum, and one cannot talk anymore of continuum but pseudo-continuum (Figure 4, left panel). In particular, the L-dwarfs offer a tricky and puzzling case, as illustrated in Figure 4, right panel. All classified as L3-dwarfs, the shape of their spectrum are modified by the condensates that form in their atmosphere. In the near infrared, dust can lead to back warming of the atmosphere, and alters the amount of H_2O and H_2, and slight variations in the H_2O and H_2 opacities can lead to large differences. Moreover, low-gravity objects will show a triangular shaped H-band (Kirkpatrick et al. 2008) whereas low-metallicity objects will present a higher flux in the J-band while suppressed H and K-bands. The full description of the atmosphere depends on effective temperature, gravity, metallicity, cloud properties, and mixing in the atmosphere.

Low mass stars host exoplanets, and are ideal targets for searches for potentially habitable terrestrial planets: e.g. TRAPPIST-1 (Gillon et al. 2016, 2017), Proxima Cen (Anglada-Escudé et al. 2016), Ross 128 (Bonfils et al. 2018). The lowest-mass brown dwarfs more closely resemble the gas giant planets than stars (see Figure 1 by

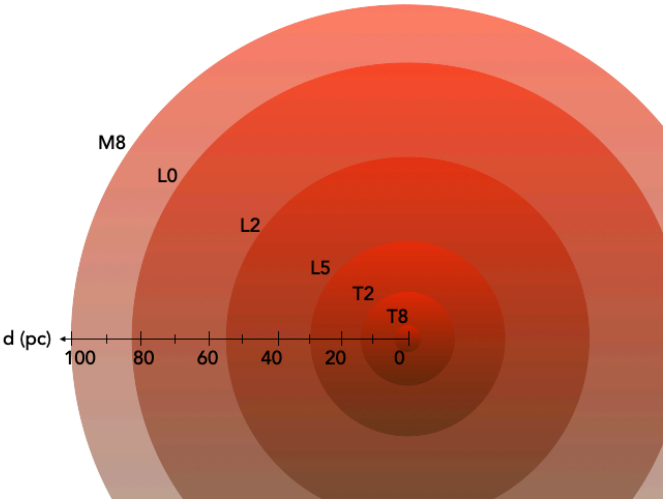


Fig. 5. Distance limits vs spectral type assuming the Gaia magnitude limit G=20.7 estimated from Smart et al. (2017).

Marley & Leggett 2009) and therefore serve as laboratories for understanding gas-giant extrasolar planets as well as the faint end of the star formation process. When these low-mass brown dwarfs are close enough and bright enough to be observed spectroscopically, their atmospheres are much easier to study than the ones of similar exoplanets, hampered by the light of the host star.

Hence their study has relevant implications for stellar, exoplanetary, and Galactic astronomy.

4 Ultra-cool dwarfs in Gaia

Gaia provides the means to uncover ultra-cool dwarfs through astrometric, rather than purely photometric, selection. They can be selected from their locus in the color absolute magnitude diagram. Gaia holds the promise of a truly volume-complete sample, with distance limit depending on the spectral type (Figure 5). Sarro et al. (2013) estimated the expected end-of-mission number of UCDs in the Gaia archive: more than 40 000 objects, 600 objects between L0 and L5, 30 objects between L5 and T0, and 10 objects between T0 and T8.

Later on, (Smart et al. 2017) compiled the Gaia ultra-cool dwarf sample (GUCDS), a catalogue of known L and T dwarfs spectroscopically confirmed at that time, containing 1010 L and 58 T having a Gaia G predicted magnitude smaller than 21.5. They cross-matched these known UCDs with Gaia first data release (DR1 Gaia Collaboration, Prusti et al. 2016) and identified 319 L and 2 T. Although incomplete, this first sample was used by the Data Processing and Analysis Consortium (DPAC) as a starting point to develop pipeline for parameter estimation purposes based on the low resolution Gaia RP spectra (see Carrasco et al. 2021; De Angeli et al. 2023). Gaia Collaboration, Babusiaux et al. (2018) found 443 L and 7 T among selected precise data from the second release (DR2, Gaia Collaboration, Brown et al. 2018). With a slightly different selection, Gaia Collaboration, Brown et al. (2018) identified 647 L and 16 T, as well as 3671 M7 to M9 dwarfs. Finally 4767 known M7 to M9, 1061 L and 16 T are identified in the third data release (DR3, Gaia Collaboration, Vallenari et al. 2023, Richard Smart, private communication).

This sample constitutes an unprecedented sample, with precise distance determination, to define absolute magnitude vs color, and vs spectral type relations (see Table 1 from Reylé 2018). It provides strong constraints to evolutionary models (Figure 6, left panel). Finally, its locus in the precise M_G vs $G - G_{\rm RP}$ diagram offers an obvious way to identify new UCD candidates (Figure 6, right panel).

To select robust UCD candidates, Reylé (2018) used Gaia DR2 data filtered following Gaia Collaboration, Babusiaux et al. (2018) who selected the most precise data (in parallax and photometry, but also handling the extinction rigorously), without trying to reach completeness. The selection in done in the M_G vs $G - G_{\rm RP}$ diagram based on the locus of the known UCD sample. As explained by Evans et al. (2018), faint objects may have spurious $G - G_{\rm RP}$ colors because the measured RP (as well as BP) flux may include a contribution of flux from bright sky background (e.g., in crowded regions or from a nearby source). Evans et al. (2018) defined a filter based on the consistency between the flux in the G, $G_{\rm BP}$, and $G_{\rm RP}$ bandpass to remove objects with spurious colors. However the use of $G_{\rm BP}$ is not appropriate for UCDs because these faint and red objects have a low flux in this band. Reylé (2018) defined a new empirical filter based on the 2MASS J band, excluding objects

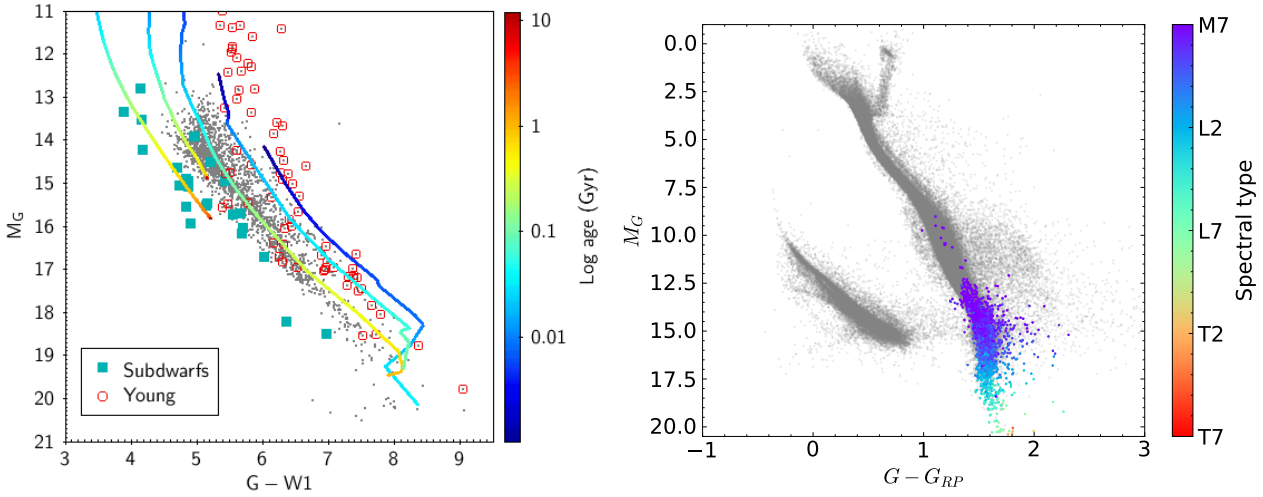


Fig. 6. Left: Known UCDs identified in Gaia DR2. The symbols depict the objects listed as potential subdwarfs or young candidates. The lines are evolution tracks at different masses and metallicities (from right to left : 0.1, 0.2, 0.5, 0.9 $M_\odot$ at solar metallicity, and 0.83 $M_\odot$ at metallicity -1). The circles on the tracks are colored by the age in Gyr. The evolutionary models are computed by Baraffe et al. (2015) that consistently couple interior structure calculations with the BT-Settl atmosphere models (Allard et al. 2013). Right: locus of known UCDs in the M_G vs $G - G_{\rm RP}$ diagram, colored by spectral type.

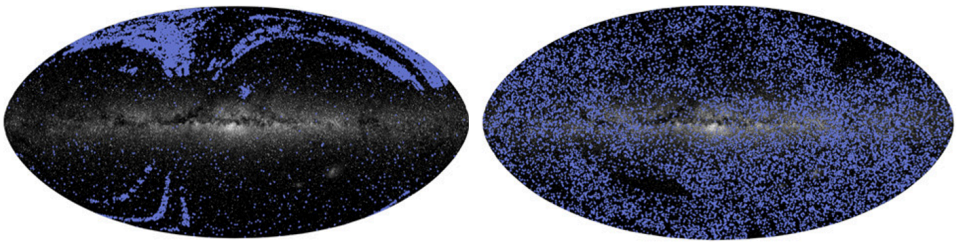


Fig. 7. Sky distribution of known UCDs (left) and new candidates found in Gaia DR2 (right).

with spurious colors, but retaining the low-mass objects (in particular L-dwarfs). This selection let to numerous new candidates, 14 176 M7 to M9 and 488 L (all earlier than L5). They are evenly distributed across the sky, filling in missing populations in the Southern hemisphere and Galactic plane (Figure 7). Their distance shows that the local census is still incomplete. At 25 pc, the number of new candidates equals the number of known objects retrieved in Gaia DR2. Beyond 30 pc, the number of new candidates exceeds that of the known sample. Using the same technique, 2879 additional candidates were identified in the GCNS based on Gaia EDR3.

The Gaia third data release (DR3, Gaia Collaboration, Vallenari et al. 2023) offered in addition the opportunity to use low-resolution spectra to refine and widen the selection. The ESP-UCD (*Extended Stellar Parametrizer Ultra Cool Dwarfs*) module infers the effective temperature from the shape of the RP spectrum (Figure 8). 94 158 UCD candidates with $T_{\rm eff}$ estimates below 2700 K were identified by Sarro et al. (2023). The readers are also referred to the lecture noted by Creevey (2024) who presents the data products from Gaia DR3 that can be exploited for stellar physics.

5 Characterization

5.1 Astrophysical parameters

Several efforts have been made and are still on-going for spectroscopic follow-up of new UCD candidates in order to confirm their nature and further characterize them. In general the spectroscopic follow-up requires the use of 4-meter facilities.

Near-infrared spectroscopic follow-up of 60 nearby (closer than 50 pc) UCDs candidates with the SOFI intrument at the New Technology Telescope (NTT) in la Silla observatory, are presented by Ravinet et al.

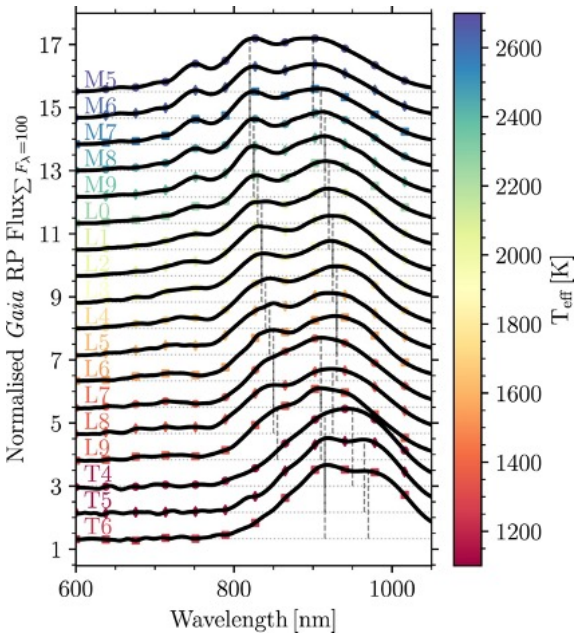


Fig. 8. Normalized median RP fluxes for M5 to T6 spectral type, indicated by labels and coloured by effective temperature. Vertical dashed lines indicate the position of the two primary spectral peaks. All spectra are normalized and linearly offset. From Cooper et al. (2024)

(2024). Their spectral type, derived from template-matching (Figure 9, left panel) using the SpeX Prism Library Analysis Toolkit (SPLAT, Burgasser & Splat Development Team 2017) is very close (standard deviation of one subtype) to the one expected using their absolute magnitude M_G, showing that M_G can be used with confidence as a proxy to select UCDs in Gaia data. Ravinet et al (in prep) also made a comparison with synthetic spectra (Figure 9, right panel) computed from various atmospheric models: DRIFT (Witte et al. 2011), BT-Settl-CIFIST and BT-Settl-AGSS (Allard 2014) (assuming different solar abundances Caffau et al. (2011) and Asplund et al. (2009), respectively), and ATMO (Phillips et al. 2020). The retrieved stellar parameters ($T_{\rm eff}$, $\log g$, [Fe/H]) can be quite different depending on the models used: comparing properties of stars obtained with different models must be handled with care. In general, ATMO synthetic spectra give the best fit, although the H-band (1.5 - 1.8 μm) is often poorly reproduced by all models, in particular at the M/L transition where dust starts to condensate in the atmosphere. It illustrates the difficulty to properly model the effect of dust in these cool atmospheres.

Spectroscopic follow-up permits also to identify binaries from "flux-reversal": binaries whose component straddle the L/T transition show a brighter secondary in the J-band than the H and K-bands (e.g. Cruz et al. 2004; Looper et al. 2008). Such unresolved binaries, named spectral binaries, can be identified by an index-identification technique (Burgasser et al. 2010; Bardalez Gagliuffi et al. 2014), using the spectral deviations between the binaries and templates, such as a stronger methane absorption in the H-band relative to the K-band, the K-band flux peak slightly shifted toward the blue in the binaries, etc (Figure 10).

A new advent came with Gaia DR3 that provides astrophysical parameters (Creevey et al. 2023). The "golden sample", which is a homogeneous sample of stars with high-quality astrophysical parameters compiled by Gaia Collaboration, Creevey et al. (2023) contains 21 068 UCDs. Figure 11 illustrates in two ways the quality and potential scientific outcome of this sample. The top panel shows the $T_{\rm eff}$ determined from the ESP-UCD module vs the radius determined from the FLAME module. It shows the expected decrease in radius as the temperature decreases with a minimum value at 2000-2200 K, in agreement with the findings of Dieterich et al. (2014), followed by an increase until 1400 K where the trend reverses. Another application is to constrain the characteristics of faint, unseen UCDs that are beyond the Gaia magnitude limit but are in binary systems with brighter objects. 11 multiple systems were identified from their common distance and proper motion (after allowing for orbital motion). Assuming that the UCD has the same age as the primary, the absolute magnitude vs age diagram, superimposed with evolutionary tracks (Figure 11, bottom rpanel) gives some indication on the

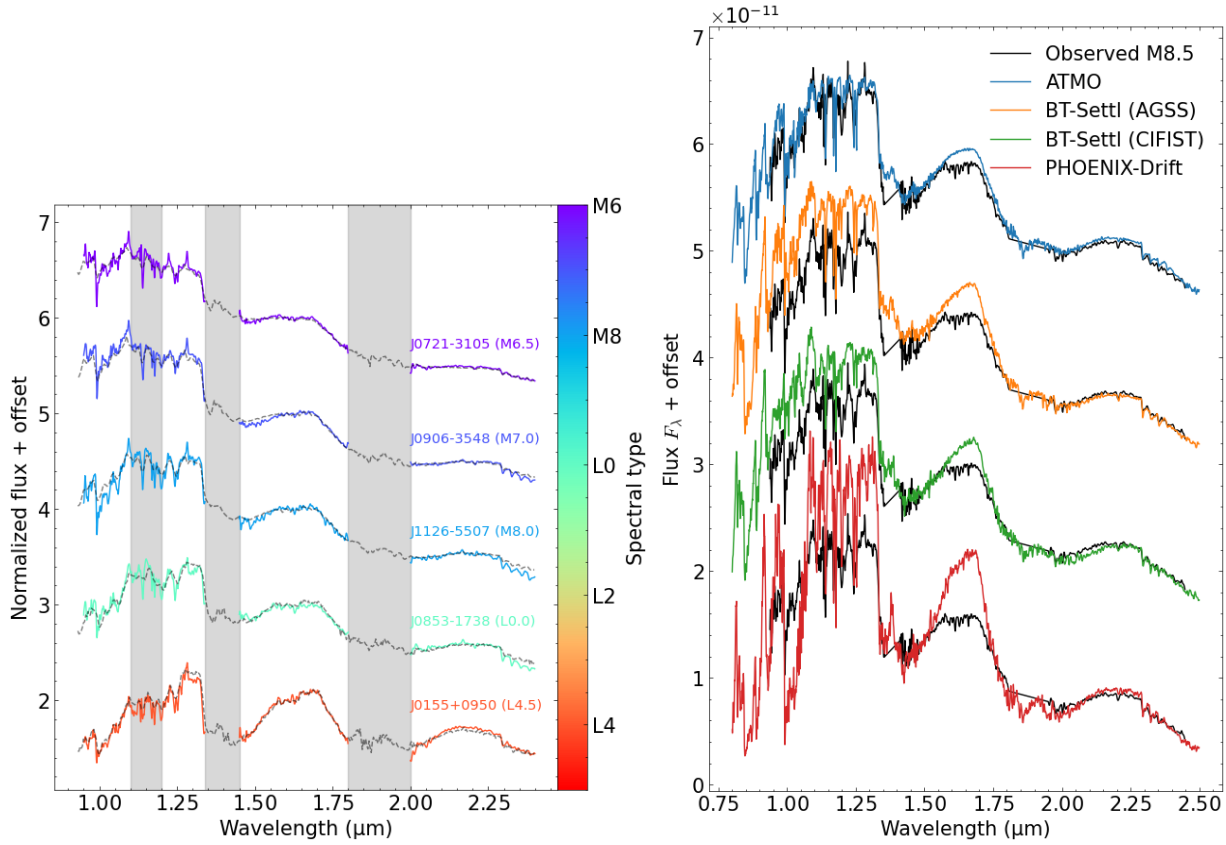


Fig. 9. Left: SOFI spectra (filled lines, colored by spectral type) of UCDs, together with standard spectra of the same spectral type. Right: SOFI spectrum of a M8.5 (black) superimposed with best fit synthetic spectra computed from different atmosphere models. Courtesy Thomas Ravinet.

stellar or substellar nature of the UCDs.

5.2 Subdwarfs

Subdwarfs are low-metallicity objects and are typically found to have thick disc or halo kinematics. The name was coined by Gerard Kuiper in 1939, to refer to a series of stars with anomalous spectra that lie between the Main Sequence and the white dwarf sequence in the Hertzsprung-Russel diagram and were previously labeled as "intermediate white dwarfs" (Figure 12, left panel). Solar-metallicity objects are classified as dwarfs, while the more metal-poor stars are classified as subdwarfs (sd), extreme subdwarfs (esd), and ultra subdwarfs (usd), in order of decreasing metallicity (Gizis 1997; Lépine et al. 2007). Because of a decreasing metallicity, TiO opacity decreases in M-subdwarf atmospheres. Less blanketing from TiO bands means more continuum flux radiated from hotter and deeper layers of the atmosphere. The M-subdwarf spectrum is closer to that of a blackbody, and subdwarfs appear bluer (Figure 12, right panel). In L-subdwarfs, the enhanced collision-induced H_2 absorptions suppress the K and K-bands.

Because low-mass subdwarfs have lifetimes far in excess of the age of the Galaxy, they are important tracers of Galactic chemical history and representative of the first generations of star formation. To date hundreds of late-type M-subdwarfs (e.g. Lépine et al. 2003; Burgasser et al. 2007; Lépine & Scholz 2008; Lodieu et al. 2012, 2017; Kirkpatrick et al. 2016), and ∼70 L-subdwarfs (e.g. Burgasser et al. 2003; Kirkpatrick et al. 2014; Zhang et al. 2017, 2018) have been discovered with modern sky surveys, often detected from their high proper motion.

Hejazi et al. (2018) used the accurate Gaia DR2 parallaxes of ∼1600 high proper motion M-dwarfs and subdwarfs with spectroscopic metallicity measurements to draw the M_G vs $G_{BP} - G_{RP}$ diagram. Stars with different metallicity ranges fall into clearly distinct loci which can be used to develop photometric metallicity calibrations, in particular for metal-poor M subdwarfs.

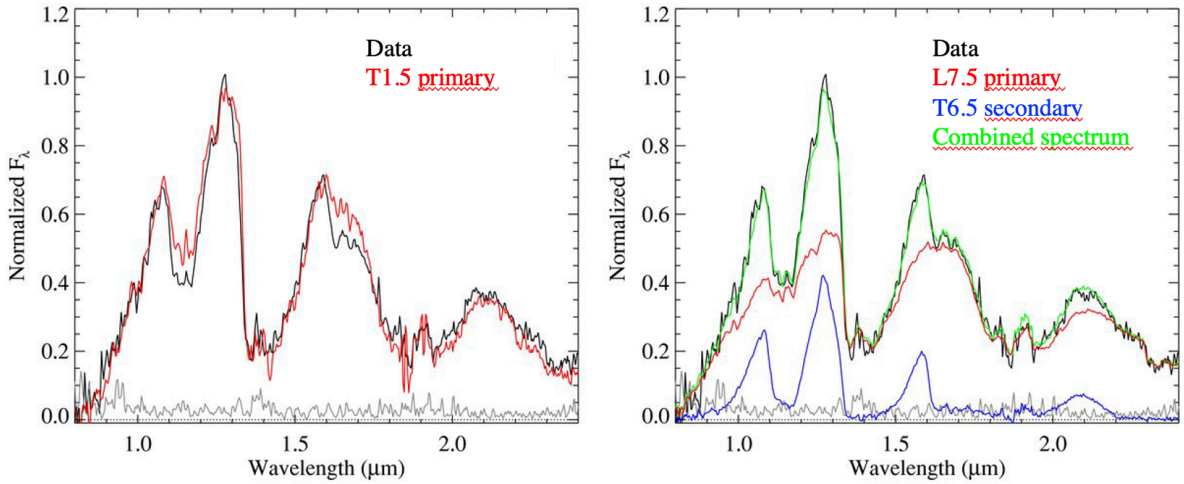


Fig. 10. Spectral binary candidate. The left panel shows the best-fit single template (red line) compared to the source spectrum (black line). The right panel shows the best-fit composite (green line), primary (red line), and secondary spectra (blue line). Adapted from Burgasser et al. (2010).

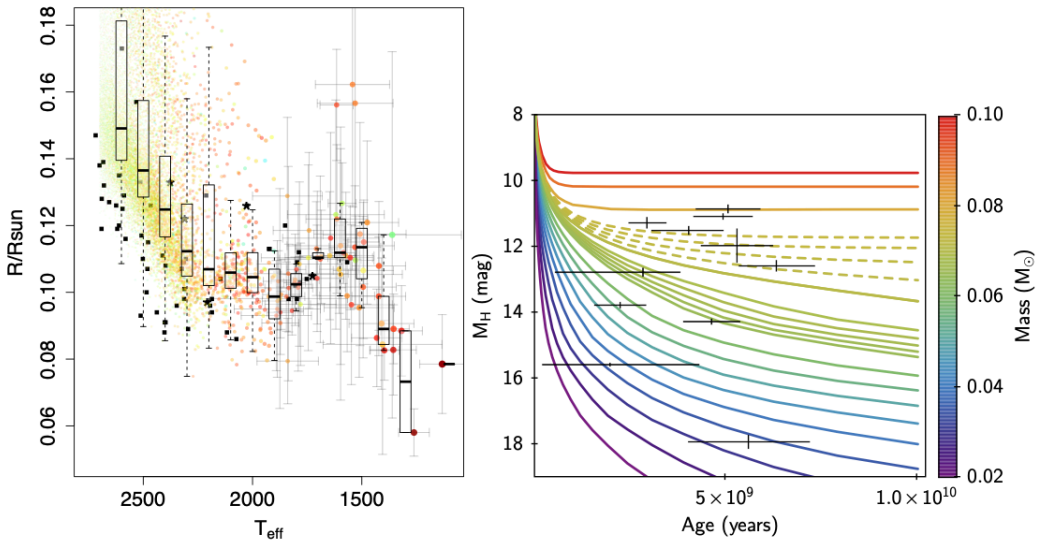


Fig. 11. Left: Radii of candidate UCDs in the Gaia golden sample (colored symbols), superimposed to the data of Dieterich et al. (2014) (black symbols). The box plots are calculated within bins of 100 K. Right: Evolutionary tracks and UCD (in binary systems) locations in the H-band absolute magnitude vs age diagram, adopting the companion age. The tracks are color coded by mass. The dashed lines indicate the stellar to substellar transition zone (from 0.072 to 0.075 $M_\odot$). From Gaia Collaboration, Creevey et al. (2023)

Zhang et al. (2018) selected L subdwarfs from their locus in optical and near infrared color-color diagrams, and confirmed their nature with spectroscopic follow-up. They computed the tangential velocity for 20 of them found in the Gaia DR2 catalogue. They found that the sdL subclass mostly have thick disc kinematics, whereas the esdL and usdL subclasses generally have halo kinematics, which is consistent to the esdM/usdM subclasses.

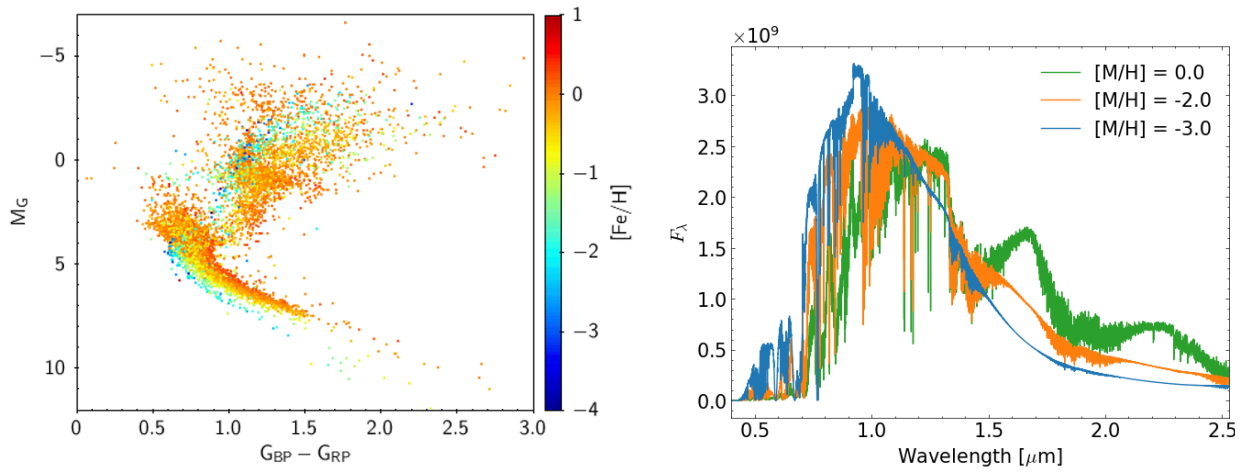


Fig. 12. Left: Color absolute magnitude diagram of Gaia DR2 stars found in the PASTEL catalogue (Soubiran et al. 2016) wich provides accurate metallicity measurements. The low-metallicity main sequence stars (in blue) form a shifted, lower, sequence. Right: Synthetic spectra computed from BT-Settl atmosphere models with $T_{\rm eff}$ = 2600 K, $\log g$ = 5, and various metallicity values.

5.3 Wide binaries

Multiplicity with Gaia can be determined from distance, angular separation and proper motion measurements (e.g. Hartman & Lépine 2020; El-Badry et al. 2021; Gaia Collaboration, Smart et al. 2021; Sarro et al. 2023). Several thousands of UCDs in binary systems have been discovered this way. They are useful benchmarks for testing stellar evolutionary models, and can be used as calibrators for age and metallicity relations from their spectroscopic characterization (e.g. Marocco et al. 2020; Low et al. 2021).

Zhang (2019) discovered the first wide M1 + L0 extreme subdwarf binary, Gaia J0452-36AB. Its kinematics is compatible with the halo population, its metallicity is about -1.4, and the projected separation is 15 282 AU. How such old wide systems can survive is still an open question, as well as other ultra-wide systems (González-Payo et al. 2023).

5.4 Young associations

The selection of members of young clusters or associations has become trivial with Gaia, but also the discovery of new groups, using the parallax, position, and motion of stars (e.g. Cantat-Gaudin et al. 2019; Galli et al. 2019, 2020a,b; Mužić et al. 2022; Tarricq et al. 2021, 2022; Sarro et al. 2023). Gaia is sensitive to 30 $M_{\rm Jup}$ up to about 500 pc in young groups. To summarize the conclusions of these numerous studies of young clusters in the solar neighborhood, the mass function in the substellar regime does not vary much, at least within the error bars (Hervé Bouy, priv. comm.).

Gaia has furthermore enabled to measure dynamic ages for some of these regions, which is very useful because age is often poorly constrained in these young regions, since evolutionary models are not very reliable for such early ages. This methodology uses the present 3D positions and 3D velocities of individual stars and computes the stellar orbits back in time with a Galactic potential. The dynamical traceback age is the time when a group of stars was most concentrated in the past, that is, when the size of the group was at its minimum (Figure 13). These model-independent ages are therefore particularly interesting. This is particularly important in the context of brown dwarfs whose brightness is very dependent on age and varies rapidly at the beginning of their life, leading to large uncertainties. Miret-Roig et al. (2022) showed that Upper Scorpius and Ophiuchus groups share a common origin. The proposed star formation scenario is likely a result of stellar feedback from massive stars, supernova explosions, and dynamic interactions between stellar groups and the molecular gas.

Using the sample of 94 000 UCD candidates (selected from their Gaia $T_{\rm eff}$), Sarro et al. (2023) performed group identification using a hierarchical mode association clustering classification algorithm. Detection and characterization of overdensities are done in the space of celestial coordinates, projected velocities, and parallaxes (Figure 14). They compared the medium RP spectrum in $T_{\rm eff}$ bins for sources outside these associations (Main

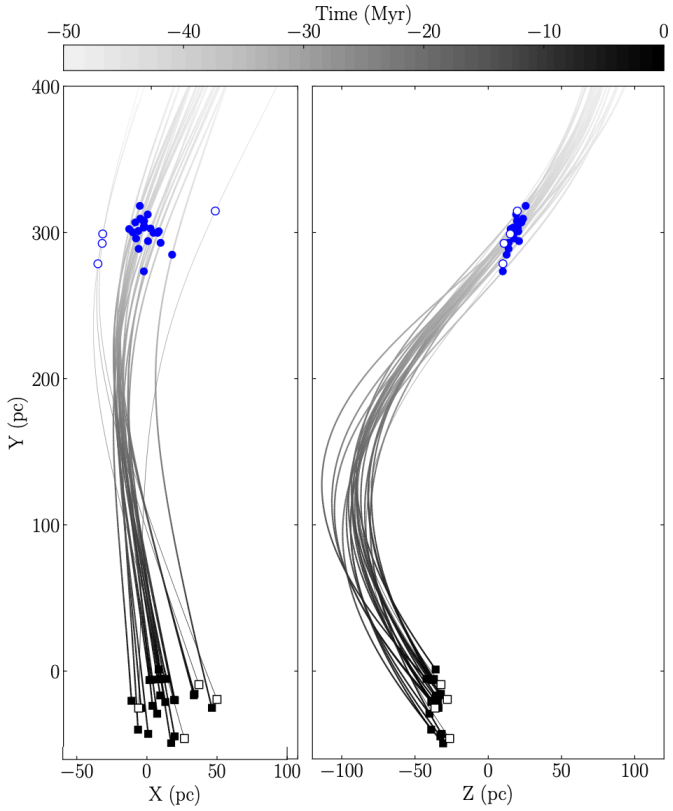


Fig. 13. 2D projection of the stellar orbits from 25 stars in the Tucana-Horologium young stellar association. The black squares indicate the present-day position of the stars and the blue circles denote the stellar positions at birth time, 38.5 Myr ago. The orbits are color-coded based on the traceback time. Adapted from Galli et al. (2023).

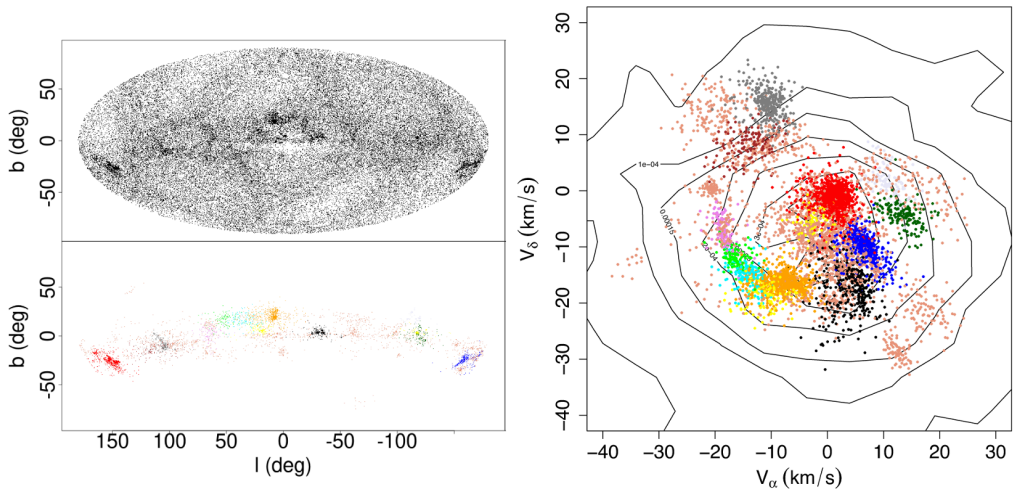


Fig. 14. Left: Distribution in Galactic coordinates of the Gaia UCD catalogue using the Aitoff projection. Clusters with more than 10 members are depicted with colors. Right: Distribution in the space of tangential velocities of the sources in clusters with more than 10 members with the same color code. From Sarro et al. (2023).

Sequence) and in several clusters. There is a systematic trend of increasing band depths as the association becomes older. By 10 Myr, the RP spectra of young and Main Sequence sources becomes undistinguishable. This statement can be used as a basis for the indication of youth in the forthcoming Gaia spectra.

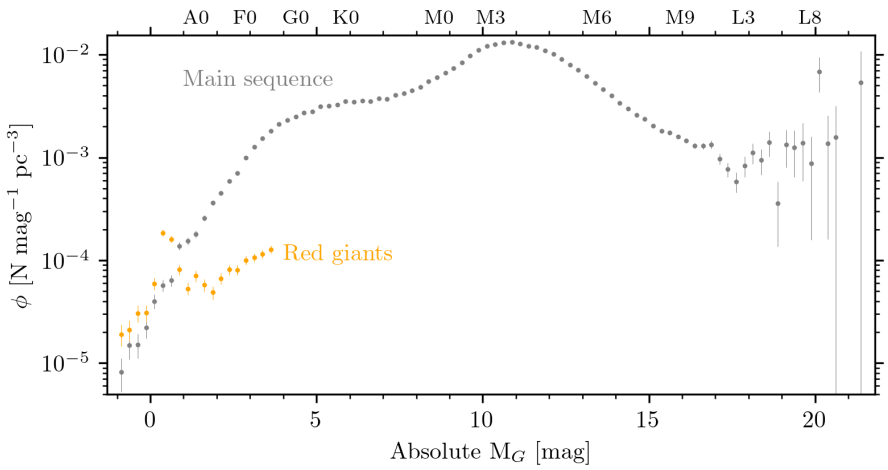


Fig. 15. Luminosity function of the GCNS, with a 0.25 bin, in log scale. The upper full curve plotted in grey shows main sequence stars. The small lower partially orange curve shows giants stars.

6 Towards a complete local census

Because Gaia revealed a huge number of UCDs, it constitutes a step towards a complete local census. The nearby sample is particularly important for the ultra-cool dwarfs which are the lowest-mass, coldest, and faintest products of star formation, making them difficult to study at large distances. Having a volume-complete sample, with good statistics, is crucial to compute precise bias-free densities, and to therefore determine luminosity, and mass functions that will bring strong constraints on stellar and substellar formation theories.

For substellar objects with no obvious mass-to-luminosity relation, the luminosity function is simulated assuming different initial mass fonctions and birth rates. Thus the comparison with the observed luminosity function allows to disentangle between the different formation scenarios, only if error bars are small enough.

Field stars and brown dwarfs are several Gyr in average. Brown dwarfs depopulate rapidly earlier spectral types to go to later ones. As a consequence, a minimum is expected in the density at the stellar substellar boundary, as predicted by simulations (Burgasser 2004; Allen et al. 2005). Observations may indicate such a minimum in the luminosity function (Cruz et al. 2007; Bardalez Gagliuffi et al. 2019). The GCNS is an exquisite dataset from which to derive the local luminosity function. This is possible for the first time using volume-limited samples with parallaxes not derived from photometric measurements that are affected by related biases (Eddington or Malmquist), from bright stars down to the substellar regime. A clear dip around the L3 spectral type is visible, that is interpreted as the stellar-substellar boundary.

7 Conclusions

There are numerous ultra-cool dwarves revealed in the successive Gaia data releases. They offer high number, high precision, information on these objects. Photometry and astrometry are so precise that the locus in the color absolute magnitude diagram can give indication on the nature of the object: youth, binary, low metallicity. In addition, Gaia offers kinematical and astrophysical parameters parameters for a large part of the catalogue. A well-characterized sample with spectroscopic follow-up is powerful to test (sub)stellar models (evolution, interior, atmosphere). A well-characterized and complete volume-limited sample provides luminosity and mass functions free of biases that plagued previous determinations, provides new insights on the stellar-substellar limit, provides strong constraints on stellar and substellar formation theories. We expect the forthcoming Gaia data releases and their additional parameters to continue the story of these puzzling and elusive objects.

This work has made use of data from the European Space Agency (ESA) mission *Gaia* (https://www.cosmos.esa.int/gaia), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, https://www.cosmos.esa.int/web/gaia/dpac/consortium). Funding for the DPAC has been provided by national institutions, in particular, the institutions participating in the *Gaia* Multilateral Agreement.

References

Allard, F. 2014, in Exploring the Formation and Evolution of Planetary Systems, ed. M. Booth, B. C. Matthews, & J. R. Graham, Vol. 299, 271–272

Allard, F., Homeier, D., & Freytag, B. 2013, Mem. Soc. Astron. Italiana, 84, 1053

Allen, P. R., Koerner, D. W., Reid, I. N., & Trilling, D. E. 2005, ApJ, 625, 385

André, P., Men'shchikov, A., Bontemps, S., et al. 2010, A&A, 518, L102

Anglada-Escudé, G., Amado, P. J., Barnes, J., et al. 2016, Nature, 536, 437

Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, Annual Review of Astronomy and Astrophysics, 47, 481, aDS Bibcode: 2009ARA&A..47..481A

Attwood, R. E., Goodwin, S. P., Stamatellos, D., & Whitworth, A. P. 2009, A&A, 495, 201

Baraffe, I., Homeier, D., Allard, F., & Chabrier, G. 2015, A&A, 577, A42

Bardalez Gagliuffi, D. C., Burgasser, A. J., Gelino, C. R., et al. 2014, ApJ, 794, 143

Bardalez Gagliuffi, D. C., Burgasser, A. J., Schmidt, S. J., et al. 2019, ApJ, 883, 205

Bate, M. R., Bonnell, I. A., & Bromm, V. 2002, MNRAS, 332, L65

Bonfils, X., Astudillo-Defru, N., Díaz, R., et al. 2018, A&A, 613, A25

Bonnell, I. A., Clark, P., & Bate, M. R. 2008, MNRAS, 389, 1556

Bowler, B. P. 2016, PASP, 128, 102001

Boyajian, T. S., von Braun, K., van Belle, G., et al. 2012, ApJ, 757, 112

Burgasser, A. J. 2004, ApJS, 155, 191

Burgasser, A. J. 2007, ApJ, 659, 655

Burgasser, A. J., Cruz, K. L., Cushing, M., et al. 2010, ApJ, 710, 1142

Burgasser, A. J., Cruz, K. L., & Kirkpatrick, J. D. 2007, ApJ, 657, 494

Burgasser, A. J., Kirkpatrick, J. D., Burrows, A., et al. 2003, ApJ, 592, 1186

Burgasser, A. J. & Splat Development Team. 2017, in Astronomical Society of India Conference Series, Vol. 14, Astronomical Society of India Conference Series, 7–12

Burrows, A. 1999, Science, 285, 847

Burrows, A., Heng, K., & Nampaisarn, T. 2011, ApJ, 736, 47

Burrows, A., Hubbard, W. B., Lunine, J. I., & Liebert, J. 2001, Reviews of Modern Physics, 73, 719

Caffau, E., Ludwig, H. G., Steffen, M., Freytag, B., & Bonifacio, P. 2011, Solar Physics, 268, 255, aDS Bibcode: 2011SoPh..268..255C

Cantat-Gaudin, T., Krone-Martins, A., Sedaghat, N., et al. 2019, A&A, 624, A126

Carmichael, T. W., Quinn, S. N., Zhou, G., et al. 2021, AJ, 161, 97

Carrasco, J. M., Weiler, M., Jordi, C., et al. 2021, A&A, 652, A86

Chabrier, G. & Baraffe, I. 2000, ARA&A, 38, 337

Chabrier, G., Baraffe, I., Leconte, J., Gallardo, J., & Barman, T. 2009, in American Institute of Physics Conference Series, Vol. 1094, 15th Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun, ed. E. Stempels, 102–111

Chabrier, G., Baraffe, I., Phillips, M., & Debras, F. 2023, A&A, 671, A119

Cooper, W. J., Smart, R. L., Jones, H. R. A., & Sarro, L. M. 2024, MNRAS, 527, 1521

Creevey, O. 2024, EES 2023, These proceedings

Creevey, O. L., Sordo, R., Pailler, F., et al. 2023, A&A, 674, A26

Cruz, K. L., Burgasser, A. J., Reid, I. N., & Liebert, J. 2004, ApJ, 604, L61

Cruz, K. L., Reid, I. N., Kirkpatrick, J. D., et al. 2007, AJ, 133, 439

De Angeli, F., Weiler, M., Montegriffo, P., et al. 2023, A&A, 674, A2

Delfosse, X., Tinney, C. G., Forveille, T., et al. 1997, A&A, 327, L25

Dieterich, S. B., Henry, T. J., Jao, W.-C., et al. 2014, AJ, 147, 94

Dupuy, T. J. & Liu, M. C. 2017, ApJS, 231, 15

El-Badry, K., Rix, H.-W., & Heintz, T. M. 2021, MNRAS, 506, 2269

Evans, D. W., Riello, M., De Angeli, F., et al. 2018, A&A, 616, A4

Gaia Collaboration, Babusiaux, C., van Leeuwen, F., Barstow, M. A., et al. 2018, A&A, 616, A10

Gaia Collaboration, Brown, A. G. A., Vallenari, A., Prusti, T., et al. 2018, A&A, 616, A1

Gaia Collaboration, Brown, A. G. A., Vallenari, A., Prusti, T., et al. 2021, A&A, 649, A1

Gaia Collaboration, Creevey, O. L., Sarro, L. M., Lobel, A., et al. 2023, A&A, 674, A39

Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., Brown, A. G. A., et al. 2016, A&A, 595, A1
Gaia Collaboration, Smart, R. L., Sarro, L. M., Rybizki, J., et al. 2021, A&A, 649, A6
Gaia Collaboration, Vallenari, A., Brown, A. G. A., Prusti, T., et al. 2023, A&A, 674, A1
Galli, P. A. B., Bouy, H., Olivares, J., et al. 2020a, A&A, 634, A98
Galli, P. A. B., Bouy, H., Olivares, J., et al. 2020b, A&A, 643, A148
Galli, P. A. B., Loinard, L., Bouy, H., et al. 2019, A&A, 630, A137
Galli, P. A. B., Miret-Roig, N., Bouy, H., Olivares, J., & Barrado, D. 2023, MNRAS, 520, 6245
Gillon, M., Jehin, E., Lederer, S. M., et al. 2016, Nature, 533, 221
Gillon, M., Triaud, A. H. M. J., Demory, B.-O., et al. 2017, Nature, 542, 456
Gizis, J. E. 1997, AJ, 113, 806
González-Payo, J., Caballero, J. A., & Cortés-Contreras, M. 2023, A&A, 670, A102
Goodwin, S. P. & Whitworth, A. P. 2004, A&A, 413, 929
Grieves, N., Bouchy, F., Lendl, M., et al. 2021, A&A, 652, A127
Hartman, Z. D. & Lépine, S. 2020, ApJS, 247, 66
Hatzes, A. P. & Rauer, H. 2015, ApJ, 810, L25
Hayashi, C. & Nakano, T. 1963, Progress of Theoretical Physics, 30, 460
Hejazi, N., Lepine, S., & Homeier, D. 2018, in 20th Cambridge Workshop on Cool Stars, Stellar Systems and the Sun, Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun, 91
Hennebelle, P. & Chabrier, G. 2008, ApJ, 684, 395
Jao, W.-C., Henry, T. J., Gies, D. R., & Hambly, N. C. 2018, ApJ, 861, L11
Kirkpatrick, J. D., Cruz, K. L., Barman, T. S., et al. 2008, ApJ, 689, 1295
Kirkpatrick, J. D., Henry, T. J., & Irwin, M. J. 1997, AJ, 113, 1421
Kirkpatrick, J. D., Kellogg, K., Schneider, A. C., et al. 2016, ApJS, 224, 36
Kirkpatrick, J. D., Schneider, A., Fajardo-Acosta, S., et al. 2014, ApJ, 783, 122
Kumar, S. S. 1963, ApJ, 137, 1121
Lecavelier des Etangs, A. & Lissauer, J. J. 2022, New A Rev., 94, 101641
Lépine, S., Rich, R. M., & Shara, M. M. 2003, AJ, 125, 1598
Lépine, S., Rich, R. M., & Shara, M. M. 2007, ApJ, 669, 1235
Lépine, S. & Scholz, R.-D. 2008, ApJ, 681, L33
Lodieu, N., Espinoza Contreras, M., Zapatero Osorio, M. R., et al. 2012, A&A, 542, A105
Lodieu, N., Espinoza Contreras, M., Zapatero Osorio, M. R., et al. 2017, A&A, 598, A92
Lomax, O., Whitworth, A. P., & Hubber, D. A. 2016, MNRAS, 458, 1242
Looper, D. L., Gelino, C. R., Burgasser, A. J., & Kirkpatrick, J. D. 2008, ApJ, 685, 1183
Low, R., Burgasser, A. J., Reylé, C., et al. 2021, Research Notes of the American Astronomical Society, 5, 26
Marley, M. S. & Leggett, S. K. 2009, in Astrophysics and Space Science Proceedings, Vol. 10, Astrophysics in the Next Decade, 101
Marley, M. S., Saumon, D., Visscher, C., et al. 2021, ApJ, 920, 85
Marocco, F., Smart, R. L., Mamajek, E. E., et al. 2020, MNRAS, 494, 4891
Miret-Roig, N., Galli, P. A. B., Olivares, J., et al. 2022, A&A, 667, A163
Mužić, K., Almendros-Abad, V., Bouy, H., et al. 2022, A&A, 668, A19
Nakajima, T., Oppenheimer, B. R., Kulkarni, S. R., et al. 1995, Nature, 378, 463
Padoan, P. & Nordlund, Å. 2004, ApJ, 617, 559
Phillips, M. W., Tremblin, P., Baraffe, I., et al. 2020, Astronomy & Astrophysics, 637, A38
Rajpurohit, A. S., Reylé, C., Allard, F., et al. 2013, A&A, 556, A15
Ravinet, T., Reylé, C., Lagarde, N., et al. 2024, arXiv e-prints, arXiv:2402.09783
Rebolo, R., Zapatero Osorio, M. R., & Martín, E. L. 1995, Nature, 377, 129
Reipurth, B. & Mikkola, S. 2015, AJ, 149, 145
Reylé, C. 2018, A&A, 619, L8
Reylé, C., Jardine, K., Fouqué, P., et al. 2021, A&A, 650, A201
Reylé, C., Jardine, K., Fouqué, P., et al. 2022, in The 21st Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun, Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun, 218
Ruiz, M. T., Leggett, S. K., & Allard, F. 1997, ApJ, 491, L107

Sarro, L. M., Berihuete, A., Carrión, C., et al. 2013, A&A, 550, A44
Sarro, L. M., Berihuete, A., Smart, R. L., et al. 2023, A&A, 669, A139
Saumon, D., Chabrier, G., & van Horn, H. M. 1995, ApJS, 99, 713
Saumon, D. & Marley, M. S. 2008, ApJ, 689, 1327
Smart, R. L., Marocco, F., Caballero, J. A., et al. 2017, MNRAS, 469, 401
Soubiran, C., Le Campion, J.-F., Brouillet, N., & Chemin, L. 2016, A&A, 591, A118
Stamatellos, D. & Herczeg, G. J. 2015, MNRAS, 449, 3432
Stamer, T. & Inutsuka, S.-i. 2019, MNRAS, 488, 2644
Tarricq, Y., Soubiran, C., Casamiquela, L., et al. 2021, A&A, 647, A19
Tarricq, Y., Soubiran, C., Casamiquela, L., et al. 2022, A&A, 659, A59
Vorobyov, E. I. & Basu, S. 2010, ApJ, 719, 1896
Whitworth, A. P. & Zinnecker, H. 2004, A&A, 427, 299
Witte, S., Helling, C., Barman, T., Heidrich, N., & Hauschildt, P. H. 2011, Astronomy and Astrophysics, 529, A44, aDS Bibcode: 2011A&A...529A..44W
Zhang, Z. 2019, MNRAS, 489, 1423
Zhang, Z. H., Galvez-Ortiz, M. C., Pinfield, D. J., et al. 2018, MNRAS, 480, 5447
Zhang, Z. H., Pinfield, D. J., Gálvez-Ortiz, M. C., et al. 2017, MNRAS, 464, 3040

École Evry Schatzman 2023: Stellar physics with Gaia
C. Babusiaux, C. Reylé (eds)

ASTROPHYSICAL PARAMETERS ASSOCIATED TO 'HOT' STARS IN GAIA DR3

Y. FRÉMAT

Royal observatory of Belgium, 3 avenue circulaire, 1180 Brussels, Belgium
e-mail: yves.fremat@observatory.be

Abstract. Hot stars, of spectral types O-, B-, and A-, represent a small fraction of the stars observed by the Gaia satellite. Their properties and the specifications of the on-board instruments make their identification challenging. In the Gaia DR3, 12 104 577 targets have been assigned a temperature greater than 7500 K. It represents ∼2.5% of the stars having an effective temperature value in the Gaia DR3 `astrophysical_parameters` table. We review the results obtained by the Apsis modules, focusing on the effective temperature, surface gravity, interstellar extinction, and V sin i parameters.

Keywords: stars: fundamental parameters - Galaxy: stellar content - dust, extinction - catalogs

1 Introduction

In this contribution, we deal with and refer to 'hot' stars, defined as stars with an effective temperature ($T_{\rm eff}$) greater than 7500 K and of spectral types O, B, or A (OBA). They are located on the upper main sequence, beyond the onset $T_{\rm eff}$ of strong convection, and are, on average, rotating significantly faster than cooler stars. For this reason, they are good laboratories to study stellar evolution at high angular rates. Usually more massive than two solar masses, hot stars also have a higher luminosity and can be observed from large distances. They belong to young stellar populations often embedded in open clusters and star forming regions found in the galactic disc, which makes their light highly absorbed by interstellar gas and dust (i.e., ISM, interstellar medium). All along their evolution, they interact dynamically and chemically with the surrounding ISM, ionising its particles and enriching its chemical composition. OBA stars are important to identify among the huge amount of Gaia data, as they are a key to a better understanding of the galaxy's chemical content, structure, evolution, and dynamics.

Their Gaia spectra are dominated by broad hydrogen lines (Balmer series in the BP/RP and Paschen series in the RVS), by the near-IR calcium triplet lines (in the RVS), and by the Balmer jump lying at the very blue end of the BP passband. These features, however, decrease very rapidly with effective temperature, leaving an almost featureless spectral energy distribution (SED, see Figs. 4 and 6 of Fouesneau et al. 2023) and making their analysis challenging (e.g., Blomme et al. 2023). The paper focuses on the astrophysical parameters (APs) of hot stars available in the Gaia DR3 `astrophysical_parameters` table (Creevey et al. 2023). We provide a rapid overview of their behaviour and limits of applicability.

2 Classification of hot stars by the Apsis pipeline

Two methods are providing astrophysical parameters for stars with $T_{\rm eff} > 7500$ K: the General Stellar Parametrizer from Photometry (GSP-Phot) and the Extended Stellar Parametrizer for Hot Stars (ESP-HS). Both modules make different assumptions and analyse epoch-combined spectroscopic data (internally calibrated BP/RP and RVS). While GSP-Phot is aimed at treating all the stars, whatever their effective temperature, ESP-HS focuses specifically on the treatment of hot stars. Their G magnitude coverage is also different(see Fig. 2 of Creevey et al. 2023); the former is treating stars brighter than $G = 19$, the latter those with $G \leq 17.65$.

GSP-Phot uses a Markov Chain Monte Carlo (MCMC) approach to derive the $T_{\rm eff}$, surface gravity ($\log g$), metallicity ([M/H]), and interstellar extinction parameters (A_0, $A_{\rm G}$, ...). Its results are based on the analysis of

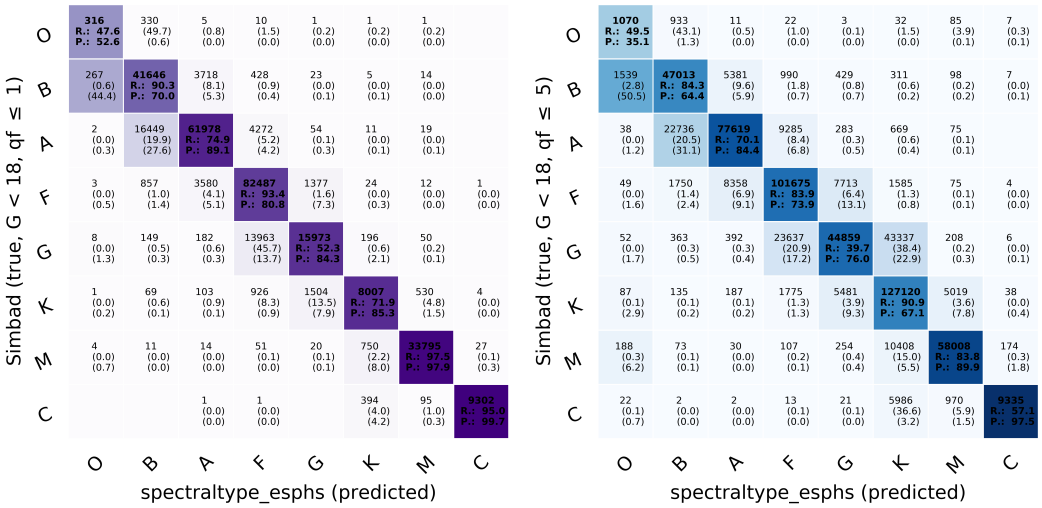


Fig. 1. Confusion matrix of `spectraltype_esphs` in Gaia DR3 (abscissa) with the spectral type label (ordinate) available in the Simbad database (CDS). The top number in each cell gives the number of targets of a given Simbad spectral type that received a certain predicted tag. The second and third numbers give the corresponding percentage along the row and column, respectively. The colour scale follows the column percentage (i.e., the third number). The right panel takes all the targets into account, whatever the quality flag (qf) value is. In the left panel, we only consider those stars classified with the highest probability.

the sampled BP/RP spectra, the apparent G magnitude, and the parallax, which are confronted with simulated spectra, photometry, and isochrones. GSP-Phot considers four different synthetic spectral libraries individually (i.e., APs per library are saved in table `astrophysical_parameters_supp`), then determines which one provides the best fit (i.e., best fit APs are saved in table `astrophysical_parameters`, with field names having the suffix `_gspphot`). More information about the algorithms is available in the dedicated paper of Andrae et al. (2023).

The ESP-HS algorithm is based on a minimum distance approach to derive $T_{\rm eff}$, $\log g$, $\mathrm{V}\sin i$ (when RVS data of sufficient quality is available), and A_0 (i.e., APs are saved in table `astrophysical_parameters`, with field names having the suffix `_esphs`). In the first step, it uses a simplex downhill algorithm to fit BP/RP and, when available, RVS simulated spectra to the observations. The distinction between these two processing modes is made in the first digit of the flag `esphs_flags`: 0 (BP/RP+RVS fit) and 1 (BP/RP-only fit). The algorithm assumes a solar chemical composition, keeping [M/H] at zero. In the process, both passbands are normalised separately. In a second step, a Levenberg-Marquadt algorithm is applied to fine-tune the solution and derive the uncertainties. The synthetic spectra used to fit the BP/RP observations are taken from a library computed with the `LLmodels` code (7500 to 20 000 K: Shulyak et al. 2004) and, for the hotter stars, with the `synspec/tlusty` computer programme (Hubeny 1988) adopting the BSTAR (Lanz & Hubeny 2007) and OSTAR (Lanz & Hubeny 2003) atmosphere models. The corresponding simulations were then generated with the SMSGen tool (Sampled Mean Spectrum Generator; see Sect. 2.3.2 in Creevey et al. 2023). Systematic mismatches found in the comparison of these simulations and observations for a sample of B and A-type stars with known APs derived from Stromgren photometry led to the masking of the spectrum above 800 nm and to the semi-empirical correction of the BP/RP data (mainly below 400 nm). The synthetic spectra used to model the RVS data were convolved with a Gaussian spectroscopic line spread function (LSF) and by assuming a median resolving power of $\mathscr{R} = 11\,500$ (e.g. Katz et al. 2023). In practice, the co-added RVS spectra analysed by the Apsis modules are a mixture of epoch spectra, each having their own LSF. This LSF is known to be non-Gaussian and may have $\mathscr{R}$ values significantly different from the documented median. It was the intention to use the fitting of the rotational broadening parameter to compensate for any deviation from the adopted assumption.

Unlike GSP-Phot, ESP-HS is not applied to all the data. It is launched on the O, B, and A-type stars flagged by a random forest algorithm trained on a data sample made of simulated and observed (see Appendix A.4. of Creevey et al. 2023) BP/RP spectra. Because it is used to pre-select the hot stars, we found it useful to save the spectral type tag in the Gaia DR3 catalogue (field `spectraltype_esphs`, whose definition and temperature

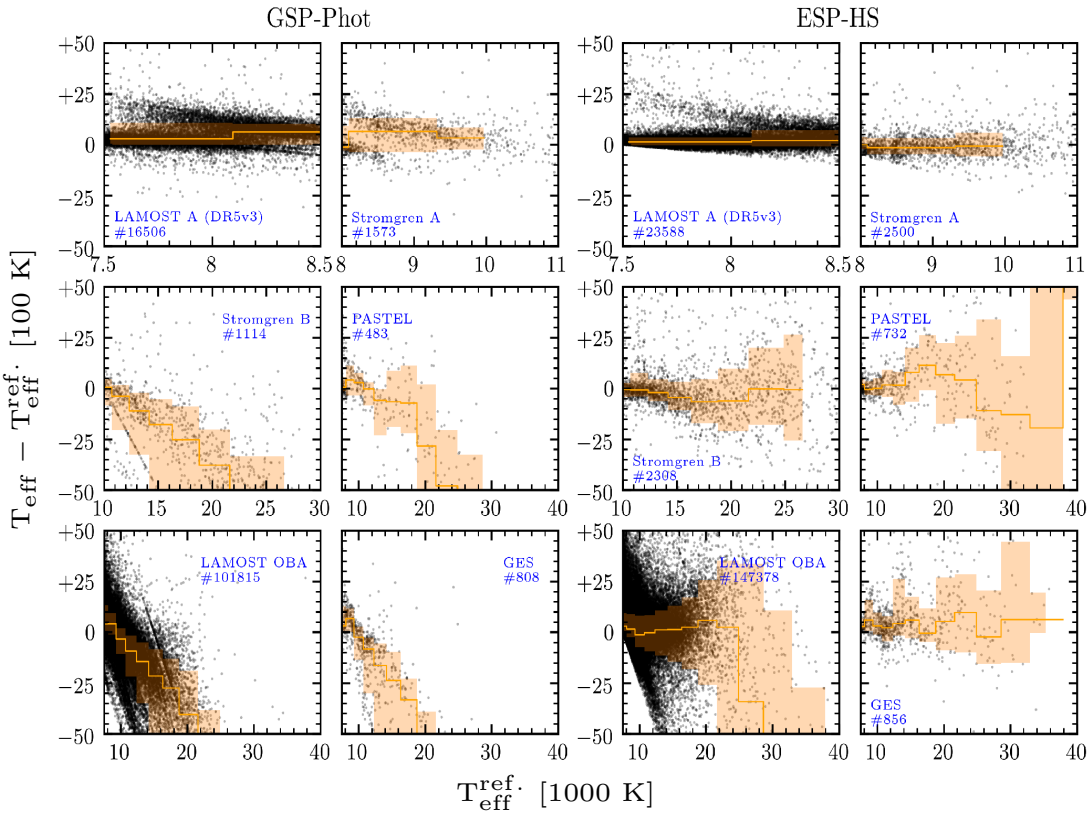


Fig. 2. Effective temperature residuals obtained by GSP-Phot (left panel) and ESP-HS (right panel) relative to various reference catalogues are plotted against the reference effective temperature. Considered catalogues and surveys are: LAMOST A catalogue (Luo et al. 2019, DR5v3), Stromgren (derived within DPAC adopting the updated calibration of Napiwotzki et al. 1993), PASTEL (Soubiran et al. 2016), LAMOST OBA (Xiang et al. 2022, DR6), Gaia ESO Survey (GES: Blomme et al. 2022).

dependance are given in Table A.1. of Creevey et al. 2023). We show in Fig. 1 the confusion matrix obtained by correlating the `spectraltype_esphs` tag to the spectral type information found in Simbad at the CDS. Because the temperature coverage of the various spectral types may differ, confusion between neighbouring types is expected. Globally, we obtain a fair level of consistency between both classifications. The O-type star classification is usually less accurate. These are often confused with B-type stars, and, at G magnitudes fainter than 15, a significant fraction of M stars are tagged as (highly reddened) O-type stars. A quality assessment is also provided with the tag. It is saved as the second digit of the field `esphs_flags` (i.e., `esphs_flags`[1:2], with integer values ranging from 1 to 5). The lower the value of this quality flag, the closer the observed spectrum is expected to match the training sample. Finally, except for the candidate Be stars identified by the ESP-ELS algorithm, ESP-HS is not processing targets with significant Hα line emissions.

3 Astrophysical parameters

GSP-Phot and ESP-HS provided for Gaia DR3 an estimate of the astrophysical parameters (i.e., $T_{\rm eff} > 7500$K) of 11 156 494 and 2 382 015 targets, respectively. The overlap between the two samples comprises 1 433 932 objects. The different biases, magnitude domains, and post-processing criteria explain why the ESP-HS sample is not completely included in GSP-Phot.

In order to assess their accuracy and precision, the $T_{\rm eff}$ and $\log g$ results were compared with those found in various catalogues. These catalogues were chosen for having, when combined, temperature and magnitude domains matching those of the Apsis' OBA sample. Figure 2 shows how the $T_{\rm eff}$ residuals (i.e., measured minus

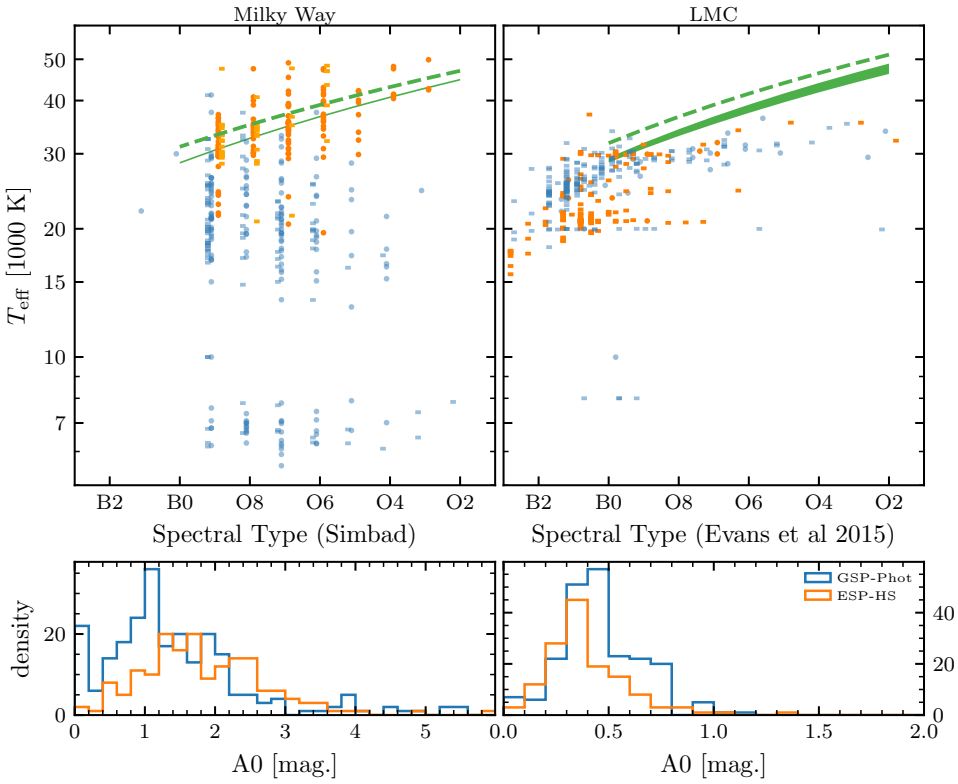


Fig. 3. Upper panels: Effective temperature derived by GSP-Phot (blue symbols) and ESP-HS (orange symbols) plotted as a function of the spectral type found in Simbad (left panel for galactic stars) and published by Evans et al. (2015) (LMC stars in right panel). The spectral type vs. $T_{\rm eff}$ calibrations found by Weidner & Vink (2010) for Milky Way and LMC dwarfs (broken green curve) and giants (solid green curve) are compared to the Apsis measurements. **Lower panels:** distribution of the A_0 extinction in the galaxy (left) and LMC (right) derived by the two Apsis modules.

expected value) vary with the catalogue $T_{\rm eff}$ estimate. In the A-type star domain, the two modules provide consistent $T_{\rm eff}$ measurements with similar, usually positive, biases (from -50 to 300 K) and spread (200 to 600 K). The sequence of positive outlying residuals seen in the 'LAMOST A' panels of Fig. 2 might be related to the combined effect of the parameter accuracy and of the hydrogen lines having their maximum intensity around 8500 K (Xiang et al. 2022). At hotter effective temperatures, ESP-HS measurements tend to be underestimated by a few thousand K above 25 000 K while the scatter also increases with the effective temperature (from ∼2000 K to ∼7000 K for the early O-type stars). The temperatures provided by GSP-Phot tend to deviate systematically above 10 000 K with negative biases. Therefore, it can be considered less reliable than ESP-HS for B-type stars. For O-type stars, we may have a better idea of the trends by comparing the APs to the Galactic and LMC spectral types vs. temperature scale (Fig. 3, upper panels). In our galaxy, ESP-HS usually derives a larger interstellar extinction (A0, lower panels of Fig. 3) and $T_{\rm eff}$, that better matches the calibration determined by Weidner & Vink (2010). GSP-Phot obtains systematically lower estimates due to non-adapted priors and a difficulty in selecting the 'best library' fit. For Large Magellanic Cloud (LMC) stars, in which direction the interstellar extinction is lower, the conclusions are reversed. GSP-Phot tends to provide a temperature scale closer to the one of Weidner & Vink (2010) for the LMC.

The distribution of the surface gravity residuals with $T_{\rm eff}$ is shown in Fig. 4. The $\log g$ offsets differ from one module to the other, usually by 0.1 dex in the 7500 - 10 000 K temperature domain, with ESP-HS providing larger estimates. When comparing the results to the GES survey, the behaviour of the residuals significantly and consistently changes around 10 000 K in the two datasets. This behaviour is also seen in the comparison with benchmark stars in Blomme et al. (Fig. 3, 2022). As explained by the same authors and also discussed by Xiang et al. (2022), it could be due to the Balmer lines reaching their maximum strength around 8500 K. Above this $T_{\rm eff}$, GSP-Phot results are impacted by the offsets found in temperature and are usually underestimated.

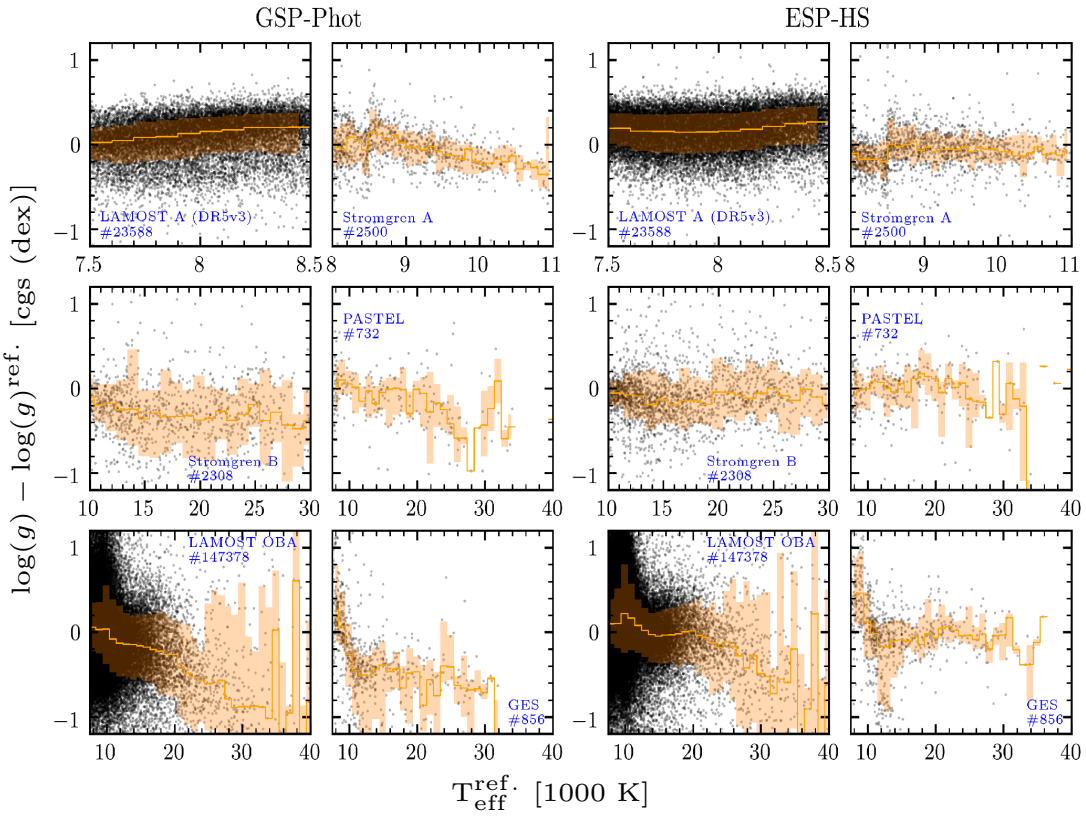


Fig. 4. Same as Fig. 2 for the surface gravity.

On the other hand, ESP-HS $\log g$ measurements show a constant bias that varies from -0.1 to $+0.1$ dex, mainly depending on the origin of the comparison values. The scatter is of the order of 0.2 to 0.4 dex in the considered magnitude range.

ESP-HS results were also confronted with the reddened PARSEC[1] isochrones (Bressan et al. 2012; Chen et al. 2015) of open clusters. In Fig. 5, the membership, extinction and ages, t, are taken from Cantat-Gaudin et al. (2020). The effective temperature offset (in BP/RP+RVS mode) is of the order of -300 K for A-type stars, $+400$ K for B-type stars, and reaching $+6000$ K in the O-type stars' $T_{\rm eff}$ range. The $\log g$ median bias is of the order of -0.02 to -0.08 dex. It is usually negative and constant with temperature. Compared to the spectroscopic ground-based estimates, the $T_{\rm eff}$ offset has a different sign. The difference due to the processing mode becomes significant above 10 000 K, where the BP/RP-only (`flags_esphs[0:1]` = 1) processing produces a somewhat larger $T_{\rm eff}$ offset ($+1200$ K in the 20000 - 30000 K domain) than the simultaneous fitting of BP/RP and RVS ($+700$ K in the 20000 - 30000 K domain). The dispersion remains similar in both cases.

From the comparison of the Apsis results to the temperature scale found in the Milky Way and LMC O-type stars (Fig. 3), the degeneracy between temperature and interstellar extinction tends to lead to an underestimate of the effective temperature in both modules. However, while GSP-Phot and ESP-HS A_0 interstellar extinction estimates are not identical, they remain consistent. Differences become more obvious at $T_{\rm eff}$ >10 000 K (Fig. 14 in Creevey et al. 2023), where ESP-HS results tend to be slightly larger because of the different treatment of the BP/RP (i.e., the module is ignoring the near-IR and applying an empirical correction to the simulated spectra). When the results obtained in open clusters are considered, the two algorithms consistently provide larger monochromatic extinctions (Fig. 30 in Fouesneau et al. 2023) than the values found in Cantat-Gaudin et al. (2020).

[1]PARSEC isochrones (v1.2S and CMD 3.6 web interface) are available from http://stev.oapd.inaf.it/cgi-bin/cmd

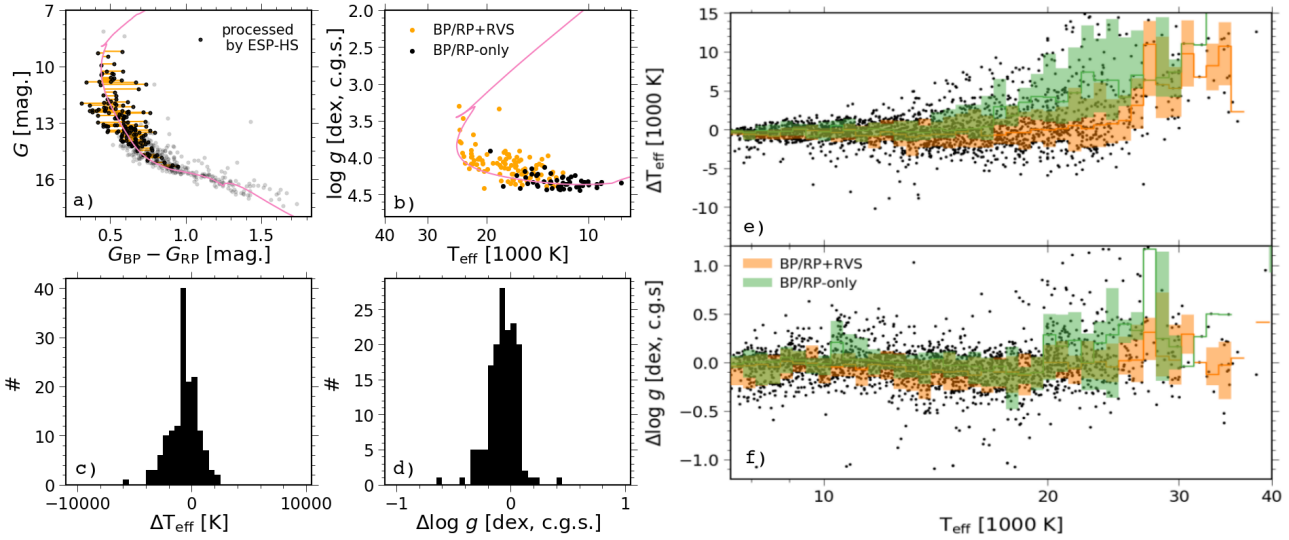


Fig. 5. The colour diagram (Panel a) of the cluster NGC869 is used to identify the closest point in the corresponding reddened PARSEC isochrone (pink curve; Nb of targets with APs: 152, $\log(t) = 7.18$, $A_V = 1.68$). The grey points are targets without any published ESP-HS AP. The corresponding Kiel diagram is shown in Panel b by adopting the APs derived by ESP-HS and by making the distinction between the two processing modes (i.e., BP/RP+RVS or BP/RP-only fit). The isochrone T_{eff} and $\log g$ are then used to compute the residuals as shown in Panels c) and d). The same approach was used for 1524 members of 42 open clusters to analyse the offset and spread of ESP-HS APs (Panels e and f). The running median is represented by the solid line step function, while the interquantile range is represented by shades. Different colours are used to make the distinction between the two processing modes.

To generate the BP/RP simulations that the algorithms confront to the observations, the Apsis team chose the interstellar extinction function derived by Fitzpatrick (1999) and fixed the ratio of total to selective extinction at $R = 3.1$. Especially for hot stars, which are often found towards highly reddened lines of sight, these assumptions may have a significant impact (e.g., Maíz Apellániz 2024). They lead to spectra vs. template mismatches, can make the T_{eff} vs. A_0 degeneracy complexier, and bias the A_0 value itself. For ESP-HS, we show in Fig. 6 the order of magnitude of the effects on A_0 and compare its results to those obtained by Maíz Apellániz & Barbá (2018) who determine both Av and R (i.e., R_{5495} in their paper) simultaneously. The relative difference between the two estimates is usually negative, A_0 being smaller, and its absolute value increases with R.

Two estimates of the rotational broadening are available in the Gaia DR3 catalogue: `vbroad` in table `gaia_source` and `vsini_esphs` in `astrophysical_parameters`. The former was determined by the CU6 pipeline on the per transit (i.e., epoch spectrum) RVS spectrum for all stars cooler than 15 000 K; the latter was measured simultaneously with the APs on the co-added RVS of stars hotter than 7500 K. The behaviour of `vbroad` is described in Frémat et al. (2023). A comparison between the two measurements is shown in Fig. 11.82 of the online Gaia DR3 documentation (Sect. 11.4.4). At lower $V \sin i$, ESP-HS estimates tend to be larger than `vbroad`. It is due to the accuracy of the radial velocities being lower in hot stars (Blomme et al. 2023) and to the assumptions made on the LSF. The median `vsini_esphs` of the stars with $T_{eff} > 7500$ K is found to be relatively constant with G_{RVS}, and ranges from 100 to 140 km s^{-1} as also found for main sequence stars in the $V \sin i$ compilation of Glebocki et al. (2001). At $G_{RVS} > 11.5$, the median is dropping rapidly, which puts a magnitude threshold above which the $V \sin i$ derived from co-added RVS spectra should be disregarded as less accurate. Fig. 7 can directly be compared to Fig. 14 of Frémat et al. (2023), which is its equivalent obtained for `vbroad`. As expected, the upper magnitude threshold of `vbroad` is, for the hot stars, lower and therefore less robust with G_{RVS} than the value obtained from co-added spectra. At temperatures greater than 10 000 K, the main criteria remaining to derive the projected rotational velocity are the broad Paschen lines. These lines are very sensitive to temperature and surface gravity variations, and therefore any offset on the APs directly affects the accuracy of `vsini_esphs`. For this reason, the results obtained above 25 000 K must be considered with caution, especially at $G_{RVS} > 10$.

In Fig. 8, we explore the impact of having metallicity mismatches on the APs computed by ESP-HS assuming

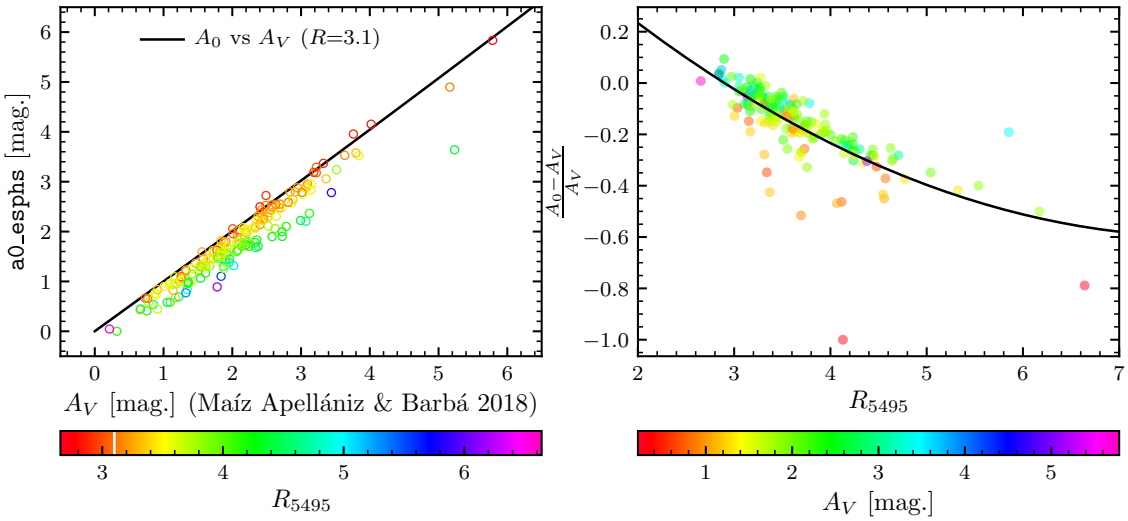


Fig. 6. Left panel: Comparison of the A_0 monochromatic extinction value obtained by ESP-HS for a sample of galactic O-type stars to the A_V found in Maíz Apellániz & Barbá (2018). The black line represents the relation that links A_0 to A_V, when R = 3.1. The colour coding follows the ratio of total to selective extinction, R_{5495}, found by the same authors (see colour bar, where the horizontal white line represents the R=3.1 value assumed in Apsis modules). **Right panel:** We plot the relative extinction difference as a function of R_{5495} (black dots). The blue curve shows a second-degree polynomial through the data. The colour scale follows the A_V (Maíz Apellániz & Barbá 2018).

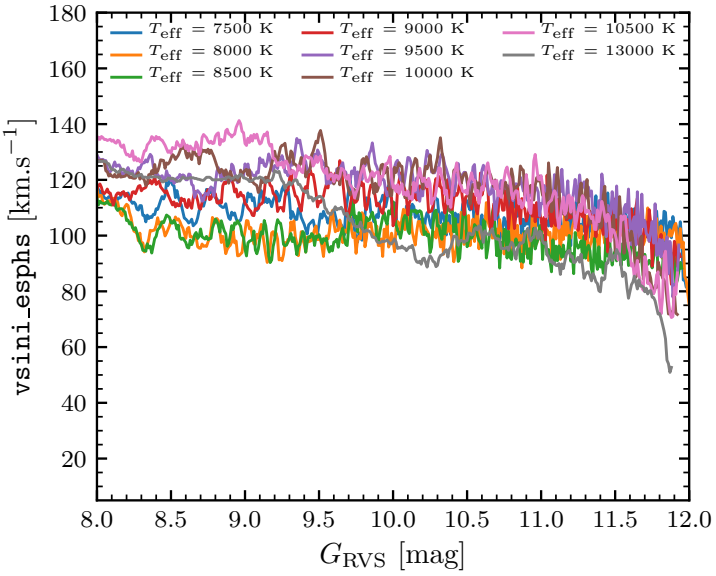


Fig. 7. Running median of `vsini_esphs` as a function of the $G_{\rm RVS}$ magnitude. Various temperature bins (`teff_esphs` ± `250 K`) are considered and represented by different colours.

a solar chemical composition. For the sample studied by Xiang et al. (2022), we compare the AP residuals to the metallicity mismatches (upper panels of Fig. 8). The plots are limited to stars with $T_{\rm eff}$ ≤10 000 K, so that we can confront these density maps to those obtained for the same targets by GSP-Phot (lower panels of Fig. 8). The surface gravity and the V sin i residuals are the most correlated with [Fe/H] mismatches.

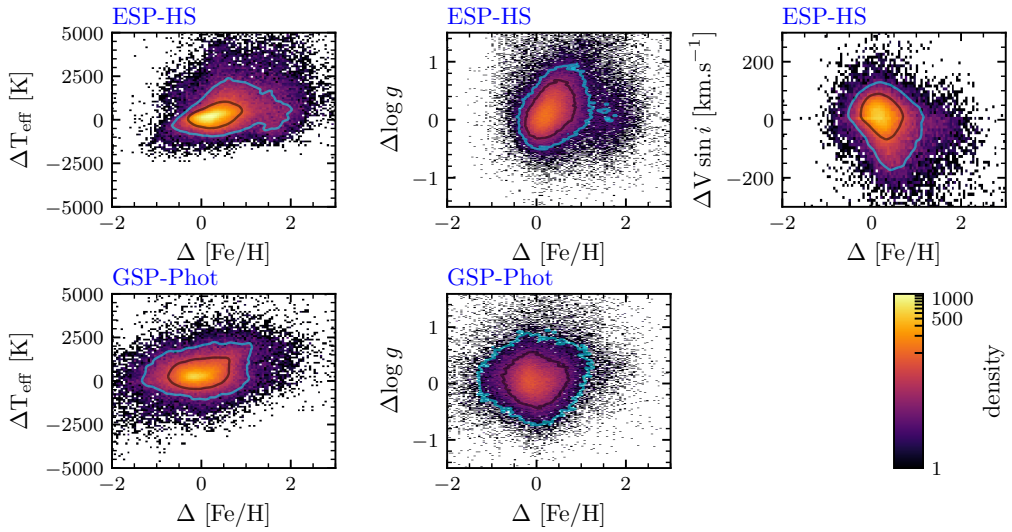


Fig. 8. Effective temperature, surface gravity, and V sin i residuals (i.e., Apsis value minus the one derived from LAMOST spectra by Xiang et al. 2022) are plotted as a function of the [Fe/H] residuals. As GSP-Phot is most reliable below 10 000 K, we limited the comparison to A-type stars. We defined the black and cyan contour lines such that they encompass 68% and 90% of the targets, respectively. Upper panels: correlation of the residuals obtained by ESP-HS. Lower panels: correlation of the residuals obtained by GSP-Phot.

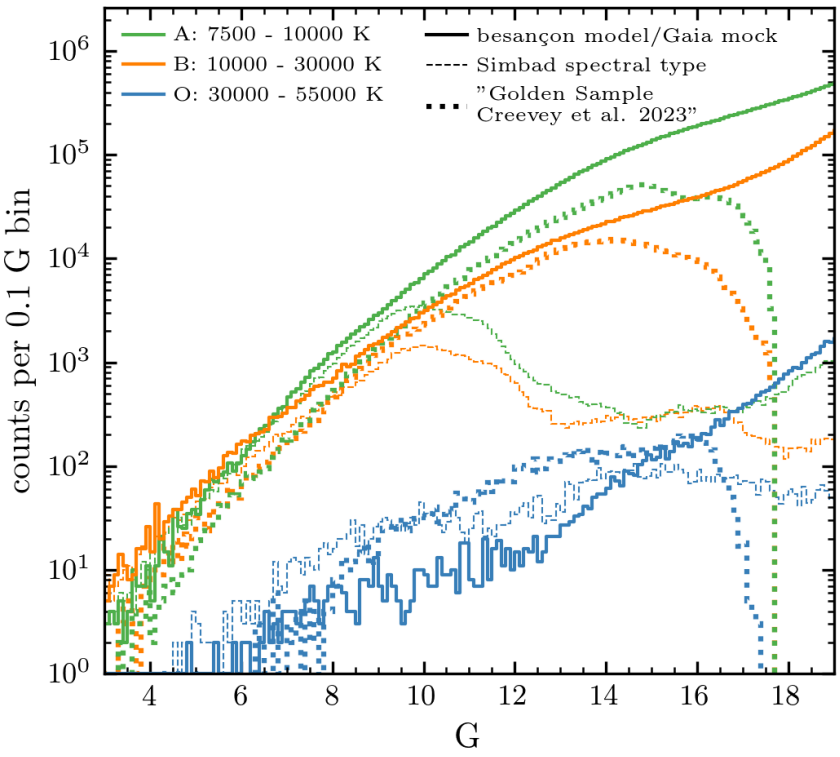


Fig. 9. Magnitude distribution of A-, B-, and O-type stars. The thick coloured lines represent the distribution as found in the GeDR3mock catalogue (Rybizki et al. 2020; Robin et al. 2012). The spectral type definition is given in the legend. The thin broken lines show the counting of hot stars done in 2022 using the basic information tab of the Simbad database. The magnitude distributions of the OBA 'golden sample' (Gaia Collaboration, Creevey et al. 2023) is plotted with thick dotted lines.

4 Discussion and conclusions

In Fig. 9, we present the magnitude distribution of A-, B-, and O-type stars as it is expected by the Besançon galactic model and the GeDR3mock catalogue (Rybizki et al. 2020; Robin et al. 2012). We also compare these distributions to those found in the Simbad database in 2022. These statistics do not take into account the most recent classification work done in various ground-based surveys, but one may still expect to identify with Gaia a significant number of hot stars at magnitudes $G > 12$. The difficulty is knowing how representative of the galactic hot star population such a sample is and how pure and complete it can be.

A significant effort was made all along in the validation of the results to remove from the catalogue spurious AP estimates. The criteria used during the post-processing were often based on the offline analysis of the goodness-of-fit distributions. However, the candidate list of O-, B-, and A-type stars is still polluted by lower-mass stars. In the list, we still find white dwarfs (WDs), subdwarf stars, and hot horizontal branch stars (hot HBs) that share the same effective temperature as more massive hot stars. The presence of hot HBs, for example, is seen through overdensities in the Kiel diagram of stars with magnitude $G > 14$ (right panel of Fig. 19 in Creevey et al. 2023), while WDs are better seen in the absolute magnitude vs. $T_{\rm eff}$ diagram. More unexpected is the presence of cooler stars, such as RR-Lyrae or LPVs, in the sample of hot stars. These misclassifications can be attributed to the combination of AP degeneracy (i.e., effective temperature and interstellar extinction), template mismatches (our simulations do not reproduce all the existing classes of stars and chemical compositions), and stellar variability (e.g., the number of combined epochs is often different in the BP and RP passbands).

To improve the cleaning of the hot star sample, other criteria were applied to the available APs after the post-processing. These are described in Gaia Collaboration, Creevey et al. (2023) and resulted in the publication with Gaia DR3 of a list of 3 023 388 candidate OBA stars (`gold_sample_oba_stars`). Their magnitude distribution is plotted in Fig. 9. For the A- and B-type stars, it fits well with the expectations of the GeDR3mock catalogue up to magnitude 14. The drawback of certain filtering criteria is the removal of bona fide OBA stars on the bright and faint ends of the magnitude domain. The impact of the filters and of the remaining $T_{\rm eff}$ vs. interstellar extinction degeneracy on the completeness of the sample is shown in Fig. 2 of Gaia Collaboration, Creevey et al. (2023), where the fraction of identified to expected OBA stars in open clusters is reported as a function of A_0. The higher A_0, the more hot stars are missing.

We thank the BELgian federal Science Policy Office (BELSPO) for funding us through the PROgramme de Développement d'Expériences scientifiques (PRODEX) and the grant 4000140452. This research has made use of the SIMBAD database, operated at CDS (Centre de Données astronomiques de Strasbourg), Strasbourg, France. This work has made use of data from the European Space Agency (ESA) mission *Gaia* (`https://www.cosmos.esa.int/gaia`), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, `https://www.cosmos.esa.int/web/gaia/dpac/consortium`). Funding for the DPAC has been provided by national institutions, in particular, the institutions participating in the *Gaia* Multilateral Agreement.

References

Andrae, R., Fouesneau, M., Sordo, R., et al. 2023, A&A, 674, A27
Blomme, R., Daflon, S., Gebran, M., et al. 2022, A&A, 661, A120
Blomme, R., Frémat, Y., Sartoretti, P., et al. 2023, A&A, 674, A7
Bressan, A., Marigo, P., Girardi, L., et al. 2012, MNRAS, 427, 127
Cantat-Gaudin, T., Anders, F., Castro-Ginard, A., et al. 2020, A&A, 640, A1
Chen, Y., Bressan, A., Girardi, L., et al. 2015, MNRAS, 452, 1068
Creevey, O. L., Sordo, R., Pailler, F., et al. 2023, A&A, 674, A26
Evans, C. J., van Loon, J. T., Hainich, R., & Bailey, M. 2015, A&A, 584, A5
Fitzpatrick, E. L. 1999, PASP, 111, 63
Fouesneau, M., Frémat, Y., Andrae, R., et al. 2023, A&A, 674, A28
Frémat, Y., Royer, F., Marchal, O., et al. 2023, A&A, 674, A8
Gaia Collaboration, Creevey, O. L., Sarro, L. M., Lobel, A., et al. 2023, A&A, 674, A39
Glebocki, R., Gnacinski, P., & Stawikowski, A. 2001, VizieR Online Data Catalog, III/226
Hubeny, I. 1988, Computer Physics Communications, 52, 103
Katz, D., Sartoretti, P., Guerrier, A., et al. 2023, A&A, 674, A5
Lanz, T. & Hubeny, I. 2003, ApJS, 146, 417
Lanz, T. & Hubeny, I. 2007, ApJS, 169, 83

Luo, A. L., Zhao, Y. H., Zhao, G., & et al. 2019, VizieR Online Data Catalog, V/164
Maíz Apellániz, J. 2024, arXiv e-prints, arXiv:2401.01116
Maíz Apellániz, J. & Barbá, R. H. 2018, A&A, 613, A9
Napiwotzki, R., Schoenberner, D., & Wenske, V. 1993, A&A, 268, 653
Robin, A. C., Luri, X., Reylé, C., et al. 2012, A&A, 543, A100
Rybizki, J., Demleitner, M., Bailer-Jones, C., et al. 2020, PASP, 132, 074501
Shulyak, D., Tsymbal, V., Ryabchikova, T., Stütz, C., & Weiss, W. W. 2004, A&A, 428, 993
Soubiran, C., Le Campion, J.-F., Brouillet, N., & Chemin, L. 2016, A&A, 591, A118
Weidner, C. & Vink, J. S. 2010, A&A, 524, A98
Xiang, M., Rix, H.-W., Ting, Y.-S., et al. 2022, A&A, 662, A66

École Evry Schatzman 2023: Stellar physics with Gaia
C. Babusiaux, C. Reylé (eds)

SOME ASPECTS OF GALACTIC DYNAMICS IN THE GAIA ERA

B. FAMAEY

Université de Strasbourg, CNRS UMR 7550, Observatoire astronomique de Strasbourg, Strasbourg, France
e-mail: benoit.famaey@astro.unistra.fr

Abstract. In the context of this school on stellar physics with Gaia, I briefly present some basic tools of theoretical Galactic dynamics, both in terms of the construction of equilibrium models and of linear perturbation theory.

Keywords: Milky Way dynamics; Gaia

1 Introduction

The Gaia mission (Gaia Collaboration et al. 2016) has been (and still is) an extraordinary milestone in offering full six-dimensional phase-space information for a large number of stars of the Milky Way over a larger volume than ever before (Gaia Collaboration et al. 2023). This offers a unique possibility to improve our dynamical models of the Galaxy, which can serve as a laboratory for a plethora of unanswered questions in the field of galaxy formation and evolution, for instance in terms of understanding the role of environment, the accretion history, the nature of internal instabilities such as bars and spiral arms, or even test models for the nature (or existence) of dark matter.

The most common typically "top-down" dynamical approach consists in generating *ab initio* simulations of galaxies resembling the Milky Way in a cosmological context (e.g., Renaud et al. 2021). While extremely valuable for understanding general features of galaxy formation, this method nevertheless lacks the flexibility needed to create a model that precisely aligns with the vast and detailed data available for our own Galaxy. The complementary "bottom-up" approach for dynamical modeling consists in starting from existing Galactic data, and constructing a model from there rather than relying on simulations. In such a model, a single-particle phase-space distribution function is used to represent all the different constituent particles, namely various stellar populations and dark matter. Such model-building typically starts with the assumptions of dynamical equilibrium and axisymmetry. These assumptions allow us to make use of Jeans' theorem constraining the distribution function to depend only on three integrals of motion, which can typically be chosen to be the radial, azimuthal, and vertical action variables of the canonical action-angle phase-space coordinates. Such dynamical models are however largely insufficient to describe our Galaxy, which is evident from its non-axisymmetric nature featuring, e.g., a prominent central bar and spiral arms that have long been known to leave their imprint in the phase-space structure of the Milky Way (e.g., Famaey et al. 2005). Additionally, recent insights from Gaia have revealed a vertical disequilibrium of the Galactic disk (Antoja et al. 2018), possibly linked to a complex interplay between external disturbances and internal non-axisymmetries (Li et al. 2023). This however does not invalidate the usefulness of the approach, since all these effects can in principle be treated, to a certain extent, through perturbation theory. In this short lecture, I lay out the general principles underlying this approach and present a few recent results along the way.

2 Fundamental equations of Galactic Dynamics

The evolution of any N-particle system (with N being for instance the number of stars, or even of stars and dark matter particles, in a galaxy) is characterized by the conservation of the probability distribution function (DF) in a phase space of $6N$ dimensions, F_N, according to Liouville's theorem. This phase-space DF in $6N$ dimensions is

strictly conserved along trajectories following the Hamiltonian flow of the system. The Bogoliubov (or BBGKY) hierarchy then relates the n-particle (with $n < N$) DF to the $n+1$ DF. The single-particle phase space density $F_1(\mathbf{x}, \mathbf{v}, t)$ (or just $F(\mathbf{x}, \mathbf{v}, t)$ hereafter) is then related to the two-particle DF only through a correlation integral term, which goes to zero in the large N and large relaxation time limit typical for gravity in galaxies (Binney & Tremaine 2008). The single-particle DF then obeys the Vlasov (or collisionless Boltzmann) equation, expressing conservation of particles in infinitesimal patches along the phase space trajectory $\{\mathbf{x}(t), \mathbf{v}(t)\}$. Coupled with the equation for the gravitational potential Φ in the weak-field limit, namely Poisson's equation, this leads to the fundamental system of equations for Galactic dynamics (or at least, and more broadly, for collisionless stellar dynamics), the Vlasov-Poisson system of equations:

$$\begin{cases} \frac{\mathrm{d}F}{\mathrm{d}t} = \frac{\partial F}{\partial t} + \mathbf{v} \cdot \frac{\partial F}{\partial \mathbf{x}} - \frac{\partial \Phi}{\partial \mathbf{x}} \cdot \frac{\partial F}{\partial \mathbf{v}} = 0, \\ \nabla^2 \Phi = 4\pi G \rho = 4\pi G \int \mathrm{d}^3\mathbf{v}\, F. \end{cases} \tag{1}$$

The single-particle Hamiltonian is simply written as $H = \frac{v^2}{2} + \Phi(\mathbf{x})$. Things become a bit more complicated when including a gas component, but this is beyond the scope of the present lecture, and gas will typically be treated as an external component to the Galactic gravitational potential Φ.

3 Equilibrium models

In a realistic axisymmetric and time-independent Galactic potential, the majority of orbits are regular or quasi-regular, exhibiting quasiperiodic patterns in the sense that their Fourier transforms have only discrete frequencies that are integer linear combinations of three fundamental frequencies. These orbits therefore possess three isolating integrals of motion, which uniquely define them within the Galaxy's gravitational potential. The Liouville-Arnold theorem then states that *if, in a Hamiltonian dynamical system with N degrees of freedom, there are also N independent, Poisson commuting integrals of motion, then there exists a canonical transformation to coordinates in which the transformed Hamiltonian is dependent only upon the new generalized momenta, and their canonically conjugated variables evolve linearly in time.* These are called the action-angle variables $(\mathbf{J}, \boldsymbol{\Theta})$, where the actions $\mathbf{J}$ (new generalized momenta) are integrals of the motion and the angles $\boldsymbol{\Theta}$ evolve linearly with time. Said otherwise, Hamilton's equations simply become:

$$\begin{cases} \frac{\mathrm{d}\mathbf{J}}{\mathrm{d}t} = -\frac{\partial H}{\partial \boldsymbol{\Theta}} = 0, \\ \frac{\mathrm{d}\boldsymbol{\Theta}}{\mathrm{d}t} = \frac{\partial H}{\partial \mathbf{J}} = \boldsymbol{\Omega}(\mathbf{J}). \end{cases} \tag{2}$$

According to Jeans' theorem, the equilibrium stellar phase-space DF, $F(\mathbf{x}, \mathbf{v})$, for any Galaxy component should then depend solely on the actions, $F = F(\mathbf{J})$. While analytical relations between action-angles $(\mathbf{J}, \boldsymbol{\Theta})$ and usual phase-space coordinates $(\mathbf{x}, \mathbf{v})$ are seldom attainable for most potentials Φ, these variables present numerous advantages: in an equilibrium configuration, the stars' angles, $\boldsymbol{\Theta}$, are uniformly distributed on phase-space orbital tori defined solely by $\mathbf{J}$, and the phase-space density of stars, $F(\mathbf{J})\mathrm{d}^3\mathbf{J}$, directly relates to the number of stars dN in an infinitesimal action range, divided by $(2\pi)^3$; additionally, the actions remain adiabatically invariant during a gradual change in the Galactic potential; finally, as we shall see in the next section, these variables serve as natural coordinates for perturbation theory. Notwithstanding the absence of an easy analytical transformation for most potentials, in order to transform from actions and angles to positions and velocities, one typically first expresses the Hamiltonian in the action-angle coordinates $(\mathbf{J}_T, \boldsymbol{\Theta}_T)$ of a *toy* potential, for which the transformation to positions and velocities is fully known analytically (generally with an isochrone potential). One then searches for a type 2 generating function $G(\boldsymbol{\Theta}_T, \mathbf{J})$ expressed as a Fourier series expansion on the toy angles $\boldsymbol{\Theta}_T$, for which the Fourier coefficients are such that the Hamiltonian remains constant at constant $\mathbf{J}$. This generating function fully defines the canonical transformation from actions and angles to positions and velocities. This method is known as the 'Torus mapping' (McGill & Binney 1990). For the inverse transformation, an estimate based on separable potentials can be used. These potentials are known as Stäckel potentials (e.g. Famaey & Dejonghe 2003), for which each generalized momentum depends on its conjugated position through three isolating integrals of the motion. These potentials are expressed in spheroidal coordinates defined by a focal distance. For a Stäckel potential, this focal distance of the coordinate system is related to the first and second derivatives of the potential. One can thus use the true potential at any configuration space point to compute the equivalent focal distance *as if* the potential were of Stäckel form, and compute the corresponding isolating integrals of the motion and the actions. This method is known as the Stäckel fudge.

Both transformations are part of the Action-based Galaxy Modelling Architecture (`AGAMA`; Vasiliev 2019) code. Other methods relying on unsupervised machine learning methods have also been developed (Ibata et al. 2021).

To construct an equilibrium model, one can then devise parametric DFs representing the different components of the Galaxy. Typically, one can (i) start from a guess for the Galactic potential (already taking into account some constraints such as the rotation curve, and including a gas surface density) and for the action-space DF expressed as a linear combination of different components (young thin disk, intermediate-age thin disk, old thin disk, thick disk, stellar halo, dark matter halo, bulge/fat disk in the central regions,...), (ii) compute the configuration space density associated with this guess-DF when integrating it over velocity space, (iii) compute a new Galaxy potential from this density with Poisson equation, (iv) take a weighted mean of this newly computed potential and the original guess-potential, and iterate until convergence. This procedure ensures one to construct a self-consistent equilibrium model. In principle, the likelihood of such a model can then be obtained on an individual star-by-star basis in observables space (sky positions, parallaxes, proper motions from Gaia, radial velocities from the RVS or ground-based surveys if available) taking into account the selection function, and the whole parameter space can be explored fully. Interestingly, taking away the baryonic component also allows to compute the shape of the dark matter halo without the contraction related to the presence of baryons, which can give clues to the nature of dark matter. Such a titanesque procedure has however not been attempted yet on the full Gaia dataset, but the closest to this endeavour has been achieved by Binney & Vasiliev (2023), who presented a qualitative fit, and Binney & Vasiliev (2024) who fit a small subset of stars with APOGEE spectroscopy, including an additional dependence on [Fe/H] and [Mg/Fe] in the DF. These models reveal in detail the zeroth order orbital structure of our Galaxy in action space.

4 Perturbation theory

Such equilibrium axisymmetric models are however insufficient to describe the current phase-space structure of the Galaxy, which has clear imprint of non-axisymmetries. Understanding these is particularly important also in view of understanding the secular evolution of the Galaxy, whose main internal driving mechanisms are instabilities leading to non-axisymmetric modes (bar, spiral arms). To take these into account, let us now consider the influence of a small perturbation to the Hamiltonian,

$$\Delta H(\mathbf{x},t) = \psi(\mathbf{x},t) \ll \Phi. \tag{3}$$

In action-angle coordinates, such a perturbing potential is 2π-periodic in the angles and can therefore typically be written as a Fourier series over the angles

$$\Delta H(\mathbf{J},\mathbf{\Theta},t) = \sum_{\mathbf{n}} \Delta H_{\mathbf{n}}(\mathbf{J},t)\exp(\mathrm{i}\mathbf{n}.\mathbf{\Theta}) \tag{4}$$

The total perturbed DF can then be expressed as $F(\mathbf{J}) + f(\mathbf{J},\mathbf{\Theta})$ with the response DF, $f \ll F$, obeying the *linearized* collisionless Boltzmann equation, obtained by dropping all the high-order terms in the original Vlasov equation:

$$\frac{\partial f}{\partial t} + \mathbf{\Omega}\cdot\frac{\partial f}{\partial \mathbf{\Theta}} - \frac{\partial F}{\partial \mathbf{J}}\cdot\frac{\partial \Delta H}{\partial \mathbf{\Theta}} = 0. \tag{5}$$

Expressing the response DF as a Fourier series over the angles,

$$f(\mathbf{J},\mathbf{\Theta},t) = \sum_{\mathbf{n}} f_{\mathbf{n}}(\mathbf{J},t)\exp(\mathrm{i}\mathbf{n}.\mathbf{\Theta}), \tag{6}$$

and assuming that the response is zero at $t = 0$, each Fourier coefficient is then:

$$f_{\mathbf{n}}(\mathbf{J},t) = \mathrm{i}\,\mathbf{n}\cdot\frac{\partial F}{\partial \mathbf{J}}\int_0^t \mathrm{d}\tau\,\Delta H_{\mathbf{n}}(\mathbf{J},\tau)\exp[-\mathrm{i}\,\mathbf{n}\cdot\mathbf{\Omega}\,(t-\tau)]. \tag{7}$$

Let us consider that the Fourier coefficients of the perturbing potential depend on time only through a logistic function controlling the amplitude of the perturbation and through a periodic sinusoidal function of frequency ω_p, which can account for a perturbing potential of m-fold symmetry rotating with a fixed pattern speed $\Omega_p = -\omega_p/m$. Then, at $t = \infty$, we have (e.g., Monari et al. 2016; Al Kazwini et al. 2022):

$$f_{\mathbf{n}} = \Delta H_{\mathbf{n}}\,\frac{\mathbf{n}.\frac{\partial F}{\partial \mathbf{J}}}{\mathbf{n}.\mathbf{\Omega} + \omega_p}. \tag{8}$$

Knowing the perturbing potential, one can thus compute the response DF in this way. For instance, in Monari et al. (2016), using the epicyclic approximation to get an analytic relation between the action-angle variables and the positions and velocities, and considering a 3D spiral arm perturber with corotation in the outer Galaxy, we showed that the first order moments of the perturbed DF describe "breathing" modes of the Galactic disc in perfect accordance with simulations. But obviously the denominator in Equation 8 diverges at resonances (problem of small divisors). This is due to the fact that, close to resonances, the orbital tori are radically modified and the linear perturbation theory breaks down. It is however possible to circumvent this problem when only one main perturber is at play: for each resonance, one can define a new set of actions and angles to describe the orbits, through two consecutive canonical transformations in order to find the relevant action variables to use in the resonant region (Monari et al. 2017). One can then populate the new tori by phase-averaging the original unperturbed DF $F(\mathbf{J})$ over these new resonant tori, which reproduces reasonably well what happens in simulations. However, this method breaks down in the presence of multiple perturbing patterns because of chaos related to resonance overlaps (Minchev & Famaey 2010), which have otherwise interesting consequences in terms of secular evolution of the disk through their effect on heating and radial migration. Radial migration can in turn be probed through the chemo-orbital distribution of equilibrium models (see previous section), which in principle can help constrain the non-axisymmetric patterns driving it: for instance, the presence of super-metal rich stars in the Solar neighborhood (Kordopatis et al. 2015) is a powerful chemo-orbital probe of the effect of past spiral perturbations and their coupling with the bar. To circumvent the problem of resonance overlaps in linear perturbation theory, one can resort to the backwards integration method originally developed by Vauterin & Dejonghe (1997), based on the conservation of the DF in infinitesimal phase-space patches following the Hamiltonian flow, encoded in the Vlasov equation. One can then compute the current DF $F(\mathbf{x}, \mathbf{v})$ by integrating orbits backward in time to an axisymmetric equilibrium state $F(\mathbf{J})$. Using this method, most kinematic groups observed in the vicinity of the Sun can be reproduced with a multi-modal bar model (Monari et al. 2019), but in regions beyond the Sun, the average radial velocity obtained from such a bar-only model clearly indicates the necessary inclusion of spiral arms (Khalil et al. 2023). The challenge is, however, to recover the velocity field measured with Gaia DR3 without destroying the already pretty satisfactory local velocity field in the presence of the bar alone. This is work in progress, and should shed light on the present-day dynamical structure (pattern speed etc.) of spiral arms in the Milky Way.

5 Self-consistency

In the previous section, I considered the response to a perturbation treated as an external one, even when dealing with an internal one such as a spiral pattern or the bar. While this can catch some highly interesting features of the response, especially relatively far away from the main density perturbation itself, it is not a self-consistent procedure. For instance, it does not allow to follow self-consistently the growth and/or decay of an instability, or to treat the self-consistent response of the Milky Way to an external perturber, for instance to the Sgr dwarf or to the Large Magellanic Cloud (LMC), which are both influencing the current dynamical state of the Galaxy.

For such an endeavour, the perturbation to the Hamiltonian considered in the previous section can be divided into an external ψ^{e} and a self-consistent ψ^{s} internal part:

$$\Delta H(\mathbf{x}, t) = \psi^{\mathrm{e}}(\mathbf{x}, t) + \psi^{\mathrm{s}}(\mathbf{x}, t). \tag{9}$$

When considering an external perturber that can itself accelerate the reference frame of the Galaxy, which is the case when considering an infalling LMC with a mass of $1.8 \times 10^{11} \mathrm{M}_{\odot}$, one can also add to the perturbing Hamiltonian a potential term accounting for a pseudo-force associated to this accelerated frame (Rozier et al. 2022). One can also consider as an "external" perturber some simple shot-noise perturbation to the disk, which disappears quickly but lets the bar and spiral instabilities develop self-consistently as a consequence.

With a perturbation given by Equation 9, the solution to the linearized collisionless Boltzman equation is still given at any time by Equations 6-7 hereabove, but now at each time one computes the response potential via the integral of f over velocity space through Poisson's equation, so that we can update the perturbing Hamiltonian self-consistently. But now we have a problem: the linearized collisionless Boltzmann equation is best expressed in action-angle phase-space coordinates, but we also need to solve the Poisson equation to get the potential of the response, and the latter is best solved in configuration space coordinates, not actions and angles. One can bypass this difficulty by resorting to Kalnajs trick (Kalnajs 1977): projecting everything on a bi-orthogonal basis of potentials $\psi^{(\mathbf{p})}$ and densities $\rho^{(\mathbf{p})}$ that solve the Poisson equation once and for all. The choice of the

bi-orthogonal basis obviously depends of the problem at hand. For instance, when computing the self-consistent response of the dark and stellar halos to the LMC infall onto the Milky Way, one can simply choose spherical harmonics multiplied by a radial component of Clutton-Brock form (Clutton-Brock 1972; Rozier et al. 2022). Another basis must be used when treating the response of the Galactic disk.

The key aspect of a bi-orthogonal potential-density basis is that it obeys the bi-orthogonality condition:

$$\int \mathrm{d}\mathbf{x}\, \psi^{(\mathbf{p})}(\mathbf{x})\, \rho^{(\mathbf{q})*}(\mathbf{x}) = -\delta^{\mathbf{q}}_{\mathbf{p}}, \tag{10}$$

where $*$ indicates the complex conjugate while $(\mathbf{p})$ and $(\mathbf{q})$ stand for multiplets of indices, for instance in the Clutton-Brock case triplets of indices (l, m, n), where l and m are just spherical harmonic degree and order, and n is the radial order. Note that this inner product could be interpreted as proportional to the interaction potential energy between two disturbances, which is thus zero between two distinct basis elements.

Now, in configuration space, the perturber and the self-consistent response can just be expressed as vectors in this basis of potentials and densities, which we call vectors $\mathbf{b}(t)$ and $\mathbf{a}(t)$ respectively. Thanks to the bi-orthogonality condition, in order to obtain the vector decomposition of the self-gravitating response, one just needs to take the inner product of the perturbed density and each of the basis elements:

$$a_{\mathbf{p}}(t) = -\int \mathrm{d}\mathbf{x} \int \mathrm{d}\mathbf{v}\, f(\mathbf{x}, \mathbf{v}, t)\, \psi^{(\mathbf{p})*}(\mathbf{x}), \tag{11}$$

This can be canonically transformed into an integral over actions and angles, in which one simply needs to insert the solution of the linearized collisionless Boltzmann equation for f, given by Equations 6-7. The solution for $a_{\mathbf{p}}(t)$ then simply becomes an integral over time, in which everything that depends on the background state $F(\mathbf{J})$ is absorbed into a "*response matrix*" that multiplies the vector $\mathbf{b} + \mathbf{a}$ (see Rozier et al. 2022, for details). As an example, one can for instance reconstruct the LMC with 3216 basis elements, and study the response of an isotropic halo to its infall, in 20 timesteps over 2 Gyr (Rozier et al. 2022). Contrary to the case of N-body simulations, here the system is still following orbits (and is still responding to the perturber) between timesteps. An interesting aspect of the response matrix specific to non-rotating spheres is that there is no coupling between different angular harmonics: each harmonic in the response is only induced by the corresponding harmonic in the perturber, mediated by that same harmonic in the matrix. This allows one to make a harmonic decomposition of the response for different halo anisotropies. In Rozier et al. (2022), we found that the large scale over/underdensity induced by the LMC in the stellar halo corresponds to the $l = 1$ terms, and is purely associated to the reflex motion: the underlying kinematics of the halo do not change this. What is sensitive to the underlying kinematics of the stellar halo population one considers is the local wake around the LMC. The response is much stronger in the case of a radially-biased halo. The orientation and winding of the quadrupolar ($m = 2$) response is also very different. Importantly, the response of the stellar halo does not allow to infer the presence of a particle-made dark matter halo. We also looked at the effect of the different Fourier numbers in our Fourier decomposition to see which resonance dominates in different anisotropic cases. For a radially anisotropic stellar halo, the inner Lindblad resonance dominates: in that case, instead of attracting particles which can move with it, the LMC rather attracts orbits which can precess with it, the wake is relatively slow, trails behind the LMC, and remains strong. For a tangentially anisotropic stellar halo, we found that the corotation frequency dominates: the wake is then basically following the LMC and dissipates quickly as the LMC moves inducing a weaker response. These insights on the dynamics at play can only be gotten from such an analytical treatment, which is thus very complementary to simulations. Similar studies should, in the future, be carried out on the response of the Galactic disk to its interaction with the Sgr dwarf, whose signature might be partially encoded in the phase-spiral discovered in Gaia data.

6 Conclusions

Our Galaxy can serve as a laboratory to answer a plethora of yet unanswered questions in the field of galaxy formation and evolution. In this brief presentation, I outlined the basics of theoretical/analytical Galactic dynamics, both in terms of the construction of equilibrium models and of linear perturbation theory, which are very powerful tools to dissect the dynamical mechanisms at play (which are often difficult to disentangle in numerical simulations), and are in this sense highly complementary to numerical simulations of galaxy formation in a cosmological context. I also gave a few examples of recent results using such tools. Constructing equilibrium models reveals the zeroth order axisymmetric orbital structure of the Galaxy, and can be coupled with a chemical

decomposition to reveal the formation history of its different stellar components. This chemical decomposition of phase-space already reveals that the disk of the Galaxy is constantly reshuffling its orbits under internal secular evolution processes, such as stellar radial migration, driven by non-axisymmetric (short-lived and/or long-lived) modes and their respective couplings. Understanding the exact nature of spiral arms and of their role in the secular evolution of the Galaxy has remained a surprisingly pressing challenge for several decades, and a chemo-orbital analysis of stars of the Milky Way is a powerful probe to advance on this question. Gaia data, complemented by ground-based spectroscopic surveys, have also revealed a present-day highly complex picture of our Galaxy, both affected by internal instabilities (spirals, bar) and external perturbations (LMC, Sgr dwarf). Here again, dynamical modeling and linear perturbation theory are powerful complementary tools to numerical simulations in order to disentangle the various effects at play, and reconstruct past events such as the Gaia-Sausage-Enceladus merger. Some of these effects at play (possible slow-down of the rotation of the Galactic bar, dynamical friction on the Sgr dwarf) also involve – in principle – interactions of the stellar components of the Galaxy with the dark matter halo, providing a potentially powerful probe of the nature of dark matter. More generally, dynamical modeling can probe the shape and substructure of the putative dark matter halo, and compare it with expectations from the standard ΛCDM model and in alternative frameworks. Regarding the formation history of the Galaxy, many questions still remain unanswered too, related to, e.g., the origin of the chemical thick disk, or the nature of the disk-halo interface. For these questions, the tools presented here complemented by future better and higher resolution measurements of stellar chemical abundances and of more precise stellar ages (e.g., Miglio et al. 2017) should allow us to be confident that the complex formation history of our Milky Way will be revealed in full in the coming decades.

References

Al Kazwini, H., Agobert, Q., Siebert, A., et al. 2022, A&A, 658, A50
Antoja, T., Helmi, A., Romero-Gómez, M., et al. 2018, Nature, 561, 360
Binney, J. & Tremaine, S. 2008, Galactic Dynamics: Second Edition
Binney, J. & Vasiliev, E. 2023, MNRAS, 520, 1832
Binney, J. & Vasiliev, E. 2024, MNRAS, 527, 1915
Clutton-Brock, M. 1972, Ap&SS, 16, 101
Famaey, B. & Dejonghe, H. 2003, MNRAS, 340, 752
Famaey, B., Jorissen, A., Luri, X., et al. 2005, A&A, 430, 165
Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, A&A, 595, A1
Gaia Collaboration, Vallenari, A., Brown, A. G. A., et al. 2023, A&A, 674, A1
Ibata, R., Diakogiannis, F. I., Famaey, B., & Monari, G. 2021, ApJ, 915, 5
Kalnajs, A. J. 1977, ApJ, 212, 637
Khalil, Y., Famaey, B., Monari, G., & Siebert, A. 2023, in SF2A-2023, 91–94
Kordopatis, G., Binney, J., Gilmore, G., et al. 2015, MNRAS, 447, 3526
Li, C., Siebert, A., Monari, G., Famaey, B., & Rozier, S. 2023, MNRAS, 524, 6331
McGill, C. & Binney, J. 1990, MNRAS, 244, 634
Miglio, A., Chiappini, C., Mosser, B., et al. 2017, Astronomische Nachrichten, 338, 644
Minchev, I. & Famaey, B. 2010, ApJ, 722, 112
Monari, G., Famaey, B., Fouvry, J.-B., & Binney, J. 2017, MNRAS, 471, 4314
Monari, G., Famaey, B., & Siebert, A. 2016, MNRAS, 457, 2569
Monari, G., Famaey, B., Siebert, A., Wegg, C., & Gerhard, O. 2019, A&A, 626, A41
Renaud, F., Agertz, O., Read, J. I., et al. 2021, MNRAS, 503, 5846
Rozier, S., Famaey, B., Siebert, A., et al. 2022, ApJ, 933, 113 https://github.com/simrozier/LiRGHaM
Vasiliev, E. 2019, MNRAS, 482, 1525 http://agama.software
Vauterin, P. & Dejonghe, H. 1997, MNRAS, 286, 812

Acronyms

Acronym	Description
2MASS	Two-Micron All Sky Survey
ADQL	Astronomical Data Query Language
AF	Astrometric Field (in Astro)
AGN	Active Galactic Nucleus
ANN	Artificial Neural Network
AP	Astrophysical Parameter
APOGEE	APO Galactic Evolution Experiment
AU	astronomical unit
BP	Blue Photometer
CCD	Charge-Coupled Device
CDS	Centre de Données astronomiques de Strasbourg
CFHTLS	Canada France Hawaii telescope legacy survey
DENIS	deep near-infrared survey of the Southern Sky
DF	probability distribution function
DPAC	Data Processing and Analysis Consortium
DR1	Gaia Data Release 1
DR2	Gaia Data Release 2
DR3	Gaia Data Release 3
DR4	Gaia Data Release 4
DR5	Gaia Data Release 5
EDR3	Gaia Early Data Release 3
ESA	European Space Agency
ESO	European Southern Observatory
ESP-CS	Extended Stellar Parametrizer for Cool Stars
ESP-ELS	Extended Stellar Parametrizer for Emission Line Stars
ESP-HS	Extended Stellar Parametrizer for Hot Stars
ESP-UCD	Extended Stellar Parametrizer for Ultra Cool Dwarfs
FLAME	Final Luminosity, Age, and Mass Estimator
FPR	Focused Product Release
GCNS	Gaia Catalogue of Nearby Stars
GES	Gaia ESO Survey
GSP-Phot	Generalised Stellar Parametrizer PHOTometry
GSP-Spec	Generalised Stellar Parametrizer SPECtroscopy
GUCDS	Gaia Ultra-Cool Dwarf Sample
HB	Horizontal Branch stars
HEALPix	Hierarchical Equal-Area iso-Latitude Pixelisation
HRD	Hertzsprung-Russell Diagram
IAU	International Astronomical Union
ID	Identifier (Identification)
ISM	InterStellar Medium
LAMOST	Large Sky Area Multi-Object Fibre Spectroscopic Telescope
LMC	Large Magellanic Cloud (special, high-density area on the sky)
LPV	Long-Period Variable (star)
LSF	Line Spread Function
LSST	Legacy Survey of Space and Time of the Rubin Observatory (formerly known as Large-aperture Synoptic Survey Telescope)
LTT	Light Travel Time
MCMC	Monte-Carlo Markov Chain
MSC	Multiple-Star Classifier
NSL	Nominal Scanning Law

NSS	Non-Single Star
OGLE	Optical Gravitational Lensing Experiment
PanSTARRS	Panoramic Survey Telescope And Rapid Response System
PARSEC	PAdova and TRieste Stellar Evolution Code
PVP	Performance Verification Paper
QSO	Quasi-Stellar Object
RAVE	RAdial Velocity Experiment
RP	Red Photometer
RVS	Radial Velocity Spectrometer
SB2	Double-lined Spectroscopic Binary
SDSS	Sloan Digital Sky Survey
SED	Spectral Energy Distribution
SIMP	Sondage Infrarouge de Mouvement Propre
SMC	Small Magellanic Cloud (special, high-density area on the sky)
SNR	Signal-to-Noise Ratio (also denoted SN and S/N)
SVO	Spanish Virtual Observatory
TDI	Time-Delayed Integration (CCD)
TESS	Transiting Exoplanet Survey Satellite
TGAS	Tycho-Gaia Astrometric Solution
TOPCAT	Tool for OPerations on Catalogues And Tables
TRAPPIST	Transiting planets and planetesimals small telescope
UCD	Ultra-Cool Dwarf (star, brown dwarf)
UKIDSS	UKIRT Infrared Deep Sky Survey
WD	White Dwarf (star)
WISE	Wide-field Survey Explorer
XP	Blue or Red Photometer