Science with Gaia: 2024

Essays on the scientific results from Gaia
michaelperryman.co.uk

Michael Perryman

Essays 157–209 (Jan – Dec 2024)

Preface

GAIA IS a mission of the European Space Agency dedicated to astrometry – the measurement of the positions of celestial bodies. The satellite was launched in 2013, and operated until January 2025. Gaia provides the distances and motions (and much other related data) of more than two billion stars in our Galaxy and beyond, with an unprecedented accuracy barely imaginable 25 years ago. It builds on the success of ESA's pioneering Hipparcos mission, which was operated in orbit between 1989–93. The Hipparcos Catalogue of nearly 120 000 stars was published in 1997.

Because of the enormous amount of data processing involved, improved Gaia catalogues are being released progressively, with Data Release 1 in 2016, Data Release 2 in 2018, Early Data Release 3 in 2020, and Data Release 3 in June 2022. Data Releases 4 and 5 are scheduled for the end of 2026, and 2030, respectively.

SINCE THE beginning of 2021, I have been writing a (mostly weekly) 2-page 'essay', picking out some scientific highlights of the mission as they are emerging, or as they caught my attention, and mixing them with occasional asides on related topics, including the history of astrometry, and some more technical, managerial, or developmental aspects of both Hipparcos and Gaia.

Who are they written for? Anyone who might have a general interest in science and astronomy, including amateur astronomers, young scientists starting out on their careers, mid-career scientists looking in on Gaia for the first time to get a feeling of what is possible, and even specialists looking in from different areas of astronomy, or physics more generally.

THE SCIENTIFIC TOPICS I select each week are not necessarily the most important. They do not follow any specific sequence. They are not a complete review of a given topic. Many will be quickly superseded by new results. But together, they are a look at what this long journey of space astrometry is achieving. They offer a snapshot of some of the discoveries that Gaia is making, and written in a form that I hope will be reasonably accessible to those not so deeply involved. I post these weekly essays on my own www site michaelperryman.co.uk.

In each, I have included a footer (DR1, DR2, EDR3, DR3) to indicate which of the (latest) data releases the essay refers to. I have used DR0 to signify technical or historical material not connected with any specific data release. This is intended to communicate how current (or out of date!) any particular essay is likely to be.

Only a few references are included, and these are (generally) 'discreetly' hyperlinked for those who want to read more. Where references appear in the form (Einstein 1908) or www.gaia.com, clicking on the text (even though not generally highlighted) should lead to the relevant (ADS) online article.

HERE, I gather all of my essays from one year in a single compilation, which can also be displayed on-screen in a 'flip book' format.

Michael Perryman

Contents

157. Many more Gaia sources in ω Cen

Gaia Data Release 3 was issued in June 2022, and the next major release, DR4, is not scheduled until 2025. But on 10 October 2023, the Gaia Data Processing and Analysis Consortium (DPAC) published five papers forming part of a special 'Focused Product Release'. I will look at each of these in my next five essays.

The first, which I will describe here, relates to the globular cluster ω Cen. The key point is that a special observing mode, implemented to overcome some of the limitations of the extreme spatial crowding in the cluster's core, has resulted in an additional *half a million new stars* observed by Gaia in its central region!

I looked at some of the first Gaia results on ω Cen in essay #40 (October 2021). As I explained there, at a distance of about 5200 pc, it is the largest and most massive globular cluster in the Milky Way. With a diameter of 50 pc, it contains some 10 million stars, has a total mass of $4 \times 10^6 M_\odot$, and is thought to be the core remnant of a disrupted dwarf galaxy. But it is too distant, too crowded, and just too faint, for *any* of its stars to have been observed by Gaia's predecessor, Hipparcos!

In essay #40, I also gave some of the key results on the numbers of cluster stars included in EDR3, and the resulting estimates of its distance and proper motion. Specifically, Soltis et al. (2021) selected 178 548 stars from EDR3 within 45 arcmin of the cluster centre. They estimated a mean cluster proper motion of $\mu_\alpha = -3.25\,\mu$as, $\mu_\delta = -6.76\,\mu$as, in good agreement with the value found for DR2 (Baumgardt et al., 2019).

Selecting stars participating in this common space motion then resulted in 108 054 candidate members. Further restriction according to their location in the colour–magnitude diagram resulted in 66 467 members with good Gaia astrometry and two-colour photometry.

Their mean cluster parallax was 0.191 ± 0.001 mas, corresponding to a distance of 5236 ± 28 pc, a value in good agreement, for example, with the classical photometric distance of 5.2 kpc from Harris (1996).

I also described the calibration of the luminosity of the tip of the cluster's red giant branch, and the implications for the distance scale, and the value for H_0.

However, the DR3 catalogue for ω Cen, along with a handful of other high-density stellar regions on the sky, shows a striking absence of detected sources in its most crowded central region. To understand how a dedicated observing mode, and the associated 'Focused Product Release', has now allowed the detection of many more stars in the central regions of ω Cen, let me outline the principles and limitations of Gaia's normal observational mode, and explain how these 'holes' arise.

In Gaia's normal mode of operation, the images of sources entering each of the two fields of view are detected by the instrument's 'sky mapper'. Information on detected and confirmed sources (their magnitude, as well as their 2-d coordinates and trajectory in the focal plane) is passed in real time to the subsequent CCDs of the Astrometric Field (AF), the Blue and Red Photometer (BP/RP), and the Radial Velocity Spectrograph (RVS).

In normal mode, accurate astrometry is obtained from the astrometric field. These observations preserve the full CCD pixel resolution *along* scan, so yielding the most accurate astrometry along-scan. But for sources fainter than 13 mag, and in order to limit the volume of data sent to the ground, a number of across-scan pixels are 'collapsed' (on-board) into a single data sample.

Gaia's resulting astrometric crowding limit is around 1.05 million objects deg^{-2}, or about one object per 12 arcsec^{-2}. The completeness of the Gaia catalogue is, accordingly, generally very high. But for a small number of very densely populated regions on the sky, this standard 'windowing' and readout strategy for the astrometric field observations becomes saturated: multiple sources begin to overlap, leading to conflicting readout windows, to windows containing multiple sources, or to no readout at all.

In a program of dedicated 'Service Interface Function' scans, an alternative approach to the on-board observations of a number of such high stellar density regions has been implemented. These special observations of crowded fields have been obtained for the globular clusters ω Cen, 47 Tuc, M4, M22, and NGC 4372, as well as Baade's Window in the Galactic bulge, the LMC and SMC, and the Sagittarius I galaxy.

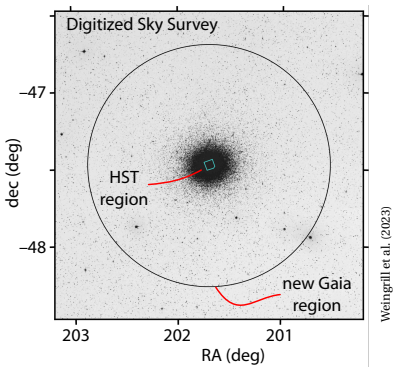
In this special mode of operation, the satellite's 'scanning law' and spin rate are not modified. But sources, and their associated astrometric data, are instead derived from the binned (2 × 2) pixels of the sky mapper. This has the advantage of preserving much more of the *across-scan* spatial information, and thereby significantly reducing source confusion. A further improvement in the effect of crowding results from the fact that the sky mappers are only illuminated by a *single* field-of-view, while the astrometry field observations result from the superposition of the two fields of view (in order to provide relative astrometry over widely separated regions of the sky).

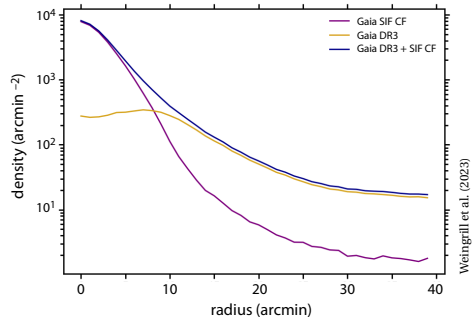
Disadvantages of these dedicated sky mapper observations are twofold. First is the lower astrometric accuracy compared with the astrometric field data, the latter resulting from nine along-scan CCDs, each exploiting the full along-scan pixel resolution. The second is the absence of colour and spectral information for the sources which, for the nominal mode of operation, is obtained from the Blue and Red Photometer (BP/RP), and the Radial Velocity Spectrograph (RVS) fields.

As I mentioned at the start, results for ω Cen were made available (as part of program of five 'Focused Product Releases', supplementing DR3), in October 2023. As described in detail by Weingrill et al. (2023), these new results were derived from special scans obtained between 1 January 2015 and 8 January 2020, and a dedicated processing pipeline. Incidentally, further such crowded fields observations are still ongoing.

As a result, the previous 'hole' in the core of the ω Cen cluster, evident in Data Release 3, is now 'filled' with 526 587 additional sources, compared with the 321 698 already present in Gaia DR3. Their fidelity, and their associated astrometry, was validated by comparing the results to a dedicated Hubble Space Telescope dataset covering the cluster core.

Magnitude uncertainties are in the range 1–10 mmag over $G = 15 - 20$ mag. Astrometric standard errors are significantly higher than for Gaia DR3, with median uncertainties of ~ 3 mas in RA/Dec, ~ 2 mas in both components of proper motion, and ~ 4 mas in parallax.





Pre-dating these latest data on the core region, many studies of ω Cen have been published exploiting the Gaia data. These include:

- calibration of the luminosity of the tip of the red giant branch, and the cluster's 3-d structure and Galactic orbit (Soltis et al., 2021; Li et al., 2023c);
- kinematics and orbital anisotropy of the various stellar sub-populations as a function of metallicity (Jindal et al., 2019; Cordoni et al., 2020; Sanna et al., 2020), including retrograde substructures (Myeong et al., 2018a), and internal rotation as a function of stellar mass (Scalco et al., 2023);
- identification of stellar streams (tidal debris) torn off the cluster as it orbits the Galaxy (Ibata et al., 2019a; Ibata et al., 2019b; Kuzma et al., 2021; Canducci et al., 2022), and confirmation of its disrupted dwarf galaxy origin (Youakim et al., 2023);
- chemical tagging (including s-process elements) of the Fimbulthul tidal stream (Simpson et al., 2020);
- confirmation of the cluster's relationship to the Gaia–Sausage/Enceladus stream (Limberg et al., 2022);
- the distribution of dark and luminous mass, with evidence that the dark-mass component is more centrally concentrated (Evans et al., 2022a);
- studies of its metal-poor stars (Johnson et al., 2020b), the distribution of hot subdwarfs (Luo et al., 2021), and of far ultraviolet sources (Prabhu et al., 2022a; Prabhu et al., 2022b);
- distortion mapping of the HST instruments ACS, UVIS and WFPC2 (Kozhurina-Platais et al., 2018; Casetti-Dinescu et al., 2021), and proper motions from HST (Libralato et al., 2018; Scalco et al., 2021);
- axion and neutrino bounds, from the TRGB (Capozzi & Raffelt, 2020), and constraints on the monopole–dipole potential (Poddar & Pachhar, 2023).

I am writing this at a time when this wealth of new observations has only just been made available, and before its scientific exploitation has begun. Given the range of scientific results that have been obtained on ω Cen so far, we can look forward to many deeper insights based on this new source catalogue in the future.

158. Radial velocity time-series of LPVs

GAIA DATA RELEASE 3 was issued in June 2022, and the next major release, DR4, is not scheduled until 2025. But on 10 October 2023, the Gaia Data Processing and Analysis Consortium (DPAC) published five papers forming part of a special 'Focused Product Release'. The first, which I looked at in essay 157, relates to many new Gaia sources in the globular cluster ω Cen.

The second, considered here, involves the application of Gaia's own radial velocity *time-series* measurements to the class of long-period variables (LPVs). The associated study, by Trabucchi et al. (2023), gives some first insights into how the full radial velocity time-series (in contrast with a single mean value) complements the availability of Gaia's time-series *photometry* in characterising this broad and important class of variable star.

AS USUAL, some context. Gaia's epoch *photometry* is in three broad wavelength bands (G, BP and RP). It yields some 70 high-accuracy photometric data points (focal-plane passages) over 5 years, albeit at (non-uniform) epochs dictated by the deterministic scanning law; with more at intermediate ecliptic latitudes, but fewer at higher and lower latitudes. These data have proven immensely powerful in confirming and/or discovering, and characterising, huge numbers of variable stars of all types, both intrinsic and extrinsic.

Specifically Gaia DR3, based on the first 34 months of data, includes nearly 10 million variables (essay 76). Supervised machine learning has classified the photometric time-series into 24 variability classes, amongst them 15 000 Cepheids, 270 000 RR Lyrae, nearly 500 000 rotational variables, more than 2 million eclipsing binaries, and more than 1.7 million long-period variables.

LPVs are intrinsic variables with periods range from days to 1000 days or more, embracing pulsating red giants and supergiants, including Mira variables, as well as semi-regular variables generally pulsating in an overtone. Also included by virtue of their similar light curves are some binary systems, including the 'ellipsoidal variables' (ELL) in which the variability originates from the body's gravitationally distorted shape (essays 79, 133).

WHILE MULTI-EPOCH radial velocity observations are acquired by Gaia's radial velocity spectrometer (RVS) for all sufficiently bright sources, $G_{\rm RVS} \lesssim 16$ (essay 86), and quasi-simultaneously with the epoch photometry, only *mean* values have generally been made available so far in DR3. An exception is for the Cepheids (Ripepi et al., 2023) and RR Lyrae (Clementini et al., 2023) variables, for which the radial velocity time series were included as part of the Gaia DR3 release.

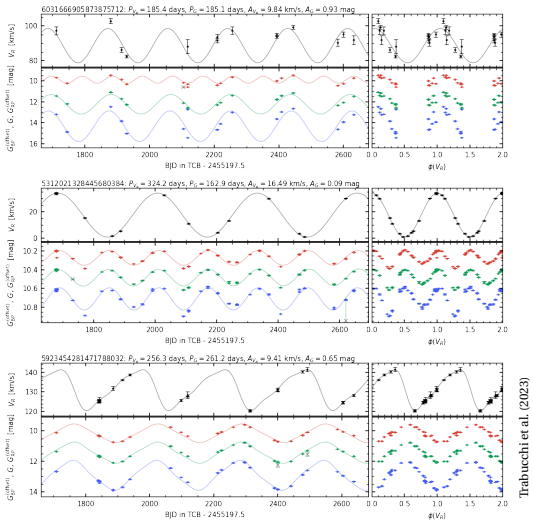
Anticipating the publication of full multi-epoch radial-velocity data with Data Release 4 in 2025, the 'Focused Product Release' described by Trabucchi et al. (2023) provides the radial-velocity time series (and derived variability parameters) for 9614 long-period variables with high-quality observations in the range $G = 6 - 14$ mag, critically selected from the 1.7 million LPV candidates identified in Gaia DR3 (Lebzelter et al., 2023).

Amongst these are their 'top-quality subsample' of 6093 stars whose radial-velocity periods are fully consistent with the values derived from the G, BP, and RP photometric time series. The radial-velocity time series have an average of 24 measurements per source (but ranging between 12–90), non-uniformally over the 34-month span of the DR3 catalogue observations.

Three examples are shown in the following figure: in each, the radial velocity time-series (in black) is acquired simultaneously with the photometric time series in BP (blue), G (green), and RP (red), each suitably offset, and shown phase-folded in the right panels. The three examples show, top to bottom: a source showing a 'mixed consistency' between the radial velocity and photometric time series (see below); a (binary) ellipsoidal variable; and a more classical pulsating long-period variable.

The detailed analysis of Trabucchi et al. (2023) shows that the majority, 88%, are genuine long-period variables. The other 12% are mainly ellipsoidal binaries.

Quality checks against radial velocities available in the literature show excellent agreement, based on 89 matches with sources from APOGEE, 167 from GALAH, and 1226 from RAVE (of which 60, 119, and 784, respectively, are in their 'top-quality sample').

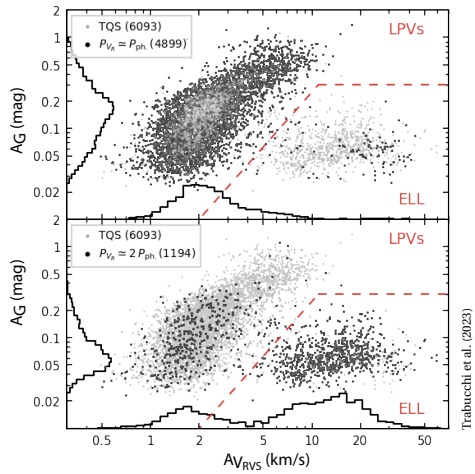


THE SIMULTANEITY of Gaia's radial velocity and photometric time series assists with various challenges associated with the identification, classification, and physical interpretation of long-period variables.

As one example, since the periods of long-period variables extend to several hundred days, obtaining velocity curves at high resolution with a good phase coverage has been an observational challenge, and the total sample of well-observed Mira variables still numbers only a few tens of objects. Over wide ranges of period and metallicity, these suggest a consistent pattern in the velocity variations, with s-shaped velocity curves in the near-infrared, and velocity amplitudes of 20–30 km s^{-1}.

Radial velocities are particularly important for ellipsoidal variables, viz. close binaries typically comprising a red giant and a main sequence star. Despite the absence of eclipses, the gravitationally induced elongation of the red giant leads to rotational brightness variations (essay 133). In turn, this generally results in a different period for the system's photometric variability compared with the orbital period from the radial velocities. The pre-Gaia compilation of radial velocities accompanying the detailed photometric monitoring of objects by the MACHO and OGLE projects counts just 81 objects in the Large Magellanic Cloud (Nie & Wood, 2014).

AS Trabucchi et al. (2023) demonstrate, the distinction between the ellipsoidal and long-period variables can be made more robustly than by simply comparing the ratio between the radial velocity and photometric periods. As shown in the figure below (their Figure 10), the diagram of photometric amplitude, A_G, versus radial velocity amplitude, A_{RVS}, delineates one region populated by smaller A_G and larger A_{RVS} which characterises the ellipsoidal variables (ELL), and a distinct region defined by larger A_G and smaller A_{RVS} which characterises the pulsating long-period variables (LPV).



For the pulsating long-period variables, the photometric variability amplitude increases with pulsation period. But a second grouping is characterised by long periods and comparatively small G-band amplitudes. The current consensus seems to be that these sources are also pulsating long-period variables, but where the dominant period identified by the variability processing pipeline is probably a long *secondary* period (LSP).

Nevertheless, their adopted classification is broadly consistent with the ratios of radial velocity to photometric periods. Among sources in their 'top-quality sample' (TQS), 89% of those identified as ellipsoidal variables show the expected 2:1 ratio between their radial velocity and photometric periods, while around 86% and 96% of the long-period variable candidates showing respectively pulsation or 'long secondary period' variability are more consistent with a 1:1 ratio.

THE PUBLICATION OF the radial-velocity time series for almost 10 000 long-period variables constitutes by far the largest such database now available. And many more will follow with Gaia DR4! As Trabucchi et al. (2023) stress, the high-quality time-series data of this 'Focused Product Release' will also serve a valuable role in providing the community with a means of preparing for the far more extensive time series content of DR4.

In the early days of the technical studies for Gaia, one of my tasks was to underline to ESA's advisory committees and decision-making bodies the importance of contempraneous multi-epoch multi-colour photometry, and the even greater challenge of making simultaneous multi-epoch radial velocity measurements.

The scientific case for doing so was strong (see essay 86), as was the resistance to making observations from space that could, even if only in principle, be made from the ground. It is thus gratifying to read in Trabucchi et al. (2023), the statement that *'The availability of simultaneous photometric measurements gives a unique added value to the Gaia catalogue'.*

159. Improved solar system astrometry

Gaia Data Release 3 was issued in June 2022, and the next major release, DR4, is not scheduled until 2025. But on 10 October 2023, the Gaia Data Processing and Analysis Consortium (DPAC) published five papers forming part of a special 'Focused Product Release'. I looked at the first two in essays 157 and 158.

The third, by David et al. (2023), concerns the epoch astrometry and orbit re-construction for 157 000 asteroids. I will explain how the data in this study differ from earlier releases, and summarise their main results.

I described the observational principles of solar system objects in essay 64. Briefly, and in common with stars and other point-like images, sources are detected as they enter Gaia's two fields of view (as the satellite scans the sky) as long as they are brighter, at each specific observation epoch, than the detection limit.

As foreseen in the mission's early studies, the multi-epoch astrometry of all objects brighter than 20–21 mag would yield a deep and uniform detection of minor planets and other bodies (including near-Earth asteroids and Kuiper belt objects), permitting profound studies of their dynamics, structure, and taxonomy.

This is important because these minor bodies retain a record of the conditions in the proto-solar nebula, and their properties can therefore provide great insights into the formation and evolution of our solar system.

For solar system objects, DR2 (April 2018) was based on 22 months of observations (July 2014–May 2016), and provided limited data on around 14 000 asteroids. EDR3 provided no further updates.

DR3 (June 2022, essay 76) was based on 34 months of observations (July 2014-May 2017), and provided astrometry for 158 000 solar system objects, orbits for 154 787, and BP/RP reflectance spectra for 60 518.

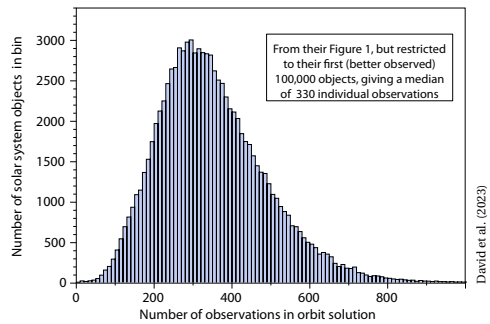
Orbit determination based on DR3 was described by Tanga et al. (2023), who also gave a summary of the associated photometric and spectral data. There is no further discussion of the photometric and spectral data in this latest 'Focused Product Release'.

In this context, the work reported by David et al. (2023) treats the *same* 157 000 asteroids as in DR3, but with a much greater observational time coverage: being 66 months, compared with the 34 months of DR3. In fact, the 66-month coverage corresponds to the time interval being used for DR4, whose full release is only expected in late 2025.

Importantly, this new data set covers more than a full orbital period for the majority of objects, and this critical time coverage results in a transformational improvement in the accuracy of the orbit reconstruction.

The terminology used here to describe an observation, or observation epoch, merits clarification. A passage or transit across one of the two Gaia fields of view comprises (a maximum of) nine position measurements at CCD level, together referred to as an (elementary) 'observation'. The nature of the satellite scanning law is such that a (varying) number of successive transits generally occurs within a 1–2 day period, followed by a gap of (typically) several weeks before the start of a new period. The clustered observations within such a visibility period are considered to define an observation 'epoch'.

The 100 000 objects best observed in the sample of David et al. (2023) have a median of 330 elementary observations over the 66 month period, corresponding to an average of around 30 ± 10 visibility periods or epochs. The range of distinct epochs, and the orbit coverage (averaging about 1.2 orbital periods for the best-observed), strongly influences the orbit accuracy.



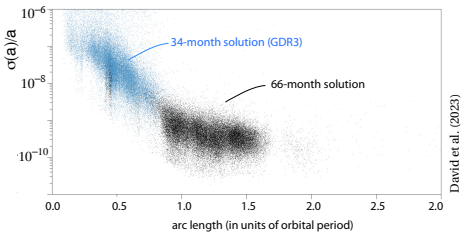
THE OTHER factor strongly influencing the final orbit accuracy is the (along-scan) accuracy of the elementary observations, which is better than 1 mas for $G < 18$ mag, degrading to ~10 mas at $G \simeq 20$. Compared to historical ground-based observations, often over more than a century, the Gaia observations are of unprecedented accuracy and homogeneity, and reasonably densely sampled, but extend over a limited duration that is, nonetheless, sufficient to allow good orbit fitting.

I should emphasise that, as in GDR3, the study considers only those asteroids with known (elliptical) orbits; the orbit determination of new discoveries are made within another dedicated 'pipeline' (Carry et al., 2021).

FOR THE ORBIT DETERMINATION, heliocentric positions are computed from numerical integration within the solar system model INPOP19a (Fienga et al., 2019), which considers the Earth–Moon barycentre with the seven other planets, the dwarf planet Pluto, a selection of 343 asteroids, and the ten most massive trans–Neptunian objects (TNO).

Mutual asteroid perturbations can be significant. In contrast to other major databases of asteroid orbits (including JPL and the Minor Planet Center, MPC), these are not yet included in the present orbit solutions.

Orbit fitting for each asteroid, determined as a least-squares fit to the Gaia data, results in the published 6-d 'state vector', which specifies the initial conditions (viz. heliocentric position and velocity vectors within the ICRF reference frame) at the epoch corresponding to the midpoint of the observations. All times are referred to TCB (Barycentric Coordinate Time).



This figure shows the formal uncertainty of the semi-major axis, σ_a/a, as a function of the arc length, for the current 66-month solution (black dots) and the earlier 34-month GDR3 solution (blue). The improvement once the observations cover a full orbit is clearly visible.

WHILE THE TRANSFORMATION from state vector to the more usually published (unperturbed) 'osculating elements' is considered routine in celestial mechanics, the unprecedented accuracies of Gaia highlight the fact that there is, as yet, no agreed transformation from the ICRS reference *system* to the ecliptic reference system. The issue is a complex one (their Section 4.6), and potentially complicates a rigorous comparison between the Gaia orbits and those from say, JPL or MPC.

THE STUDY by David et al. (2023) includes a detailed discussion comparing the 66-month Gaia orbits with the best existing orbits published by JPL and the Harvard Minor Planet Center (MPC).

They focussed on comparing the resulting semi-major axis which, in contrast with the angular Keplerian parameters, has the merit of being a true geometric quantity that is independent of the reference frame.

They concluded that, despite the different observation sets, their time coverage, and their detailed differences in dynamical modelling, there is an impressive consistency between the various solutions. In particular, they demonstrated systematic relative differences in semi-major axis, σ_a/a, that are below 5×10^{-10}.

ANOTHER WAY of confirming orbit accuracies is to use stellar occultations observed from Earth.

Occultations events are, today, predicted on the basis of stellar astrometry from Gaia (essays 24, 137). But a test of the orbit accuracies of solar system objects can be made by comparing the 66-month orbits with the timings of successful occultations. Their test sample comprised 978 main-belt asteroids, associated with 5774 astrometric occultation measurements.

The observed minus computed (O–C) values showed a large spread, reaching several hundred milliarcsec. But when only occultation events contemporary with the Gaia mission were selected, the residuals were much smaller, peaking at ~10 milliarcsec, and with a substantially suppressed tail of high values.

THE NEXT MAJOR improvements to these existing orbits will come from the combination of the larger number of observations, and the extended arc length, of the approximately 10-year DR5 solution, whose release is expected around 2030.

This will essentially double the time interval of the 66-month data used by David et al. (2023), further improving the accuracy of the orbit determination of small solar system bodies. They anticipate, for example, that the final accuracies of the Trojan orbits will then be as good as those of the main-belt asteroids today.

AMONGST THE 'visualisations' of Gaia Sky (essay 147), I referred to some of those constructed from earlier data releases which illustrate, powerfully and visually, the direction of this work. For example, this tour through the asteroids is based on DR2. These orbits of 154 741 asteroids are based on DR3, colour-coded according to (reflectance) colour, which serves as a proxy for chemical composition. Objects are clearly resolved into the main-belt asteroids, Mars crossers, Jupiter Trojans, near-Earth objects, and the brightest 24 trans-Neptunian objects.

This recent Gaia Sky video illustrates the remarkable precision of the new Gaia Focused Product Release, based on the asteroid 18520 Wolfratshausen.

160. More on diffuse interstellar bands

GAIA DATA RELEASE 3 was issued in June 2022, and the next major release, DR4, is not scheduled until 2025. But on 10 October 2023, the Gaia Data Processing and Analysis Consortium (DPAC) published five papers forming part of a special 'Focused Product Release'. I looked at the first three in essays 157–159.

The fourth, by Schultheis et al. (2023), concerns further detailed spatial mapping of two diffuse interstellar bands, or DIBs. I will explain the context, describe the data used, and summarise their main results.

I HAVE ALREADY devoted a previous essay (essay 92, October 2022) to some early results on the mapping of DIBs with Gaia. Let me summarise a few key points, but refer to that essay for further background.

Briefly, DIBs are absorption features seen in stellar spectra, superimposed on a smoother wavelength-dependent interstellar extinction. They were first reported just over a century ago, but their origins still remain something of a mystery. Several hundred bands have now been discovered, at ultraviolet, visible and infrared wavelengths (e.g. Fan et al., 2019). And while there is a general consensus that they are probably due to large and complex molecules in the Galactic interstellar medium, only one carrier has been securely identified: ionised buckminsterfullerene ($C60^+$), with five absorption bands in the near-infrared (Linnartz et al., 2020).

The choice of wavelength range for Gaia's radial velocity spectrometer was driven by a number of considerations (see essay 85). Amongst these, and within the adopted wavelength range (845–872 nm), was a known medium-intensity DIB at 862 nm, first identified in 1975. The correlation of its equivalent width with absorption was known to be rather tight, suggesting that it could be used to build a detailed reddening map, especially for high values of interstellar extinction (Munari, 1999).

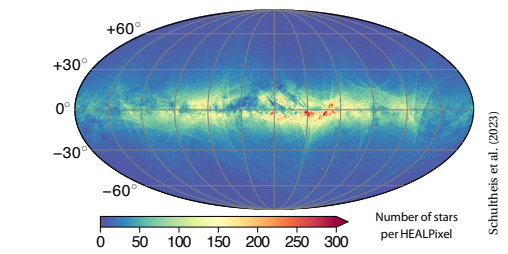
The first major study of the 862 nm DIB, from Gaia DR3, was reported by Schultheis et al. (2023). Their sample contained 476 117 stars with valid DIB measures, with their high-quality sample comprising 141 103 stars. The former was, by an order of magnitude, the largest homogeneous full-sky sample obtained to date.

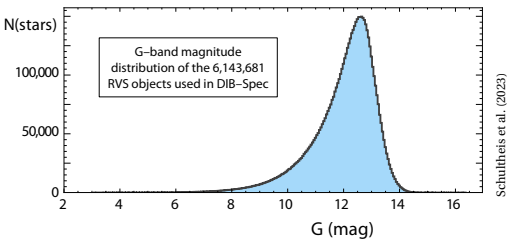
ALTHOUGH THE strong line at 862 nm was the only confirmed DIB in Gaia's RVS spectral range at the time of its consideration, Munari et al. (2008) also considered a weaker one at 864.8 nm to be another candidate. From 8458 DR3 spectra, Zhao et al. (2022) confirmed its existence, and showed that both equivalent widths are well correlated, albeit with the 862 nm line being better correlated with the E(BP–RP) extinction.

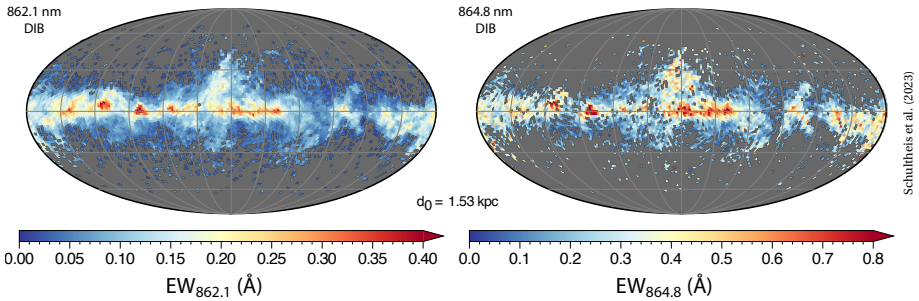
That is, in outline, the status of the DIB investigations with Gaia until the present 'Focused Product Release' study reported by Schultheis et al. (2023).

THE LATEST STUDY, by Schultheis et al. (2023) makes use of same DR3 data, but with an improved 'pipeline' treatment of the relevant RVS spectra to address two major limitations of the earlier studies.

The first was that, due to the limited signal-to-noise ratio of the individual RVS spectra, the 862.1 nm DIB could be successfully measured in only about 10% of the sources. Restricting the sample to more reliable high-quality measurements further reduced the sample size to around 140 000 sources. Another limitation was in the use of synthetic spectra in extracting the DIB signal.







Galactic distributions of the two DIBs at a distance $d = 1.53$ kpc. Other distance distributions are given in the study paper.

SCHULTHEIS ET AL. (2003) have developed a new module within DPAC Coordination Unit 8, responsible for source classification and parameterisation (essay 89). The module, DIB–SPEC, processed all 6.8 million RVS spectra, and made the new DIB measurements.

Based on the RVS spectral quality, and the results of the existing CU8 processing module GSP–SPEC, DIB–SPEC constructs two samples: a set of 'target' stars whose spectra are expected to contain possible DIB signals, and a set of 160 000 'reference' stars, at high Galactic latitudes ($|b| \geqslant 65°$), and covering a wide range of stellar parameters, less likely to contain DIB features.

The reference stars effectively define the corresponding contribution from the ISM spectra for each target star. Matching each target spectrum to its closest reference spectra in stellar parameter space then allows an empirical removal of the stellar spectrum, i.e. without reference to stellar models, and leaving a set of 6 million spectra characterising the interstellar medium (ISM).

Using the sky coordinates and parallax distances, they then assigned each ISM spectrum to a 3-d grid (of 'VOlume piXELs', or 'voxels') defined by Galactic coordinate and heliocentric distance, with an angular size of 1.8° (level 5 HEALPix), and 29 unequally sized distance shells. DIB–SPEC then stacks the spectra of the target stars within a given 'voxel', and fits the two DIBs in the stacked ISM spectra.

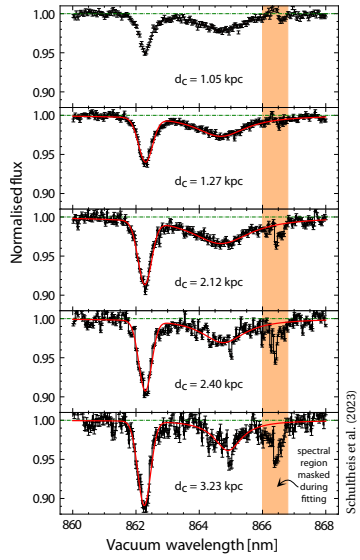
DIB–SPEC measures both DIB signals (at 862.1 nm and 864.8 nm) out to larger distances than DR3, so defining the large-scale spatial distribution of the DIB carriers. The results included in this specific 'Focused Product Release' include the best-fit parameters of the two DIBs (depth, central wavelength, width, and equivalent width), along with the stacked ISM spectra (i.e. containing only the interstellar features) in each voxel.

The width of the DIB profile, for example, contains substantial information about the properties of the intervening ISM clouds and the DIB carriers, such as the profile broadening that is related to the gas kinetic and rotational temperature in different physical environments.

AMONGST THEIR conclusions, they found that the strength and spatial distribution of the 862.1 nm DIB in the stacked spectra are consistent with what was found from Gaia DR3, but with a much higher signal-to-noise ratio, a higher spatial resolution, and extending to much larger distances. This allowed them to trace DIBs in the outer spiral arm, and beyond the Scutum–Centaurus spiral arm.

They produced all-sky maps, below Galactic latitudes ±65°, and out to ~4000 pc, of both DIB features and their correlations. They found a reasonable correlation between the dust reddening found from stellar absorption and the equivalent widths of both DIBs. And they detected the signals of the 862.1 nm feature inside the Local Bubble ($\lesssim 200$ pc).

AT PRESENT, this significant advance in mapping these two specific DIBs, out to much larger distances, nevertheless leaves their origin as an unknown!



Example fits to the 862.1 nm and 864.8 nm DIBs (red lines) in stacked ISM spectra for five distance voxels in one direction.

161. Strongly lensed quasars

Gaia Data Release 3 was issued in June 2022, and the next major release, DR4, is not scheduled until 2025. But on 10 October 2023, the Gaia Data Processing and Analysis Consortium (DPAC) published five papers forming part of a special 'Focused Product Release'. I looked at the first four in essays 157–160.

The fifth, by Krone-Martins et al. (2024), is a new search for strongly lensed quasars, which I look at here.

I have already discussed the application of Gaia data to quasar research in previous essays, notably their role in defining the celestial reference frame (essay 27), and consequently in detecting Galactic aberration (essay 32), and perhaps our Galaxy's 'tumbling' disk within the dark matter halo (essay 95). In addition, of just over 1 million previously known quasars analysed in DR3, a host galaxy was detected around 65 000 (essay 82).

Already by the end of 2020, nearly 100 publications had used the DR1 or DR2 data to examine the reliability of earlier quasar surveys, and to quantify effects important in refining the extragalactic reference frame link.

As an aside to the main topic of this essay, and in an attempt to bring the Gaia/quasar research field more up-to-date, let me mention the Gaia–unWISE all-sky quasar catalogue, Quaia (Storey-Fisher et al., 2024).

This merges the 6 649 162 quasar candidates from Gaia DR3 having redshift estimates from its low-resolution BP/RP spectra (Delchambre et al., 2023) with the unWISE infrared catalog (reprocessed from WISE) to construct a large volume spectroscopic quasar sample suitable for cosmological studies. It has 1 295 502 quasars with $G < 20.5$ mag, and 755 850 candidates in an even more well-defined sample to $G < 20.0$ mag.

Amongst its application, Alonso et al. (2023) examined the angular clustering of the quasars to derive cosmological constraints on the amplitude and growth of structure, importantly breaking the usual degeneracy between the matter density parameter (Ω_m) and the amplitude of the matter power spectrum (σ_8), yielding $\Omega_m = 0.343^{+0.017}_{-0.019}$ and $\sigma_8 = 0.766 \pm 0.034$.

An important contribution of Gaia to quasar science is in detecting and classifying gravitationally lensed systems (essays 16 and 58). On-board CCD sampling provides source images with an ultimate resolution of 0.1–0.2 arcsec in the scan direction, much better than can be routinely measured from ground. Accordingly, Gaia is able to discover new (and numerous) lensed quasars with image separations below ~1 arcsec.

Strongly lensed quasars, i.e. those which result in multiple *discrete* images, provide an important diagnostic for many applications. These include measuring the Hubble constant at intermediate cosmic distances (Chen et al., 2019b; Shajib et al., 2020); constraining the properties of dark matter (Gilman et al., 2019; Nierenberg et al., 2020); inferring structure near the event horizon of supermassive black holes (Pooley et al., 2009; Chartas et al., 2016); measuring the accretion disk size (Blackburne et al., 2011; Fian et al., 2021); constraining the quasar broad emission-line region geometry (Sluse et al., 2011; Braibant et al., 2017); measuring black hole spin (Walton et al., 2015; Middleton, 2016); and testing general relativity in the strong-gravity regime (Collett et al., 2018; Melo-Carneiro et al., 2023).

Amongst the most keenly hunted are the quadruply-imaged 'Einstein crosses' (essay 58), which are some of the most constraining for detailed physical modelling.

To give a flavour of the ongoing studies, a series of papers 'Gaia GraL' (for gravitational lensing) has focussed on the lensed systems discovered in Gaia DR2.

These have dealt with: new quadruple candidates around known quasars (Krone-Martins et al., 2018); the statistics of previously known quadruple-image quasars, where 12 systems were found to have all four images detected (Ducourant et al., 2018); a systematic blind search for new lensed systems (Delchambre et al., 2019); ground-based confirmation and modelling of new discoveries (Wertz et al., 2019; Krone-Martins et al., 2019; Stern et al., 2021); XMM–Newton X-ray observations of discoveries between $z = 1-3$ (Connor et al., 2022); and a radio census of 24 lensed systems (Dobie et al., 2024).

A COMPILATION of known lenses (with images) is held at research.ast.cam.ac.uk/lensedquasars. Of these, for example, some 80 quadruple-image systems are known today (Stern et al., 2021).

Most lensed systems currently known have image separations above 1 arcsec (e.g. Ducourant et al., 2018). Nevertheless, the expected distribution of lenses is estimated to peak at somewhat smaller separations, below 1 arcsec, making most of them more-or-less undetectable from the ground.

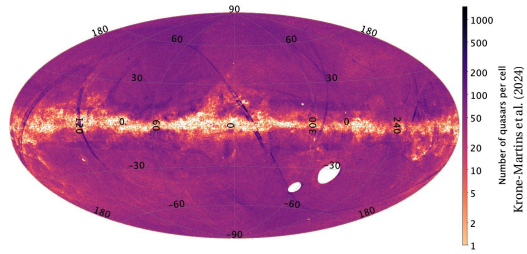
Estimates of the lensed quasars detectable by Gaia (in standard ΛCDM cosmology) have suggested that ~3000 multiply imaged quasars could be detected, of which > 250 would have ≥ 3 images and the rest doubles (Finet & Surdej, 2016). By providing a homogeneous survey at ~180 mas angular resolution, and with precise astrometry for all lensed images, Gaia should yield a significant increase in the number of known lensed systems.

But even the GDR2 and GDR3 catalogues are known to be incomplete at separations ≲ 2 arcsec (Arenou et al., 2017; Arenou et al., 2018; Fabricius et al., 2021). This is the result of a robust selection on the astrometric and photometric quality indicators of the published sources.

CAPITALISING ON Gaia's unprecedented angular resolution of ~0.18 arsec in the optical, and in order to improve the completeness of lensing searches at small angular separations, Krone-Martins et al. (2024) developed the GravLens pipeline, implemented within the Gaia DPAC Consortium. Their goal was to analyse all Gaia detections around quasars, and to produce a catalogue of secondary sources around each of them, along with their mean astrometry and photometry.

Their dedicated processing started from an input list of quasars, resulting from the merger of a number of major catalogues of quasars and AGN (active galactic nuclei) candidates, including the Milliquas catalogue (Flesch, 2023); the R90 and C75 selections of the AllWISE catalogue (Assef et al., 2018); a catalogue of AGN candidates (Shu et al., 2019); a subset of the GDR3 quasar candidates (Bailer-Jones et al., 2023); and additional quasars whose morphology was analysed by Ducourant et al. (2023). Most contain stellar contaminants. Their resulting starting list contained 3 760 480 quasars or quasar candidates with an entry in GDR3.

Of the data from the first 3 years of operations, corresponding to the time interval of DR3, they analysed 183 368 062 transits matched to the 3 760 480 quasars from their initial list. GravLens then identified all Gaia transits within 6 arcsec of each quasar, based on machine-learning (specifically the DBSCAN unsupervised clustering algorithm), and without relying on the standard Gaia object cross-matching. The output is a list of sources within 6 arcsec of each quasar, with mean positions, fluxes, and magnitudes of the components.

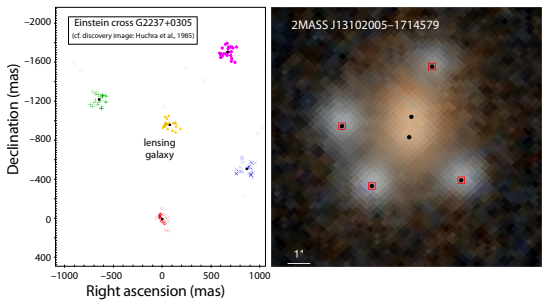


Distribution of quasars in the input list, in Galactic coordinates.

THEY FOUND a total of 4 760 920 sources, including the quasars themselves, within 6 arcsec of the quasar positions. Of these, 103 000 are new sources complementing those already present in GDR3. In most cases (87%), the quasar is a single source, without close neighbour. In 501 385 cases, neighbouring sources were detected, of which 9% have two components. There are 159 000 systems (4%) with more than two components. They also found 9000 multiplets with more than 10 components, generally corresponding to extended galaxies decomposed into many sources.

Their resulting GravLens catalogue includes 450 known or candidate lenses previously published in the literature, 76 with four images, and the rest being two-component systems. For 67 of the 76 quadruple systems, GravLens complements the existing measures from GDR3 by measuring one or more additional components, or the deflecting galaxy itself. In total it found 1293 components in the fields of known lenses, of which 1207 were present in GDR3. The 86 newly detected components in the fields of known lenses are mostly faint components surrounding lenses with a small number of existing GDR3 counterparts.

To search for new lenses, they employed two methods: an outlier 'scoring' algorithm, and an application of the Extremely Randomised Trees algorithm (Delchambre et al., 2019), making use of Gaia spectra when available. They identified 381 new lensed candidates, of which they considered 49 as the most promising. These have angular sizes from 1.03–5.97 arcsec, with a minimum image separation of 0.41 arcsec.



Two examples of the Gaia multiple-lens detections (Krone-Martins et al., 2024; Figs 3 and 8b)

162. Cosmology with Gaia's quasars

IN THE EARLY studies of Gaia in the 1990s, we recognised that the large numbers of measurable quasars, 500 000 or more, would contribute fundamentally to the determination of the quasi-inertial reference frame (e.g. Bachchan et al., 2016; Gaia Collaboration et al., 2021a).

Today, Gaia's quasar survey is also being applied to two topical observational questions in ΛCDM cosmology: the kinematic dipole anomaly, and the S_8 'tension' in the amplitude of cosmological structures.

WHILE ΛCDM has passed a vast range of impressive experimental tests, there remain big gaps in the theoretical understanding of its key ingredients, including the nature of dark matter and dark energy, the asymmetry of matter versus antimatter, and whether the dimensionless parameters of physics evolve with time.

Among observational questions are certain discrepancies in the cosmological model parameters inferred from different data sets. Prominent amongst these is the $\sim 5\sigma$ 'tension' in the cosmic expansion rate H_0 as measured from the cosmic microwave background compared with more local distance calibrators (essay 44).

There is a similar discrepancy in the the value of the dipole amplitude of the cosmic microwave background radiation, and the 'kinematic dipole' defined by the distribution and motion of objects on scales approaching the Hubble length (e.g. Rubin & Heitlauf, 2020; Ananthanarayan & Mohanty, 2021; Mohayaee et al., 2021; Peebles, 2022; Lombriser, 2023). It arises as follows.

The amplitudes of the spherical harmonic expansion of the CMB temperature of degree $\ell > 1$ are interpreted as remnants of the decoupling of acoustic oscillations in the early Universe. But the dipole amplitude, $\ell = 1$, is much larger than predicted from this effect, and is instead attributed to the Sun's motion relative to the CMB of amplitude 370 km s^{-1} in the Galactic coordinate direction $l = 264°$, $b = 48°$ (Aghanim et al., 2020).

Adjusting for the solar motion within our Galaxy, and thereafter for the barycentric motion of the Local Group, implies that the Local Group is moving relative to the CMB at 620 km s^{-1} in the direction $l = 272°$, $b = 30°$.

OBJECTS DISTANT enough that their distribution should approximate to a homogeneous isotropic universe should display a similar dipole anisotropy in their number counts (Ellis & Baldwin, 1984). Early tests at both radio (Baleisis et al., 1998; Blake & Wall, 2002) and X-ray wavelengths (Scharf et al., 2000) have been followed by many similar studies, arguing both for and against a discrepancy (see Peebles, 2022, §3.2.4).

Recent tests have been made using a million quasars from the WISE survey based on their mid-infrared colours (Secrest et al., 2021; Singal, 2021). A joint analysis of various catalogues of radio galaxies *and* quasars concluded that their sky distribution is *inconsistent* with the standard ΛCDM model, at the 5.1σ level, similar to that of the Hubble tension (Secrest et al., 2022).

BUT THE PENDULUM has swung the other way with the construction of the Quaia quasar catalogue (Storey-Fisher et al., 2024). This merges the 6 649 162 quasar candidates from Gaia DR3 having redshift estimates from its low-resolution BP/RP spectra (Delchambre et al., 2023) with the unWISE infrared catalogue (reprocessed from WISE) to construct a large volume spectroscopic quasar sample. It has 1 295 502 quasars with $G < 20.5$ mag, and 755 850 candidates in an even more well-defined sample to $G < 20.0$ mag.

Taking into account selection effects, and especially contamination near the Galactic plane, Mittal et al. (2024) concluded that the Quaia quasar dipole *is* consistent with the cosmic microwave background dipole, in terms of both amplitude and direction. This lends further support to the ΛCDM model, in the sense that the observed quasar dipole can be attributed to our local departure from the Hubble flow.

ANOTHER EXPLANATION for the cosmic dipolar tension has been given by Cai et al. (2025). They postulate that it originates from a gigaparsec-scale void, whose own existence has been discussed as another observational discrepancy facing ΛCDM (e.g. Peebles, 2022, §4). Evidently, the final word has not yet been written!

IN THE SECOND of these recent applications of the Gaia quasar data to cosmological studies, I will return to a topic that I made passing reference to in essay 161: a study of their angular clustering in order to place constraints on the amplitude and growth of cosmological structure, and to compare these observations with the predictions of ΛCDM cosmology.

This is a contribution to what is referred to as the 'S_8 tension'. This is a distinct discrepancy, of around 3σ, in the amplitude of the predicted matter density fluctuations based on the cosmic microwave background data, and the many observations, at low redshifts, inferred from direct measurements of galaxy clustering data, or weak lensing ('cosmic shear') data.

Incidentally, the principle of these 'weak lensing' surveys is that distant radiation passing massive foreground objects undergoes a deflection whose magnitude depends on the total lensing mass, whether luminous or dark. The deflection distorts the observed shape of extended galaxies, in a manner which is spatially correlated. The mass distributions that cause the lensing can then be constrained from these spatial correlations.

The Gaia contribution that I will discuss here is also based on the Gaia–unWISE all-sky quasar catalogue, Quaia (Storey-Fisher et al., 2024). But let me first explain the context in a little more detail.

THE AMPLITUDE of density fluctuations, of galaxies or clusters, is characterised by this S_8 parameter, defined as $S_8 = \sigma_8(\Omega_m/0.3)^{0.5}$, where σ_8 is the variance of the linear matter overdensity field in spheres with a $8h^{-1}$Mpc radius, and Ω_m is the fractional energy density of non-relativistic matter, both defined at $z = 0$.

The value of S_8 has been measured by various galaxy clustering and weak lensing surveys, notably from the Dark Energy Survey (Abbott et al., 2018; Troxel et al., 2018), from Subaru (Hikage et al., 2019), and from KiDS (Heymans et al., 2021). These, and a number of others, have resulted in various degrees of 'tension' compared with the value inferred from the Planck satellite. The best-constrained cosmic shear data tend to suggest smaller values of S_8 than the Planck value, 0.832 ± 0.013.

From combined data, García-García et al. (2021) derived $S_8 = 0.7781 \pm 0.0094$, with errors comparable with the Planck CMB measurements, but deviating by 3.4σ. They also reconstructed the *evolution* of S_8 from $z \sim 2$ to the present epoch, and argued that the main contribution to the tension in the measured value of S_8 is due to a lower amplitude of fluctuations than the Planck prediction in the range $z = 0.25 - 0.7$, unfortunately with limited constraints at lower or higher redshifts.

The importance of resolving the S_8 tension is that, if systematic uncertainties in the microwave background or cosmic shear analyses can be excluded, the results may point to new physics giving rise to a slower growth of perturbations at later times (e.g. Adil et al., 2024).

FURTHER CONSTRAINTS, and an extension of the redshift range over which the growth history can be reconstructed, can be derived from quasar clustering measurements, both through three-dimensional analyses of their auto-correlation (Laurent et al., 2017; Castorina et al., 2019; Neveux et al., 2022), as well as their cross-correlation with other tracers.

With this objective, García-García et al. (2021) also considered 343 708 objects with measured redshifts in the range $z = 0.8-2.2$ covering 4800 deg^2 from the SDSS–eBOSS survey. But as the sparsest of their clustering samples, it provided limited new constraints.

AND SO to the constraints enabled by the Gaia quasar data, again making use of the Quaia catalogue resulting from the merger of the 6 million quasar candidates from Gaia DR3 with the unWISE infrared catalogue (Storey-Fisher et al., 2024). This large-volume spectroscopic sample, with redshifts, has 1 295 502 quasars with $G < 20.5$ mag.

Alonso et al. (2023) examined the angular clustering of the quasars in the Quaia catalogue to derive improved constraints on the amplitude and growth of structure. The main attributes of the Quaia catalogue in this respect are its full-sky coverage, high spatial density, and redshift coverage to $z \lesssim 4$ with good redshift accuracy (e.g. $|\Delta z/(1+z)| < 0.01$ for 62% of the sources). They used the full $G < 20.5$ Quaia catalogue, divided into two redshift bins ($z < 1.47$ centred on $z = 1.0$, and $z > 1.47$ centred on $z = 2.1$) each containing approximately the same number of sources.

This allowed them to constrain the growth of structure over a range of cosmic times, complementary to those accessed by previous analyses, but importantly giving access to significantly larger spatial scales.

There is an important consequence of this broad coverage: since it is usually not possible to fully break the degeneracy between growth and geometry with existing datasets, results are generally reported in terms of the combination $S_8 = \sigma_8(\Omega_m/0.3)^{0.5}$. The large-scale coverage of the Quaia/Gaia sample allows independent estimates of σ_8 and Ω_m, thus providing a constraint on the amplitude of matter fluctuations, $\sigma_8 = 0.766 \pm 0.034$, and on the fractional abundance of non-relativistic matter, $\Omega_m = 0.343^{+0.017}_{-0.019}$, together yielding $S_8 = 0.819$.

THEIR STUDY concluded that the Quaia quasar results are compatible with Planck at the 1.4σ level, and therefore do *not* provide support for the 'S_8 tension'. Nevertheless, their higher redshift sample favours a value of σ_8 that is lower than the Planck measurement at the $\sim 2.5\sigma$ level.

Future Gaia releases should give improvements in quasar numbers and redshift accuracy. Meanwhile, we have hints that the Gaia quasar sample has alleviated both the dipole anomaly, and the S_8 tension.

163. Dual active galactic nuclei

THIS ESSAY IS another on the theme of Gaia's contributions to cosmology. It is centred on the perhaps surprising fact that Gaia can identify *binary* supermassive black holes in active galaxies, and with sufficiently small angular separations that can pinpoint in-spiralling black holes well on their way to their eventual merger.

THE REALISATION that active galaxies and quasars host supermassive black holes at their centres is a long story. It dates back some 60 years to Maarten Schmidt's early attempts to explain the emission-line spectrum of the radio source 3C 273, and Hoyle & Fowler's proposal that H-burning supermassive stars could explain the compact dimensions and high energy output of quasars.

In 1970, Wolfe & Burbidge argued that the large velocity dispersion of stars in the central regions of elliptical galaxies could only be explained by a large mass concentration, of order $10^{10} M_{\odot}$. Dynamical evidence for a massive dark object at the core of M87 was presented in 1978, and further supported by high-angular resolution observations by the Hubble Space Telescope in 1994.

Fast forwarding over many other theoretical and observational contributions, and other candidate supermassive black holes (including at the centre of our own Galaxy), the consensus today is that all quasars, Seyfert galaxies, and active galactic nuclei (AGN) in general, host a matter-accreting massive black hole at their cores.

LEAVING ASIDE details of their initial formation, especially of primordial black holes early on in the history of the Universe, cosmological models of structure formation within the ΛCDM framework all predict the widespread existence of a population of *dual* supermassive black holes inside their common host galaxy.

Such dual supermassive black hole systems are held to be the result of galaxy mergers. The in-spiralling black holes, within the merged galaxy, will themselves eventually merge, giving rise to a burst of intense gravitational waves. The first observation of such gravitational waves was made in 2015, when a black hole merger signal was picked up by the LIGO gravitational wave detectors.

AT A TIME when there was little observational evidence for the existence of massive *binary* black holes, Yu (2002) studied their occurrence, merger history, and lifetimes in the hierarchical model of galaxy formation, and examined the consequences *'... if no evidence of binary black holes is found in active galactic nuclei'*.

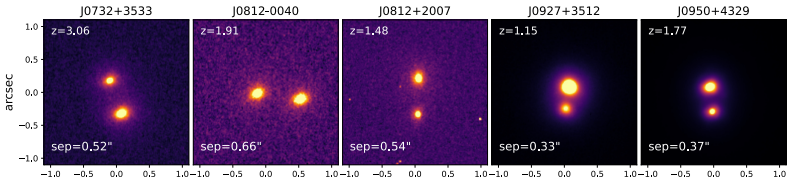
To my knowledge, the first clear evidence for a supermassive binary black hole system came with the Chandra X-ray observations of 3C 75, the double radio source at the centre of Abell 400. Hudson et al. (2006) resolved the two active galactic nuclei, concluding that they are a bound system resulting from a previous merger.

These and subsequent observations were guided by significant improvements in large-scale hydrodynamic cosmological simulations, which began to make detailed predictions for the evolution of the baryonic structures in the Universe, including galaxy mergers.

Drawing on these simulations of galaxy mergers, close pairs of supermassive black holes can form when both (pre-merger) progenitor galaxies host such an object. When separated by a few kiloparsec, they can be observed either as dual AGN (when both black holes are observed as active), or as 'offset AGN' (and displaced from the galaxy centre) either when only one is active, or when signalling a recoiling black hole after a binary merger (Steinborn et al., 2016; Capelo et al., 2017).

OBSERVATIONAL EFFORTS to find dual AGN have intensified over the past decade, in part to verify the findings of these structure formation and merger simulations. And although simulations suggest that the fraction of dual active galactic nuclei increases with redshift (Rosas-Guevara et al., 2019), progress is challenged by the fact that discovering them, at kpc-scale separations, even in the local Universe, requires high-resolution observations of large numbers of candidate objects.

I should also stress that the appearance of multiple active galactic nuclei at small angular separations can also be due to gravitational lensing of single sources (essay 161), with follow-up spectroscopy often being crucial in distinguishing between the two effects.



High-resolution Ks band Large Binocular Telescope images (2.2×2.2 arcsec^2) of five Gaia 'multi-peak' AGN discoveries, showing redshifts and separations (Mannucci et al., 2023).

A NUMBER OF candidates have been discovered with SDSS spectroscopy, from which Liu et al. (2011b) estimated the fraction of dual AGN at just 1.3% within 30 kpc. The first triple system was also discovered from SDSS spectroscopy by Liu et al. (2011a): the three black holes, with masses $\sim 10^8 M_\odot$, are predicted to merge in ~40 Myr. Others were found from the SDSS–MaNGA spectroscopic survey (Fu et al., 2023). Bhattacharya et al. (2023) searched one million SDSS images, finding 159 dual AGN candidates, with just two triple systems.

Some 100 candidates (but with just 3 confirmed) have been found with the Hyper Suprime–Cam on Subaru (Silverman et al., 2020). High-resolution follow-up observations have been made with JWST–NIRSpec (Perna et al., 2025; Perna et al., 2023), and at multi-wavelengths including X-ray by De Rosa et al. (2023).

While a key question is whether their numbers match the predictions of cosmological simulations, especially at high-z (Perna et al., 2025), it remains the case that only a few confirmed examples have been found at $z > 0.5$: Mannucci et al. (2022) noted that just four had separations below 8 kpc at $z > 1$.

I WILL NOW focus on some recent contributions with Gaia, where the instrument's sub-arcsec along-scan angular resolution is generating its own survey of active galactic nuclei. Various approaches are being followed.

The first, which the authors have dubbed 'varstrometry' for 'intrinsic variability induced astrometric jitter', exploits one of the approaches used for the astrometric detection of unresolved binary stars with Hipparcos, viz. one component varying in luminosity leads to the astrometric motion of the system's *photocentre*. The phenomenon was already termed 'VIMs', for 'variability induced movers' (Wielen, 1996), and 288 multiple star solutions were classified as VIMs in the Hipparcos catalogue (Lindegren et al., 1997). A similar scheme is being exploited for Gaia (Pourbaix et al., 2022, §7.2.6).

As developed in the Gaia–AGN context by Hwang et al. (2020), the idea is to detect any temporal displacements of the photocentre in unresolved (sub-kpc separation) optically selected dual active galactic nuclei.

Shen et al. (2019) then used DR2 astrometry, including the '`astrometric excess noise`' (Lindegren et al., 2018, Eq. 12), to constrain the off-nucleus population of a low-redshift ($z = 0.3-0.8$) sample of active galactic nuclei. These were selected from the Sloan Digital Sky Survey as showing significant photometric variability.

They found that Gaia DR2 already provided strong constraints on the projected off-nucleus distance at these redshifts: 99%, 90%, and 40% must be centred to better than 1 kpc, 500 pc, and 100 pc, respectively. They concluded that 'offset AGN', those displaced by more than a few hundred pc, must be rare at low redshift.

Chen et al. (2022) used the same approach to detect unresolved quasar pairs, also based on Gaia DR2. Their follow-up snapshot imaging observations with HST–WFC3 showed a high fraction of resolved systems amongst the 45 observed, but without be able to distinguish dual nuclei from lensed systems.

Shen et al. (2023) searched for multiple Gaia sources corresponding to a single SDSS target, and found 60 resolved candidates, although again unable to distinguish dual nuclei from lensed systems.

Similar studies by Ji et al. (2023) identified 143 SDSS spectroscopically confirmed quasars that have multiple Gaia EDR3 detections within 1 arcsec of the quasar position. Based on spectral-fitting, they found two *bona fide* double quasars, and 56 other candidates.

Chen et al. (2023b) searched in DR2, but made their analysis with DR3 astrometry. Follow-up VLBA radio observations of 27 candidates revealed 8 systems with either multiple radio components, or system offsets.

A DIFFERENT approach was used by Mannucci et al. (2022), who searched for AGNs showing multiple peaks in their Gaia light profiles. All of their 221 candidates (at $z = 0.3-4$) followed up with high-resolution images (26 archival HST, and 5 from LBT) showed multiple components with sub-arcsec separation, confirming the high reliability of this selection method.

Various follow-up studies have been made for these multiple-peak candidates. Ciurlo et al. (2023) found 2 dual AGNs out of 4 candidates observed with integral-field spectroscopy. Gross et al. (2023) established one as a lensed quasar. Scialpi et al. (2024) observed 12 with VLT–MUSE, and classified 5 as dual AGNs, 2 as lensed systems, and 5 as chance alignments of foreground stars. Wang et al. (2023a) observed 5 candidates with e–MERLIN, and detected significant radio offsets in four.

DUAL ACTIVE GALACTIC NUCLEI, we now know, provide an important probe of the physical processes that drive the in-spiralling of supermassive black hole pairs inside a single merged galaxy. Difficult to detect and characterise, Gaia is making a significant contribution to their discovery.

164. Sub-subgiants and (tiny) black holes

FIFTY YEARS AGO, Hawking (1971) published a much-cited paper which posited that there may be a large number of primordial black holes of very small mass, 10^{-8} kg upwards, formed in the first second or so following the Big Bang. Their existence would obviate the need (through processes he detailed) for the very special initial conditions required to explain the Universe's high degree of isotropy, and indeed the very existence of galaxies. These primordial black holes might, in the process, provide a solution to the dark matter problem.

He contended that a mass of 10^{14} kg of such objects could have accumulated at the centres of sun-like stars. If such a star later became a neutron star, there would be … *a steady accretion of matter by the central collapsed object, which could eventually swallow up the whole star in about ten million years.*

Hawking further proposed that the Sun itself may harbour such a primordial black hole at its core, whose accretion would supply part of the solar luminosity.

LET ME PLACE that remarkable idea to one side (I will return to it later), and introduce the seemingly unrelated topics of subgiants, and blue and red stragglers.

First, subgiants. Their discovery, their role in the early understanding of stellar evolution, their importance in the age dating of the Galactic disk, and the insights provided by Hipparcos, were reviewed by Sandage et al. (2003). The term was first used by Strömberg (1930): their discovery did not fit the picture of stellar evolution developed by Russell in the years 1914–30, in which stars were believed to be born as giants near $M_V = 0$, after which (it was surmised) they contracted, becoming hotter, until reaching the main sequence.

A deep M67 colour–magnitude diagram provided the key to the now-accepted explanation, and post-1960 terminology (Johnson & Sandage, 1955): that the giant-branch is the locus of stars on the 'first ascent' of the post-main-sequence in a H shell burning phase, after core-H exhaustion. The He flash, at the top of the first-ascent branch, is followed by the descent to the core He-burning phase, populated by the 'clump' giants.

SUBGIANTS capture the transition from the main sequence to the red giant branch, and occupy the roughly 'horizontal' part of of the evolutionary track between the main sequence and giant stages: the start of H-shell burning is accompanied by an increase in radius, and a movement away from the main sequence. They lie just above the main sequence and to the left of the Hertzsprung gap for the more massive stars, and on the slow rise up the lower giant branch for low-mass stars. Their rapid evolution means that they are rare in the solar neighbourhood: Hipparcos observed only 12, including the nearest and brightest, β Hyi.

Blue stragglers, meanwhile, were first identified by Sandage (1953). Most easily identified in stellar clusters (by exploiting their common ages), they are more luminous, and have a higher effective temperature (therefore bluer) than the cluster's main sequence turnoff point. Inconsistent with standard stellar evolutionary theory – since they should have consumed their nuclear fuel, and evolved to become white dwarfs long ago – the consensus is that they are either binary stars in the process of merging (or that have recently merged), or a result of mass-transfer in a binary system (Mapelli et al., 2006). 'Yellow stragglers' are believed to be in an even more advanced evolution stage. A workshop in 2012 was devoted to their collective understanding (Boffin et al., 2015).

BLUE STRAGGLERS are not the central topic of this essay, but let me note that Gaia is contributing to their understanding, mainly by providing the much improved location of previously suspected candidates in the colour–magnitude diagram, and rejecting others.

The catalogue of Ahumada & Lapasset (2007) listed 1887 candidates in 427 open clusters. Using Gaia DR2, Rain et al. (2021) were only able to confirm a small subset, amounting to 897 blue stragglers, and 77 yellow stragglers, in 408 open clusters. Around the same time, also from DR2, Jadhav & Subramaniam (2021) found 868 blue stragglers in 228 clusters. And they established that their numbers *increase* with cluster age and mass. Using Gaia DR3, Li et al. (2023a) found a further 138.

THE NAMING became more confusing when Belloni et al. (1998) discovered a new class of star, in M67, that they termed 'sub-subgiants'. Defined empirically in the optical colour–magnitude diagram, they fall redward of the normal main sequence, while being *fainter* than the subgiant and giant branches (see also Mathieu et al., 2003). Similar stars in the globular cluster 47 Tuc were termed 'red stragglers' (Albrow et al., 2001). Other examples have since been discovered in both open and globular clusters, as well as in the field.

Geller et al. (2017) attempted to clarify the confusing nomenclature, and (with reference to their Figure 1) recommended use of the term 'sub-subgiant' to refer to stars that are redder than the normal main-sequence and *fainter* than normal subgiants, and the term 'red straggler' to stars similarly redder than the normal red giants but *brighter* than normal subgiants. The latter were originally postulated to be either former blue stragglers evolving toward the giant branch, or binary star systems, or 'something more exotic' (Clark et al., 2004).

Based on a literature search, and their location in the colour–magnitude diagram, Geller et al. (2017) identified 65 sub-subgiants and red stragglers in 16 open and globular clusters. They found that 58% are X-ray sources (with properties similar to RS CVn active binaries), and at least 65% are variables, 21 of which are radial-velocity binaries. Taken together, they suggested that binarity is important for sub-subgiant formation. They nonetheless emphasised that, unexplained by standard theories of single-star evolution, their origin remained a mystery.

THREE EXPLANATIONS for the existence of these sub-subgiants were evaluated by means of (MESA) evolutionary models by Leiner et al. (2017): binary mass transfer, envelope stripping, and magnetic activity. In the latter scenario, sub-subgiants result from the rapid rotation in subgiants or giants due to tidal synchronisation in a close binary. The strong magnetic fields inhibit convection, which in turn produces large starspots, inflated stellar radii, and lower mean surface temperatures and luminosities. Of the three mechanisms, the magnetic-field mechanism was predicted to create the largest numbers of sub-subgiants in open clusters.

Parallaxes and photometry from Gaia EDR3 were later used by Leiner et al. (2022) to position known RS CVn (active giant) binaries in the colour–magnitude diagram. They classified stars that fall below a 14 Gyr, metal-rich isochrone as candidate field sub-subgiants.

Out of a sample of 1723 RS CVn, they found 448 sub-subgiant candidates from Gaia, a factor seven more than the 65 previously known. They concluded that the ubiquity of sub-subgiants amongst this population indicates that they are a normal evolutionary phase for RS CVn-type systems, rather than rare by-products of dynamical encounters found only in dense star clusters.

ALTHOUGH THE EXISTENCE of sub-subgiant stars is perhaps, at last, explicable by the normal evolution of RS CVn-type systems, and not least thanks to the high quality astrometry and photometry from Gaia, a more exotic origin, for at least some, has been suggested by Bellinger et al. (2023). And this links to the possible existence of primordial black holes that I mentioned at the start, and in particular to the formation of stellar-mass black holes if they grow through stellar capture.

The central idea is as follows: the Milky Way is expected to contain $\sim 10^8$ black holes from normal stellar evolution pathways, with an average separation of 21 pc (Sweeney et al., 2022). If dark matter is comprised of primordial black holes at the classical Hawking evaporation limit of $10^{-18} M_\odot$, the number increases to 10^{30} black holes with an average separation of ~1 au, raising the possibility of their capture by stars (Ilie et al., 2021).

Primordial black holes are considered to be an attractive solution to the dark matter problem because they require no modification to the standard model of particle physics. Possible approaches to their detection include gravitational wave signatures (e.g. Phukon et al., 2021) and gravitational microlensing (Carr et al., 2024).

Neither is the possibility of black holes existing inside normal stars a new idea, with evolutionary implications already considered by Clayton et al. (1975).

Incidentally, normal evolution models (i.e. without a central black hole) suggest that the Sun will evolve via a crystallising white dwarf into a black dwarf, and thereafter live on for a preposterously long time. In the absence of proton decay, pycnonuclear fusion, due to zero-point oscillations of nuclei around their equilibrium point, will slowly process their composition to ^{56}Fe, over a lifetime of 10^{1100} yr (Caplan, 2020).

FROM EVOLUTIONARY models of stars with a central black hole, Bellinger et al. (2023) found that the lightest black holes have little evolutionary influence, while the more massive progressively consume the star, and with various observable consequences. Models of the Sun born with a central black hole of mass $10^{-6} M_\odot$ are, they argue, compatible with current observations.

In this scenario, the Sun would first fade to half its current luminosity over a period of 100 Myr as the accretion starts to generate enough energy to quench nuclear reactions. It would then expand into a fully-convective star, where it would radiate for several Gyr with an enriched surface helium abundance, first as a sub-subgiant, and later as a red straggler, before becoming a sub-solar-mass black hole.

The unique internal structures of stars harbouring black holes may make it possible, they argue, for asteroseismology to confirm their existence.

Perhaps the Gaia-identified sub-subgiants, and red stragglers, offer a particularly interesting search sample.

165. Runaway stars

RUNAWAY STARS are stars with such high space velocities, generally considered to be above a threshold of $30\,\mathrm{km\,s^{-1}}$, and sufficiently distinct from the overall stellar velocity distribution, that they must have been imparted by a particular formation process. Most are young massive O stars, or Wolf–Rayet stars (W–R stars being the final He-burning phase in the evolution of massive O stars of initial mass $\gtrsim 25 M_\odot$; essay 105).

The discovery of runaway stars followed from the realisation that many O stars (Blaauw, 1961; Gies & Bolton, 1986; Gies, 1987) and Wolf–Rayet stars (Moffat & Isserstedt, 1980; Moffat et al., 1998) lie well outside their likely birth places in open clusters and OB associations, with some clearly 'pointing back' to their likely origin.

They often create spectacular 'bow shocks' as they plough through the interstellar medium, most recently beautifully imaged by HST–ACS and JWST.

THERE ARE TWO main theories for the origin of runaway stars, and both appear to operate: the binary-supernova scenario, BSS, and the dynamical ejection scenario, DES, in which a star is ejected via dynamical interactions in a young, compact cluster.

Dynamical ejection models employ increasingly detailed N-body simulations to study the effects of mass segregation, binary fraction, period distribution, binary mass ratios, and eccentricities (e.g. Oh & Kroupa, 2016).

IN THE binary-supernova scenario, runaway stars result from the evolution of massive binary systems (van den Heuvel, 1973; Tutukov & Yungelson, 1973):

$$\mathrm{O+O} \;\rightarrow\; \mathrm{WR+O} \underset{\mathrm{SN}}{\rightarrow} \mathrm{c+O} \;\rightarrow\; \mathrm{c+WR} \underset{\mathrm{SN}}{\rightarrow} \mathrm{c\,(+)\,c}$$

Here, c denotes the compact companion (neutron star or black hole), left after the supernova explosion of its progenitor. Simply stated, the more massive star evolves fastest, while wind-driven mass-loss (assisted by Roche-lobe overflow in the closest massive binaries) makes O stars evolve into lower-mass (and themselves evolving) Wolf–Rayet stars. At the end of the Wolf–Rayet phase, in both cases, the star explodes as a supernova.

If the first supernova explosion is symmetric, the binary system remains bound, leading to the class of high-mass X-ray binaries. If the first supernova is asymmetric, the binary system may disrupt, depending on the magnitude and direction of the extra kick velocity (de Cuyper, 1982). In either case, with the supernova explosion duration being very short compared to the orbital period, the star receives a recoil velocity, and becomes a runaway, with velocities reaching $200\,\mathrm{km\,s^{-1}}$ for the 'tightest' and most massive pre-supernova binaries.

If the system has not already separated after the first supernova, the system will normally become unbound as a result of the second, producing two high-velocity, single pulsars. In rare cases, the binary can survive this second supernova, producing a binary pulsar (e.g. de Cuyper, 1985; Moffat et al., 1998).

PRE-HIPPARCOS STUDIES of runaway stars were based largely on radial velocities. The broad emission lines resulting from the high outflow velocities means, however, that the determination of radial velocities by classical Doppler techniques is itself greatly restricted.

More than 1000 O–B5 stars were included in the Hipparcos catalogue, including 153 of the 162 proposed runaway candidates known at the time. It also observed 67 Galactic Wolf–Rayet stars, including almost all down to $V = 12\,\mathrm{mag}$, and representing some one third of all Wolf–Rayet then known.

In the study by Moffat et al. (1998), only four of the O stars, and just one of the Wolf–Rayet stars, had reliable parallaxes, and only photometric distances were adopted for the remainder. At $1\,\mathrm{mas\,yr^{-1}}$ accuracy, runaways with transverse velocities of some $100\,\mathrm{km\,s^{-1}}$ could be detected out to distances of about 7 kpc.

Hoogerwerf et al. (2000, 2001) retraced the orbits of 56 runaway O and B stars within 700 pc, linking at least 21 of them to nearby associations and young open clusters, including the classical runaways 53 Ari (Ori OB1), ξ Per (Per OB2), and λ Cep (Cep OB3). Amongst them were a number of runaway 'pairs', including the 70° wide-separation AE Aur+μ Col, which they ascribed to a common origin, coinciding with the λ Ori cluster.

THE PRE-GAIA understanding of runaways was constrained by the limited availability of distances and proper motions (e.g. Moffat et al., 1998; Chevalier & Ilovaisky, 1998; van den Heuvel et al., 2000). Larger numbers, and accurate space motions, were considered essential to distinguish between the BSS/DES formation mechanisms, and their relation to known supernovae, supernova remnants, and high-mass X-ray binaries.

Also debated were suggestions that many massive runaways have high He abundance and rapid rotation, suggesting that at least these were formed through BSS. The 'blue straggler' nature of some was attributed to (hydrogen) mass transfer prior to the supernova explosion. Another discriminant amongst the possible mechanism should be the multiplicity of runaways, where a high multiplicity fraction would point to an efficient dynamical ejection rather than a predominantly binary-supernova origin (e.g. Kobulnicky & Chick, 2022).

These sorts of investigations would be expected to provide further insight into the formation of very massive stars, as well as an improved understanding of stellar kinematics in clusters and associations. They could also elucidate the existence of 'two-step ejections', in which dynamically ejected non-compact binaries can then experience re-acceleration of the surviving star following the subsequent supernova explosion of the primary (Pflamm-Altenburg & Kroupa, 2010).

I should note that the conventional definition of a 'runaway' adopts a $30\,\mathrm{km\,s^{-1}}$ 3-d space velocity threshold (e.g. Gies & Bolton, 1986), while unbound stars with a lower velocity are frequently referred to as 'walkaways'.

'Hypervelocity stars', moving with space velocities of 500–1000 $\mathrm{km\,s^{-1}}$ or more, cannot be explained by either the BSS supernova or DES ejection mechanism, and point to an even more extreme formation scenario. I will not detail these here, but refer to essay 22 (May 2021), and an update to appear as essay 166.

Neither will I say more on the related work on high-mass X-ray binaries, which is also benefitting from the Gaia data (e.g., Hambaryan et al., 2021; van der Meij et al., 2021), but also merits a separate treatment.

INVESTIGATIONS with Gaia DR2 include many individual runaways, together shedding light on the frequency and nature of ejection events in young clusters.

These include: identifying new candidates (Maíz Apellániz et al., 2018); confirming the ejection of VFTS 16 from R136 some 1.5 ± 0.2 Myr ago (Lennon et al., 2018); assigning the runaway J01020100–7122208 to the Milky Way halo rather than the SMC (Massey et al., 2018); confirmation of LP 40–365 as a white dwarf remnant of a peculiar Type Iax supernova (Raddi et al., 2018); association of the Sh 2–296 nebula with three successive ejected runaway stars (Fernandes et al., 2019); identification of 26 runaways originating in the dense Orion Trapezium nebula (McBride & Kounkel, 2019); confirmation that the $150 M_\odot$ VFTS 682 in 30 Dor is a runaway ejected from the R136 cluster (Renzo et al., 2019); other studies of the 30 Dor runaways (Gebrehiwot & Teklehaimanot, 2021); that the K giant HD 137071 was ejected in a massive BSS event 32 Myr ago (Comerón & Figueras, 2020); the nature of two runaways in Vela OB1 (Azatyan et al., 2020); and the confirmation of WR6 (Gvaramadze, 2020).

A SYSTEMATIC search around the Orion Nebula Cluster using Gaia DR2 was made by Farias et al. (2020). Their starting list of 17 000 sources out to 1 kpc resulted in 25 new candidates. A similar search by Schoettler et al. (2020) identified 9 runaways and 24 walkaways. The former is consistent with their N-body simulations over an age of 1.3–2.4 Myr, confirming the current cluster age estimate but based only on runaways. Detailed studies of the AE Aur+μ Col system using EDR3 were reported by Bobylev & Bajkova (2021).

Dorigo Jones et al. (2020) used DR2-based transverse velocities for 304 OB stars in the Small Magellanic Cloud to conclude that dynamical ejections dominate over supernova ejections by a factor 2–3, with two-step runaways likely dominating the BSS runaway population.

Focussing on the nature and origin of blue stars at high Galactic latitudes, Raddi et al. (2021) identified 12 B-type runaway candidates, of $2-6 M_\odot$, out to 10 kpc, and with flight times compatible with their evolutionary ages. Three with high velocities $>450\,\mathrm{km\,s^{-1}}$ challenge the BSS or DES ejection scenarios, suggesting mechanisms more aligned to those for hypervelocity stars.

Bhat et al. (2022) described a method to determine whether two or more objects come from the same location, tested against Orion runaways, and suggested that AE Aur+μ Col might not have originated together.

TWO LARGER-SCALE searches have been reported. Using EDR3, Kounkel et al. (2022) compiled a catalogue of 3354 young candidates within 500 pc that appear to have been ejected from their parent associations. Among the possible candidates they identified various pairs, with similar traceback times, that appear to have interacted in the ejection process.

Starting from the Galactic O-Star Catalog (GOSC) of Maíz Apellániz et al. (2013), and the Be Star Spectra (BeSS) Database of Neiner et al. (2011), Carretero-Castrillo et al. (2023) used Gaia DR3 astrometry to find 106 O-type runaways (42 new), and 69 Be-type runaways (47 new), with the percentage of runaways being 25% for O-type stars, and 5% for Be-type stars.

Their sample includes seven X-ray binaries and one gamma-ray binary. The higher percentages and higher velocities found for O-type compared to Be-type runaways again suggests that the dynamical ejection scenario is more favoured.

166. Hypervelocity stars: part 2

JUST OVER two years has elapsed since my essay 22 on hypervelocity stars in mid-2021. But with more than 30 Gaia papers since then, an update is timely. Let me start with the context, and the status at that time.

Classical 'runaway' stars have been known for more than 60 years (see essay 165). Their space velocities, of 30–100 km s^{-1}, are sufficiently distinct from the overall stellar velocity distribution that they must have been imparted by some particular formation mechanism(s).

Two such processes appear to operate: a dynamical ejection scenario (DES), in which a star is ejected via dynamical interactions in a young compact cluster, and the binary-supernova scenario (BSS), where velocities are imparted in a supernova-driven binary disruption.

Even more extreme velocities in a related but more extreme environment were predicted by Hills (1988). He showed that a close encounter between a tightly bound binary, and a $10^6 M_\odot$ black hole, causes one component to become bound to the black hole, and the other to be ejected at up to 4000 km s^{-1}, or 1% of the speed of light!

THE FIRST hypervelocity star, SDSS J090745+0245077 (HVS1), was discovered, from SDSS, to have a radial velocity of ~800 km s^{-1} (Brown et al., 2005). At a distance of ~55 kpc, 30 kpc above the disk and 60 kpc from the Galactic centre, its space motion was consistent with it having been ejected from the Galactic centre. And models demonstrated consistency with its generation, via the Hills mechanism, from the $4 \times 10^6 M_\odot$ black hole (Sgr A*) at the centre of our Galaxy (Brown, 2015).

In essay 22 I described the radial velocity surveys, at the MMT and LAMOST, set up to discover more of these stars, recalled estimates of their production rate, and mentioned their potential use as tracers of the Galaxy's potential as they escape its confines over some 300 Myr.

SINCE DR2 in April 2018, Gaia has been providing proper motions and distances of unprecedented accuracies, helping to confirm candidates (and refute others, e.g. Quispe-Huaynasi et al., 2022), and pinpointing their origins, while searching the entire sky for others.

A number of the other 30–40 early Gaia discoveries were also provisionally considered to be consistent with a Galactic-centre origin, such as HVS 22 with v ~1500 km s^{-1}. But at least one, HVS3, appeared to come from the Large Magellanic Cloud, suggesting that it harbours a massive black hole of its own.

Others appeared to come from elsewhere in the Galaxy disk (e.g. Li et al., 2018; Du et al., 2019; Marchetti et al., 2019). Amongst these are at least four white dwarfs (Shen et al., 2018; Ruffini & Casey, 2019), conjectured to be companions to primary white dwarfs that exploded as Type Ia supernovae, and suggesting that some may originate from extreme binary supernova ejections.

IN BRINGING the topic up-to-date, let me first mention the other (non-Gaia) discoveries, all since making use of Gaia for their further interpretation.

Following HVS1 came a number of other B-type hypervelocity halo stars: HVS2 (Hirsch et al., 2005), HVS3 (Edelmann et al., 2005), HVS4–HVS24 (Brown et al., 2006; Brown et al., 2007; Brown et al., 2014), HD 271791 (Heber et al., 2008), HIP 60350 (Irrgang et al., 2010), and LAMOST–HVS1 through LAMOST–HVS4 (Zheng et al., 2014; Huang et al., 2017; Li et al., 2018).

The AAT–2dF S5 project (Southern Stellar Stream Spectroscopic Survey) found its first, S5–HVS1 (Koposov et al., 2020), a $G = 16$ star with radial velocity 1020 km s^{-1} (and with a Gaia DR2 proper motion $\mu_\alpha = 35.328 \pm 0.084$, $\mu_\delta = 0.587 \pm 0.125$ mas/yr). Together this yields a 3-d Galactocentric velocity 1755 ± 50 km s^{-1} and, according to the authors, the *'only hypervelocity star confidently associated with the Galactic Centre'* at that time.

With the availability of Gaia DR2, Boubert et al. (2018) revisited ~500 (mainly late-type) high-velocity stars identified in the literature over the years, and found that most do not fall into the category of true hypervelocity stars. Li et al. (2021) emphasised the distinction between hypervelocity stars, 'hyper'-runaways (runaways that are unbound to the Galaxy), runaways, and fast halo stars (often simply reflecting the different rotation states of the disk and halo).

ADDITIONAL CONSTRAINTS on formation models are the 40 or so 'S stars' in close orbit around Sgr A* (e.g. Gillessen et al., 2017), which may represent the remaining captured stars (Generozov & Madigan, 2020).

Meanwhile, progress in characterising hypervelocity stars originating from the Galactic centre has led to an improved understanding of the conditions in the inner few parsecs of the Milky Way (Evans et al., 2022b; Evans et al., 2022c; Evans et al., 2023; Generozov et al., 2022; Generozov & Perets, 2022; Chu et al., 2023).

For example, from his models, Generozov (2020) suggested that S5–HVS1 is part of a larger stream, and predicted its spatial and velocity distributions.

AS IT BECAME clearer that many hypervelocity stars could not have originated from the Galactic centre (e.g. Irrgang et al., 2021), other similarly highly disruptive ejection mechanisms have been put forward.

In addition to the original hypothesis of tidal disruption of binary stars by a single massive black hole in the Galactic centre (Hills, 1988; Yu & Tremaine, 2003; Bromley et al., 2006), these include single star encounters with a binary black hole (Yu & Tremaine, 2003; Sesana et al., 2006; Sesana et al., 2007); single star encounters with a cluster of stellar mass black holes around a massive black hole (O'Leary & Loeb, 2008); or the interaction between a globular cluster with a single or binary black hole (Fragione & Capuzzo-Dolcetta, 2016).

Additionally, three-body processes may dominate the high-velocity ejection from globular clusters, perhaps explaining some of the known hypervelocity stars (Weatherford et al., 2023; Cabrera & Rodriguez, 2023).

In the very different 'dynamically driven double-degenerate double-detonation' (D6) mechanism (for details and history, see Shen et al., 2018, §1), Type Ia supernovae may occur during unstable mass transfer between two white dwarfs in a binary. If the donor survives the explosion, it can be released as a hypervelocity star, with the velocity of its 1000–3000 km s^{-1} pre-supernova orbital speed. Further studies continue (e.g. Evans et al., 2020; Bauer et al., 2021; Shields et al., 2022; Braudo & Soker, 2024; Igoshev et al., 2023; Liu et al., 2023).

I SHOULD STRESS that the kinematic interpretation of most of these hypervelocity star candidates, including their confirmation (or otherwise), has used the Gaia distances and proper motions to determine the transverse component of the star's space velocity.

Of Gaia searches, the earliest were based on DR2 (Du et al., 2019; Marchetti et al., 2019; de la Fuente Marcos & de la Fuente Marcos, 2019; Li et al., 2021). Subsequent searches used EDR3 (Marchetti, 2021; Irrgang et al., 2021; Li et al., 2023b), and most recently DR3 (Marchetti et al., 2022; Liao et al., 2023; El-Badry et al., 2023c; Igoshev et al., 2023; Parthasarathy, 2023).

THE GAIA STUDIES are a mix of previously known candidates, unbound stars not originating from the Galactic centre, and others of a possible extragalactic origin. El-Badry et al. (2023c), for example, found six new runaways, including four with radial velocities >1000 km s^{-1} and total space velocities >1300 km s^{-1}, most likely the survivors of the 'D6-driven' scenario.

AS FOR THOSE originating from *beyond* our Galaxy, Gualandris & Portegies Zwart (2007) inferred that HE 0437–5439 was directed from the LMC, although this was later refuted by Brown et al. (2010). But stronger evidence that HVS3 originated in the LMC (Erkal et al., 2019) has prompted further work in this direction.

Evans et al. (2021) modelled stars ejected from the Milky Way and the LMC via the Hills mechanism, propagated in a combined potential in which the LMC is on its first infall. They identified simulated hypervelocity stars which should be recognisable in the stellar halo.

Lin et al. (2023) used DR3 to examine the space motions of 3119 massive O-B2-type stars in the LMC, and identified 98 as possibly escaping. They concluded that dynamical ejection is the dominant mechanism for the hypervelocity runaway stars in the LMC.

Montanari et al. (2019) used DR2 to identify the orbital intersection of two known stars, 20–40 Myr ago, with the Sagittarius dwarf spheroidal galaxy.

Li et al. (2022b) used EDR3 astrometry, radial velocities, and [Fe/H], to find 60 high-velocity stars probably from the Sagittarius dwarf galaxy. Two are hypervelocity stars, possibly produced by the Hills mechanism.

Gülzow et al. (2024) modelled hypervelocity stars being generated in Andromeda. They estimated that their number, in *our* Galaxy, is between 12–3910.

Meanwhile, studies have continued into their use as probes of the detailed gravitational potential of the dark matter halo (e.g. Koposov et al., 2020, §6; Gallo et al., 2022), and in probing MOND (Chakrabarty et al., 2022).

FOLLOW-UP observations of hypervelocity stars include spectroscopy of 14 candidates from the MMT, where the results indicated that all but one are main sequence B stars of mass $2.5-5.M_\odot$, with ages 90–400 Myr and distances 30–100 kpc (Irrgang et al., 2018).

Martinez et al. (2022) modelled the hydrodynamical interaction between the stellar wind and surrounding material, demonstrating that this leads to strong particle acceleration and non-thermal radiation. They account for perhaps 0.1% of the sub-TeV Galactic cosmic ray flux.

I'M NOT AWARE of a recent synthesis, so I can't give the number of s*ecure* hypervelocity stars: perhaps 50? Nor are we clear how many originate in the Galactic centre, or via the D6 mechanism, or from other galaxies. But Gaia is making a substantial contribution to their study!

167. Carbon stars

CARBON STARS are characterised by atmospheres with more carbon than oxygen. The 'classical' carbon stars are luminous red giants, on the asymptotic giant branch. Their plentiful atmospheric carbon gives them a deep red appearance, and they were first recognised spectroscopically by Angelo Secchi in the 1860s.

Intrinsic carbon stars arise as a result of carbon being 'dredged up' from their core at certain evolutionary phases. But some (extrinsic) dwarfs and supergiants may also show C/O>1 as a result of mass transfer from a carbon-enriched binary companion.

The abundant C combines with any O in the cooler upper layers to form CO. Residual carbon forms other C-based compounds, and results in a 'sooty' atmosphere and a vivid red colour. Spectroscopically, they include the C_2 (Swan) absorption bands at 438.3 and 473.7 nm, and the 421.6 nm band of CN. They are typically long-period variables, and include the C-rich Mira variables.

Apart from their high carbon content, their spectral properties broadly parallel main sequence K–M stars. They are sub-divided into the 'warmer' (and rarer) R stars (the carbon analogue of K stars), and the 'cooler' N stars (the carbon analogue of M stars), and others.

TO EXPLAIN HOW the excess C abundance arises, I will greatly simplify the complex post-main sequence evolution of intermediate mass stars (see figures).

Following core H exhaustion, the core contracts and heats, causing the outer layers to expand and cool, and the star develops into a red giant. Once He-core burning starts, the cooling halts, and the star moves 'down and left' on the *horizontal branch*. On completion of core He-burning, the star moves 'up and right' again – cooling, expanding and increasing in luminosity along the *asymptotic giant branch*, or AGB.

During a series of short-duration thermal pulses, which result from a complex interplay between cycles of H- and He-shell burning, deep convection 'dredges up' core material. It is the phenomenologically-defined 'third dredge-up' which brings He, C, and the various s-process elements to the surface, increasing the C/O ratio, and turning the star into a carbon star.

THEORY SUGGESTS that essentially all stars with initial masses $1.5-4M_\odot$ should go through the carbon star phase, lasting around 300 000 yr, before ending as a white dwarf. Their very high luminosity makes them important for integrated light studies of galaxies, and they have also been considered as potential standard distance indicators, along the lines of the Cepheids, TRGB, or JAGB (Richer, 1981; Battinelli & Demers, 2005; Freedman & Madore, 2020; Ripoche et al., 2020).

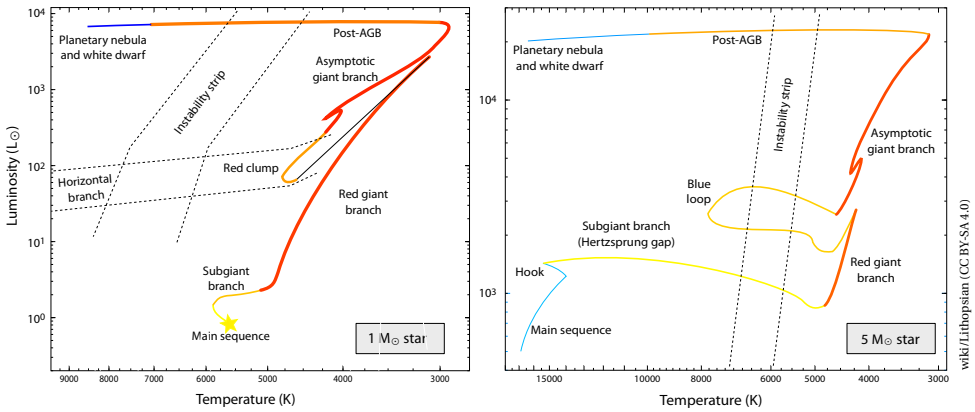
Their chemical compositions are typically nearly solar C/H, N/H, and $^{12}C/^{13}C$ ratios, indicating that much of the C and N in our Galaxy came from mass-losing carbon stars. Their mass-loss rates, up to several times $10^{-5}M_\odot\,\mathrm{yr}^{-1}$, contribute around half of the total mass returned to the interstellar medium (Wallerstein & Knapp, 1998; Straniero et al., 2023). Carbon stars are also among the main sites of heavy-element production ($A \geq 90$), through the (slow neutron capture) s-process.

TODAY, there are thousands of known carbon stars: the Third Edition of the Catalog of Galactic Carbon Stars by Alksnis et al. (2001) listed 6891. A recent catalogue of AGB stars in our Galaxy by Suh (2021) lists 11 209 O-rich and 7172 C-rich. Other compilations include those by Chen & Yang (2012) and Ji et al. (2016).

Some of the brighter and more well-known carbon stars include (wiki links provided): V Aql, Y CVn, CW Leo, R Lep, SS Vir, and TX Psc.

No individual distances were known before Hipparcos and Gaia. But even if a distance is known, and therefore an absolute magnitude, assigning an individual mass is usually impossible, and it is parameters such as mass-loss rate and age that are some of the weakest links in current model predictions (Abia et al., 2020).

SOME 320 carbon stars were identified in the Hipparcos catalogue by Knapp et al. (2001). Of the R stars (the least understood, and whose knowledge is, in part, limited by distance uncertainties) just 17 had $\sigma_\pi/\pi < 0.5$. Knapik et al. (1998) estimated a local space density of $40-70\,\mathrm{kpc}^{-3}$, with no clear spatial correlation with the local spiral arm structure or interstellar extinction.



In their detailed review of the present state of knowledge, both theory and observation, Straniero et al. (2023) noted: *'More than 40 years after the pioneering paper by Iben (1981), the efficiency of the third dredge-up and the chemical yields from AGB stars are still burdened by heavy uncertainties and disagreements among different authors, mainly due to the lack of a robust theory of convection and mass loss.'*

For example, the number of thermal pulses required to yield C/O>1 is a complex function of initial mass, metallicity, dredge-up efficiency, mixing-length and convective overshoot parameters.

It is these sorts of uncertainties, poorly constrained by calibrated luminosities given their large distances, that propagate through to uncertainties in chemical enrichment models, and thus to the chemo-dynamical history of the Galaxy (e.g. Gustafsson, 2022).

Turning to Gaia, the 'Golden Sample' of Creevey et al. (2023; see also essay 90) used all relevant Gaia data (astrometry, photometry, and the mean BP/RP and RVS spectra) to define various well-defined subsets, including a sample of 15 740 carbon stars. It reveals a strong concentration towards the Galactic plane, with others in the LMC/SMC, and the Sagittarius stream.

Various studies have been made using Gaia data. Kostandyan (2020) used DR2 to study 127 carbon stars ($G = 9-18$ mag) discovered from the First Byurakan Survey, of which 56 are N type. All are AGB stars in the Galactic halo, all within 14 kpc, all within 8 kpc of the mid-plane, and with radial velocities available for 75.

Pal & Worthey (2021) tackled the uncertain age dilemma by selecting carbon stars associated with clusters and moving groups of known age, and investigated the occurrence of Galactic carbon stars as a function of progenitor mass using the Gaia DR2 data. Amongst their conclusions, the C-star frequency agrees with that observed in the Magellanic Clouds down to $1.67 M_{\odot}$, as does the frequency of carbon stars in M31.

Abia et al. (2020) also used DR2 to determine the luminosity function and kinematics of 210 carbon stars in the solar neighbourhood with $\sigma_{\pi}/\pi < 0.2$. Amongst their conclusions were that most of the N type stars belong to the thin disk population, while a significant fraction of the R type are compatible with the thick disk.

Abia et al. (2022) extended these studies using EDR3, and a larger sample of 974 C stars (491 N type, 22 SC, 83 J, 234 R, and 276 CH). Amongst their conclusions, they found that N- and SC-type stars share a similar luminosity function, while the R type have luminosities throughout the red giant branch, favouring an external origin for their C enhancement. They identified 2660 new C stars through their 2MASS photometry and Gaia astrometry.

The AGB stars, and in particular the J-region (JAGB) method (Freedman & Madore, 2020, §2.1), have considerable potential as an extragalactic standard candle, capable of calibrating the absolute magnitudes of local Type Ia supernovae, and thus an independent determination of the Hubble constant. The topic is already a substantial field of study (e.g. Lee et al., 2021a; Lee, 2023), which I will not detail further here.

A study by Nanni (2019) characterised the properties of amorphous carbon dust condensed around carbon stars by comparing the LMC carbon stars (from 2MASS and Gaia DR2) with synthetic photometry from dust growth models. Only very specific combinations of optical data and grain size could simultaneously match the infrared photometry and Gaia astrometry.

Roulston et al. (2021) focussed on dwarf carbon (dC) stars, main-sequence stars enriched by mass transfer from a carbon-rich (AGB) companion, itself since evolved to a white dwarf. Their investigations focussed on their orbital properties, derived from the photometric light curves from the Zwicky Transient Facility for a sample of 944 dC stars. They used the Gaia EDR3 parallaxes and proper motions to verify that any periodic (orbit) candidate was indeed a dwarf carbon star.

168. S stars in the Wesenheit diagram

ONLY WITHIN the past century did astronomers establish that stars were powered by nucleosynthesis, and that the chemical elements (except for the very lightest: H, He, and Li) were created inside them.

The theory underpinning stellar nucleosynthesis replaced some earlier ideas that all elements were created in the Big Bang. Major conceptual advances were made in the 1940s and 1950s (Hoyle, 1946; Hoyle, 1954; Hoyle et al., 1956; Fowler et al., 1955; Burbidge et al., 1957).

Relevant to my topic of S stars is the early abstract of Fowler et al. (1955): *'It is supposed that the synthesis of the heavy elements will take place in stars which have reached the cool giant stage in their evolutionary path. The M as well as the S stars may be important in this connection.'* They estimated that… *'synthesis of the heavy elements in the S-star stage alone could account for a considerable fraction of the heavy elements'.*

STELLAR NUCLEOSYNTHESIS broadly explains why certain elements are more abundant than others, and predicts how and why they evolve over time.

The details of stellar evolution and its fusion products are not particularly simple or straightforward to interpret anywhere in the Hertzsprung–Russell diagram, although the reactions and internal structures of main-sequence stars are reasonably well understood. Things are more complicated for the more advanced evolutionary stages, and in particular for stars on the 'asymptotic giant branch', or AGB.

The interiors of stars on the asymptotic giant branch are characterised by a largely inert core of carbon and oxygen, an inner shell where helium is fusing to form carbon, an outer shell where H-to-He burning still operates, and an outer envelope, often of composition similar to main-sequence stars (Lattanzio & Forestini, 1999; Siess, 2006; Höfner & Olofsson, 2018; Karakas, 2019; Kobayashi et al., 2020).

Indeed, AGB stars are very important in the production of the heaviest elements. But their understanding is complicated by their dramatic and rapid evolution, and their sensitivity to the details of various crucial processes including deep mixing and strong mass loss.

AS I DESCRIBED in essay 167, a star arrives on the AGB on completion of core He-burning, where it cools, expands, and its luminosity increases. Following a series of short-duration thermal pulses (resulting from cycles of H- and He-shell burning), deep convection (the 'third dredge-up') brings He, C, and various s-process products from the core to the surface, where they are seen in the spectra and can be expelled through the stellar wind. The third dredge-up increases the C/O ratio and, in some cases, turns the star into a carbon star.

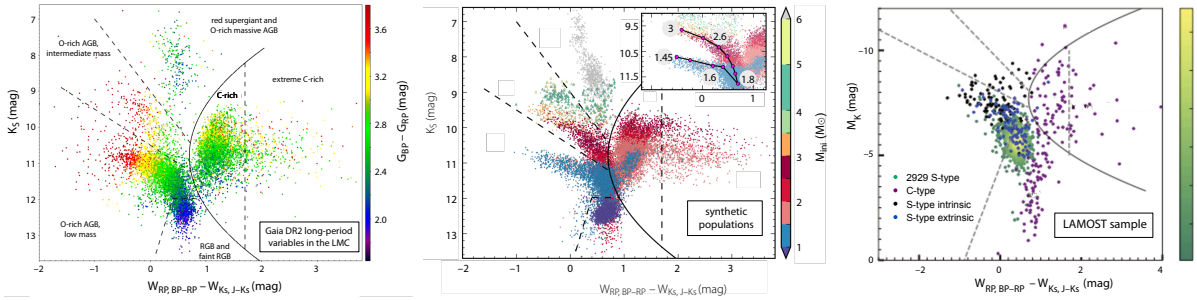
[The first dredge-up occurs when a main-sequence star enters the red-giant branch, resulting in an atmosphere displaying the products of H-fusion, including lowered ^{12}C/^{13}C and C/N ratios, and lower abundances of Li and Be. The second dredge-up occurs only in intermediate-mass stars, $4-8M_{\odot}$. It reveals the products of the CNO cycle, increasing the surface abundance of ^{4}He and ^{14}N, and decreasing that of ^{12}C and ^{16}O.]

THE S STARS are cool luminous (AGB) giants lying between the O-rich M-type giants and the C-rich carbon stars. Short-lived (10^5-10^6 yr), they make up only 1% of all M stars, but their large-amplitude long-period variability facilitates their discovery. The third dredge-up also brings s-process elements to the surface, with the result that S stars show unusual absorption lines, including the defining lines of ZrO (Merrill, 1922).

In *intrinsic* S stars, the s-process elements arise from this convective dredge-up (rather than as mass-transfer in a binary system), where they are accompanied by the presence of technetium, another of the s-process elements. With its longest isotopic half-life of only 4.2 Myr, the presence of Tc indicates its recent creation.

To summarise, the importance of the class is that the s-process (slow-neutron capture) creates around half the atomic nuclei heavier than iron. A range of elements and isotopes is produced, with relative abundances depending on the source and flux of the neutrons (Suess & Urey, 1956; Clayton et al., 1961; Kobayashi et al., 2020).

The General Catalogue of Galactic S-Stars (Stephenson, 1984) listed 741 stars, with one of the brightest being the 370-d, 7–14 mag Mira variable R Gem.



K_S versus W(Gaia) – W(2MASS) for (a) LMC, (b) evolutionary models (both Lebzelter et al., 2018); (c) LAMOST (Chen et al., 2023a)

MAJOR ADVANCES in understanding AGB stars have followed from large-scale multi-epoch multi-colour surveys, notably MACHO (e.g. Wood, 2003) and OGLE (e.g. Soszyński et al., 2009). Today, with its astrometry and 3-band photometry (G, BP and RP), Gaia is providing new means of identifying the various populations on the asymptotic giant branch. As we shall see, it can clearly distinguish between the O-rich and C-rich stars, as well as delineating the region populated by S stars.

Let me first explain the 'Wesenheit function', W. This is some combination of a star's magnitude and colour, chosen to be (largely) unaffected by distance *and* reddening. The term comes, I believe, from the German for 'entity' or 'true nature'. It was first used in the context of the Cepheid period–luminosity relation in the 1970s.

As an early proponent, Madore (1976) adopted $W = V - R(B - V)$, where R is the total-to-selective absorption ratio. The demonstrated linearity of W implies the linearity of the period–luminosity and period–intrinsic colour relations. The properties and choices for W are described further in, e.g., Ngeow & Kanbur (2005, §2).

USING THE Gaia DR2-based catalogue of 151 761 long-period variables, which itself doubled the number of known long-period variables (Mowlavi et al., 2018), Lebzelter et al. (2018) showed that a diagram of W(Gaia)–W(2MASS) versus K_S is a powerful tool to distinguish between different AGB classes.

Here, $W(\text{Gaia}) \equiv W_{G_{RP},\,BP-RP}$ is a Wesenheit function based on the Gaia magnitudes G, G_{BP} and G_{RP}, and $W(\text{2MASS}) \equiv W_{K_S,\,J-K_S}$ is similarly based on the 2MASS infrared magnitudes J and K_S. Specifically, they adopted $W(\text{2MASS}) = K_S - 0.686\,(J - K_S)$, and $W(\text{Gaia}) = G_{RP} - 1.3\,(G_{BP} - G_{RP})$.

As they argued, $G_{RP} - G_{BP}$ is strongly affected by the temperature sensitivity of molecules dominating these wavelength ranges, which themselves depend on the C/O ratio. However, a degeneracy exists between interstellar reddening, circumstellar reddening, chemistry, and temperature differences at a given luminosity. By using Wesenheit functions, they could largely eliminate the interstellar reddening component (and much of the circumstellar component) from this degeneracy.

A diagram for their 11 022 Large Magellanic Cloud candidates (selected via their Gaia parallax and proper motion) is shown in the first figure. As they stress, this is not a classical colour–magnitude diagram, since the bluest objects are at the centre, with redder ones to both sides. They identified six regions (C rich, extreme C rich, O rich in three different mass ranges, and red-giant branch), and they explain how these segregations arise as a result of the different molecular species and mass ranges.

During their evolution on the AGB, stars of low and intermediate mass show strong variations of temperature, composition and luminosity on various timescales, which have been studied in detail using evolutionary models in view of their importance in stellar and Galactic evolution (e.g., Karakas & Lattanzio, 2014; Marigo et al., 2017). Their simulated populations, obtained with the TRILEGAL evolutionary code, are shown in the middle figure, well replicating the observational diagram.

THE SUBSET of S stars has been further studied using the Gaia data, providing further observational evidence for the third dredge-up, and for the detailed evolutionary status of the Tc-rich stars, based on DR2 (Shetye et al., 2018; Shetye et al., 2019; Shetye et al., 2020; Abia et al., 2020), and EDR3 (Shetye et al., 2021; Abia et al., 2022). A recent review of the field, including the most recent Gaia results, is given by Van Eck et al. (2022).

WHERE DO THE subset of S stars lie in such a Wesenheit diagram? Although I am not aware of the systematic placing of all known S stars in such a diagram on the basis of the Gaia astrometry and photometry, two specific studies provide an encouraging picture.

Figure A2 of Shetye et al. (2021) shows their 23 (spectroscopic) S stars in the same W(Gaia)–W(2MASS) versus K_S diagram, and most fall, as expected, in the low-mass O-rich region.

On a larger scale, Chen et al. (2023a) found 2939 new S stars from LAMOST DR 10, and placed them in the same 'Lebzelter' diagram using Gaia DR3 astrometry and photometry. Their result, colour-coded according to whether they are inferred to be intrinsic or extrinsic, is shown top right. Again, most fall in the O-rich region.

169. A billion radial velocities

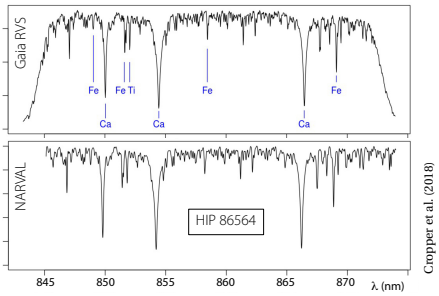
I HAVE DESCRIBED the acquisition and use of Gaia's radial velocities in several previous essays: why they are important (essay 8), the resolution and wavelength range chosen for the radial velocity spectrometer (85), their acquisition on board (86), and a summary of their content in Data Release 3 (87).

Let me summarise a few key points. The star's radial velocity (i.e. *along* the line-of-sight) is required, together with the projected motion on the sky given by its proper motion, to determine the star's full 3-d space motion. This is important, and often crucial, for many kinematic and dynamical studies (e.g. Wilkinson et al., 2005). They also contribute, deeply, to orbital studies of binaries.

However, astrometry *per se* cannot determine this radial component of motion. Accordingly, a dedicated radial velocity spectrometer (RVS) was included in Gaia's instrument focal plane. RVS is an integral-field spectrograph, covering the wavelength range 845–872 nm at a resolution $R = \lambda/\Delta\lambda \sim 11\,700$. This combination was carefully chosen to cover the important Ca II infrared triplet, the Paschen series of hydrogen in early-type stars, and certain diffuse interstellar bands (essay 85).

DUE TO PRACTICAL constraints, RVS obtains spectra only for sources $G_{\rm RVS} \lesssim 16$ mag, and with 43% fewer focal plane transits than the astrometric and photometric fields, typically ~80 over the 10-year mission.

As an example, the figure shows the $V = 6.7$ mag K5 star HIP 86564 from a single 4.4 sec Gaia exposure (top), and with the NARVAL spectrograph at the Observatoire du Pic du Midi (at the same spectral resolution, bottom).



THE PROCESSING of the RVS data is carried out, on the ground, within an extensive and dedicated processing 'pipeline', described for Gaia DR2 by Sartoretti et al. (2018), and for Gaia DR3 by Katz et al. (2023).

While Gaia is today into its tenth year of observations, only the first 34 months (July 2014–May 2017) of data have been processed and released by the Gaia Data Processing and Analysis Consortium (DPAC), as Gaia Data Release 3 on 13 June 2022.

Along with 1.8 billion sources with astrometry, DR3 provides radial velocities for 33 million sources down to $G_{\rm RVS} \sim 14$ mag. When DR4 is released in 2025, based on 66 months of data, RVS results for the 100 million sources down to $G_{\rm RVS} \sim 16$ mag should be available.

LET ME put these numbers in context. When the Hipparcos catalogue of 120 000 stars was released in 1997, radial velocities were known for just 20 000. Surveys by Coravel, amongst others (often focused on exoplanets) later measured several thousand more.

Subsequently, and on a much larger scale, and in the northern hemisphere between 2005–2010, the Sloan Digital Sky Survey's SEGUE extension obtained spectra for 240 000 stars, with typical radial velocity accuracies of 10 km s^{-1} (Yanny et al., 2009). SEGUE–2 (2008–2009) observed a further 120 000 (Rockosi et al., 2022).

Complementing SEGUE in the south, RAVE (RAdial Velocity Experiment) was a multi-fiber spectroscopic survey using the 1.2-metre AAO–UK Schmidt Telescope. Conducted between 2004–2013, and partly motivated by the prospects of Gaia, RAVE acquired some 574 000 spectra for around 483 000 stars (Kunder et al., 2017).

The LAMOST–II medium-resolution ($R \sim 7500$) spectral survey measured 1 597 675 spectra for 281 515 stars (Wang et al., 2019), achieving radial velocity accuracies of around 1 km s^{-1}. The (AAO–HERMES) GALAH+ survey includes 584 015 dwarfs and giants in the magnitude range $G = 11 - 14$ (Zwitter et al., 2021).

And the merged and homogeneous 'Survey of Surveys' contains almost 11 million stars with radial velocity precision in the range 0.05–1.50 km s^{-1}, of which half are exclusively from Gaia (Tsantaki et al., 2022).

THIS BACKGROUND sets the scene for some work recently reported by Verberne et al. (2024). They turned their attention to the possibility of deriving radial velocities from the low-resolution BP/RP spectra that Gaia measures for *all* sources. Unlike the RVS which (aside from data rate limitations) is photon-starved at $G \gtrsim 16$ mag, the low-resolution BP/RP spectra are acquired for all sources crossing the main focal plane.

Currently, DR3 (June 2022) provides the full astrometric solution for 1.5 billion sources, while the low-resolution spectra have been published for 'only' some 220 million. The BP/RP spectra for Gaia's full 2 billion or more sources will be made available with DR4 in 2025.

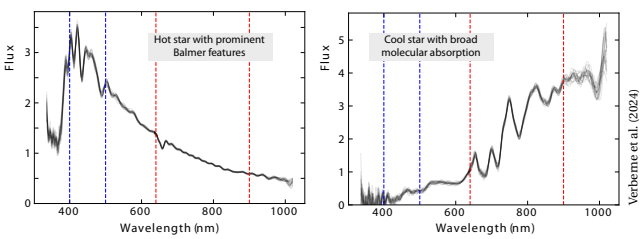
THE LOW-RESOLUTION BP/RP spectra provide astrophysical information on each astrometric source. The design, including wavelength range and resolution, was optimised to provide information needed to determine the astrophysical parameters of each source, such as $T_{\rm eff}$, log(g), [M/H], and extinction, across a wide range of spectral type and luminosity class. Within the DPAC Consortium (CU8), the 'astrophysical parameters inference system' pipeline (Apsis, Bailer-Jones et al., 2013; Recio-Blanco et al., 2023) runs 13 modules, using different combinations of data and models, to produce astrophysical parameters for stars, galaxies and quasars.

Amongst the modules identifying and classifying the extragalactic sources, the 'quasar classifier' module (QSOC) estimates the quasar redshifts using a χ^2 approach, in which the BP and RP spectra are compared to a composite quasar spectrum over trial redshifts in the range $z = 0-6$ (Bailer-Jones et al., 2023; §2.2). A similar approach estimates the redshifts of galaxies using the 'unresolved galaxy classifier' module (UGC). An animation illustrating this procedure is shown here.

As a result, 6.4 million quasars and 1.4 million galaxies have redshifts given in DR3 (Delchambre et al., 2023).

A SIMILAR APPROACH has been adopted for determining stellar radial velocities by Verberne et al. (2024). The principle is straightforward, although of course the details are not: the radial velocity (and its estimated error) is obtained by fitting the BP/RP spectra to a wide range of models based on a grid of synthetic spectra.

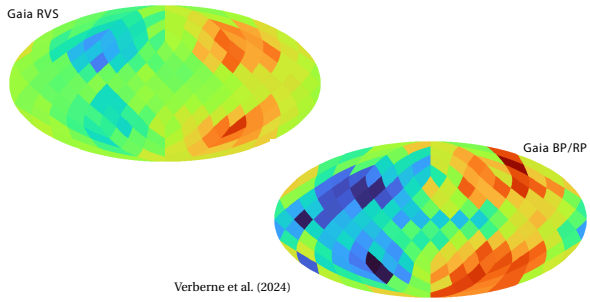
The figure shows two example Gaia source spectra, with the blue and red dashed lines indicating the BP and RP spectral ranges used in the fitting procedure.



THEIR MOST reliable subset comprises 6.4 million sources with uncertainties <300 km s^{-1}, around one quarter of which have no radial velocity yet available from RVS. They also constructed an extended catalogue with all 125 million sources for which they could obtain a valid radial velocity estimate.

The figure below shows: (top) the median radial velocity for low-metallicity stars, [Fe/H] < –1, from the BP/RP spectra as a function of sky position in Galactic coordinates, and (bottom) the median radial velocity from RVS. Colour coding ranges from –200 km s^{-1} (dark blue), through 0 (green), to +200 km s^{-1} (dark red). The dipole component attributable to the solar motion is evident in both maps.

The typical uncertainties are, of course, far higher than those obtained through precision spectroscopy. Their median uncertainty on the calibrated BP/RP radial velocity measurements is ~770 km s^{-1}, but extends to below 100 km s^{-1} (their Fig. 4). Sources with the best accuracy tend to be either very blue (BP–RP $\lesssim 0.7$) or very red (BP–RP $\gtrsim 2$). The scientific value of such imprecise values remains to be seen, but might include searches for black hole binaries or hypervelocity stars.



LET ME summarise. Gaia DR3 has provided radial velocities for 33 million sources to $G_{\rm RVS} \sim 14$, with the 66-month based DR4 expected to deliver around 100 million to $G_{\rm RVS} \sim 16$ mag in 2025. Very coarse measurements from BP/RP are now also available for 125 million.

A VERY DIFFERENT approach to estimating the missing radial velocities is to infer them from the 5 other astrometric parameters, as first implemented using neural networks by Dropulic et al. (2021). Their 6 million source training set was later used to infer radial velocities for a further 92 million, which they applied to stars in the Enceladus merger using EDR3 (Dropulic et al., 2023).

Naik & Widmark (2022) used a Bayesian neural network approach to predict the values for 16 million stars in EDR3. These predictions were tested by Naik & Widmark (2024), and DR3 astrometry in turn used to predict the missing values for 185 million stars to $G < 17.5$. They estimate that the predictions are reliable, to 25–30 km s^{-1}, for stars within 7 kpc of the Sun.

170. Metrication in the UK

In essay 96, 'Is the Earth flat', I discussed a topic that may appear a little absurd and scientifically irrelevant, but which (I think) demonstrates that leading-edge science is not necessarily bringing the rest of society along with it. This may be harmless enough when it centres around whether the Earth is round or flat, but assumes greater importance when it concerns the societal consequences of immunisation or climate change.

Here I will make an aside from my usual scientific insights with Gaia to give some thoughts on the question of metrication… albeit central to Gaia's success!

As a scientist, and one who lived and worked in continental Europe for 30 years, the question is as 'black-and-white' as the Earth's shape – as it is for all countries in Europe except the United Kingdom, and for most other major economies across the world except the United States. Indeed, almost every other country uses, or has made the transition to, metric units.

For the UK, metrication addresses the move from its historical use of yards, feet and inches to metres; from stones, pounds and ounces to kilogrammes; from fluid ounces, pints and gallons to litres; from degrees Fahrenheit to Celsius; and discarding archaic units such as horsepower (for the power of engines or motors), or British Thermal Units (BTU, the heat required to raise one pound of water by one degree Fahrenheit).

Try this: divide 5 stone 7 lb 9 oz into four equal parts. Or add 5 yds 2 ft 3 inches to 2 yds 2 ft 11 inches. To design a supersonic jet, or a high speed train, or a meteorological satellite, is unimaginable in such a system. To cling to them is like holding on to farthings and florins, perches and chains, quarts and gills.

Examples of the difficulties and absurdities of the imperial system are endless. So is the potential for errors, sometimes disastrous. So why is it perpetuated?

Why does the UK continue to do so much of its business in imperial units? Why is the issue of metrication so contentious? Is it a case of maintaining our sovereignty? Or the fear of transitioning from a system that has served us since the Middle Ages? Or simply xenophobia?

The history of the move to decimal currencies offers some insights. Prior to 1971, the UK currency system of 'pounds, shillings and pence' was based on the 240:12:1 Carolingian system (introduced by Charlemagne in the late 8th century) with 12 pennies to the shilling and 20 shillings to the pound. Until 1971, the coins in circulation were guineas, pounds, crowns and half crowns, florins and shillings, sixpences, threepenny 'bits', pennies, halfpennies and (until 1961) farthings.

Russia adopted the world's first decimal currency in 1704, with France following in 1795 along with their adoption of the metric system.

The UK Parliament first considered decimalising the British currency in 1824. It was another 17 years before the Decimal Association was founded in 1841, in support of both decimalisation and use of the (SI) metric system. A report by a standards commission in the 1860s recommended metrication for Britain.

It was a further 100 years before the British Government introduced the 'Committee of the Inquiry on Decimal Currency'. Their 1963 report resulted in the final agreement to adopt decimalisation on 1 March 1966, with the approval of the Decimal Currency Act (and the setting up of a Decimal Currency Board to prepare a public information campaign) in May 1969, and the final currency changeover on 15 February 1971. The old halfpenny, penny, and threepenny coins were promptly removed from circulation, in August of the same year.

Although the 'older generation' found it more difficult, the changeover to decimal currency went smoothly. Some resisted, but I can't believe anybody today mourns the 'good old days' of pounds, shillings and pence.

The move to the wider metric system has meanwhile been piecemeal. The UK government committed to adopting the metric system on joining the European Economic Community in 1973, and regulations in 1995 required goods to be sold in metric units. In a high profile case in 2002, five market traders lost their court battle for the right to trade in pounds and ounces. They were supported by celebrities including the comedian John Cleese, and politicians including Boris Johnson.

Sir Jacob Rees-Mogg (Minister for Brexit Opportunities & Government Efficiency, Feb–Sep 2022; and today, Secretary of State for Business, Energy & Industrial Strategy, BEIS) is a high-profile advocate of imperial measures.

In what was seen as an attempt to revive this archaic system, the BEIS conducted a consultation in Jun–Aug 2022: *'If you had a choice, would you want to purchase items: (i) in imperial units, (ii) in imperial units alongside a metric equivalent'*. The absence of a third option, metric only, was widely criticised. With more than 100 000 responses, the majority expressing little appetite for increased use of imperial measures, the Government decided not to introduce any new legislation in this area. *'But new guidance has been issued to promote awareness and use of imperial measurements'*. Difficult to believe?!

I WILL NOT review the many arguments for metrication, other than to say that the metre and kilogramme sit at the heart of a *coherent system in which units are interrelated*. Like decimal currency, it is easy to learn and to use, *because* it is mainly decimal. And, importantly, the same units, used exclusively across scientific research and industry, can also be used for shopping and cookery. Extensive background is given by the UK Metric Association, and under wikipedia's metrication in UK.

ALMOST ALL other countries in the world have managed the conversion from imperial. This includes some of the most recent, including the Commonwealth countries of Australia in 1969 (where the Government's decision was based on an all-party recommendation following a Select Committee inquiry), and Canada in 1970. Closer to home, metrication in the Republic of Ireland, including speed limits, was officially completed in 2005.

In the UK, an enormous amount of work *has* been done in assembling the case for metrication. As just one example, the question of whether to convert trade and industry was the subject of a UK Government White Paper in 1951, itself the result of the Hodgson Committee Report of 1949 which unanimously recommended currency decimalisation and metrication within 10 years. The Metrication Board, to promote and plan for metrication, was set up in 1969… but wound down in 1981.

FOR THOSE CONCERNED that the 'move to metric' is an imposition from beyond our shores, it is useful to recall that development of the system is an ongoing international effort, notably within the framework of the International Bureau of Weights and Measures (BIPM) to which many British scientists continue to contribute.

And the names of British scientists are prominent amongst the measurement units of the metric system: Isaac Newton gave his name to the unit of force, James Watt to the unit of power, Michael Faraday to the unit of capacitance, James Joule to the unit of energy, and Lord Kelvin to the unit of absolute temperature.

IN CONTRAST to the views of advocates of imperial measurements, the supporters and endorsers of metrication are many and varied. And, not surprisingly, I find their arguments considerably more compelling.

Jim Al-Khalili, physicist and broadcaster, spoke for many when he said *'As a scientist, using the metric system is for me more than just a matter of ideology, convenience or a badge of honour showing off my credentials as a progressive global citizen in the modern world. Rather, it is the only way I am able to do good science and develop my understanding and knowledge of the universe'*.

The Member of Parliament Sarah Olney has emphasised the importance for the younger generations: *'By retaining our use of the metric system, we maintain compatibility with the overwhelming majority of nations across the globe.… Ensuring our children are educated in a system that allows them to maximise their opportunities across the world is key. It is time we move away from the rhetoric and the nostalgia, and look at what is best for the educational needs of the next generation'*.

TODAY, petrol from UK pumps is sold in litres, as it has for 30 years, but efficiency is measured in miles per gallon. Milk is sold in 2.272 litre containers which, being exactly 4 pints, adheres to the letter of the legislation if not the spirit. And from 2023, wine can be sold by the pint! Roads signs are in miles (and yards), while some modern cars have reverted to speeds marked *only* in miles per hour. The Government's rationale for rejecting speed limit signs in km/h in 1970 is illustrative.

Meanwhile, the pharmaceutical industry fully embraced the metric system in the 1960s. And while medical records, by law, now use metric units, most people in the UK, if asked their weight or height, would probably still reply in stones and pounds, or feet and inches.

WIDER SOCIETY is, I believe, confused. But it is also out-of-step both with the UK's actual education system which is today founded on metric units, and with 90% of its trading partners around the world. The UK Government supports science, technology, engineering, and maths (STEM), fields that emphasise innovation, problem-solving, and critical thinking. But it fails to steer a clear course on metrication.

Of course it must exercise caution in imposing unnecessary laws and regulations. But the success of other countries offers both guidance and encouragement.

For me, the UK stance on metrication is a symptom of two issues of concern: a tendency by wider society to cling to some long-gone 'glorious' past. And the unwillingness of Government to take sufficient lead on issues critical for the long-term prosperity of the country, but that might be unpopular in the shorter term.

Is it not time to break with the US, Liberia, Myanmar, and just a few other countries that retain the imperial system? Or is it the rest of the world that has it all wrong?

171. The Small Magellanic Cloud

THE MAGELLANIC CLOUDS are two 'nearby' irregular dwarf galaxies, visible to the unaided eye in the dark skies of the southern hemisphere.

The Large Magellanic Cloud (LMC) has a diameter of 4.3 kpc, compared with about 30 kpc for the Milky Way. It lies at a distance of 50 kpc, corresponding to a parallax of 20 micro-arcsec. The Small Magellanic Cloud (SMC) has a diameter of 2 kpc, and lies about 60 kpc away (de Grijs & Bono, 2015). The two are separated by 20° on the sky, 23 kpc apart. Only the smaller Sagittarius dwarf elliptical (discovered in 1994), and the Canis Major dwarf galaxy (discovered in 2003) are closer neighbours.

Both LMC and SMC probably have large dark matter halos. The LMC is believed to be the fourth most massive of over 50 galaxies comprising the 'Local Group'. Observations and theory suggest that both have both been distorted by tidal interaction with the Milky Way. Their gravity has, in turn, affected our own Galaxy, distorting the outer parts of our Milky Way's disk.

Whether the LMC and SMC are bound as orbital companions to our Milky Way remains uncertain. If they are, their orbital period is at least 4 Gyr. The other possibility is that they are on a first (or even second) approach, and we are witnessing the start of a merger that may overlap with the Milky Way's expected merger with the Andromeda galaxy sometime in the future.

I DESCRIBED some of Gaia's first insights into the Magellanic Clouds in essay 38, based largely on an analysis of the Gaia EDR3 data release by Luri et al. (2021). Using the positions, parallaxes, and proper motions, they identified ~11 million Gaia sources in the LMC, and some 1.7 million in the SMC. Further division according to star colour allows the stars to be grouped according to age and evolutionary phase.

Smoothed maps of the proper motion field showed a clear ordered rotation of the LMC, while the SMC is more chaotic. And one of the most prominent features in their outskirts is the existence of a bridge between them, attributed to tidal forces that strip gas and stars from the least to the most massive galaxy.

THE LINE-OF-SIGHT structure of the SMC is complex and, collectively, the data have been difficult to interpret (e.g. Murray et al., 2024). Substantial advances in the past 5–10 years, both observational and modelling, have benefitted from a wide range of observations, including from HST, the VISTA–VMC survey, and APOGEE.

Today, Gaia's proper motions, radial velocities, and parallaxes (for foreground suppression), are proving to be of great importance in understanding its structure. And amongst the most recent work, Murray et al. (2024) suggest a radically new picture: that the SMC is composed of two distinct structures, with its interstellar medium arranged in two, superimposed, star-forming regions separated by ~5 kpc along the line-of-sight.

TO APPRECIATE the complexities and set the scene, I will first summarise some of the most recent work. Much of this exploits the Gaia data, which I have indicated here (as superscript) according to data release.

The oldest stellar populations are reasonably spherical within a radius of ~10 kpc, with suggestions of rotation in the central region (Gaia Collaboration et al., 2018[DR2]; Niederhofer et al., 2018; Zivick et al., 2018; Niederhofer et al., 2021[DR2]). But stars with estimable distances (red clump stars, Cepheids, and RR Lyrae) extend some 20–30 kpc along the line-of-sight (Scowcroft et al., 2016; Ripepi et al., 2017; Zivick et al., 2021[DR2]).

In contrast, stars of the young main sequence and red giant branch display a radial velocity gradient indicating rotation (El Youssoufi et al., 2023[EDR3]), along with distinct substructures along the line-of-sight, both morphological (e.g. Subramanian et al., 2017; Martínez-Delgado et al., 2019[DR2]; Tatton et al., 2021; Omkumar et al., 2021[DR2]; Cullinane et al., 2023[EDR3]; Almeida et al., 2024[DR3]), as well as chemical (Hasselquist et al., 2021; Massana et al., 2022; Mucciarelli et al., 2023).

Further complicating the observed morphological structure is the evidence of tidal disruption by the LMC (Niederhofer et al., 2018; Zivick et al., 2019[DR2]; De Leo et al., 2020[DR2]; Zivick et al., 2021[DR2]; Niederhofer et al., 2021[DR2]; Cullinane et al., 2023[EDR3]).

THE SMC IS ALSO important as a laboratory for studies of the interstellar medium and star formation at low metallicity (~20% solar; Russell & Dopita, 1992). And it too presents various complexities that remain poorly understood. Indeed, starting 70 years ago, studies indicated multi-peaked H I velocity profiles, suggesting the presence of sub-systems at different distances (Kerr et al., 1954; Johnson, 1961; Hindman, 1964).

It has remained unclear whether this structure has originated from gravitational interactions with the LMC (Murai & Fujimoto, 1980; Mathewson & Ford, 1984; Ma et al., 2023), or as a series of expanding gas shells (Hindman, 1967; Staveley-Smith et al., 1997). And whether the integrated H I velocity field is consistent with a rotating disk (Stanimirović et al., 2004; Di Teodoro et al., 2019), or with the motions of its young stars (Evans & Howarth, 2008; Dobbie et al., 2014; Murray et al., 2019).

Some of this complexity has been attributed to tidal interaction with the LMC, either on its first infall (Besla et al., 2007; Zivick et al., 2018), or second (Massana et al., 2022; Vasiliev, 2024[EDR3]), or indicative of a close impact (Zivick et al., 2021[DR2]; Choi et al., 2022[DR2]). There is stellar debris in the outer regions of both (Pieres et al., 2017; Choi et al., 2018; Martínez-Delgado et al., 2019[DR2]), including the gaseous features of the Leading Arm, Bridge and Stream (For et al., 2014; For et al., 2016).

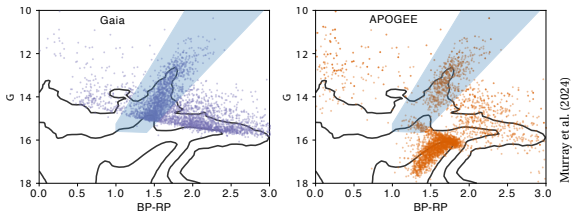
Models of the structure and dynamical history of the LMC–SMC–Milky Way system must evidently account for interactions between the LMC and the Milky Way (Besla et al., 2012; Pardy et al., 2018; Lucchini et al., 2021), but also for interactions with the SMC (Patel et al., 2020[DR2]), a challenge compounded by the uncertain morphology and dynamics of the SMC itself.

ACROSS THESE STUDIES, the Gaia data have been used to identify the systematic space motion of the SMC, providing evidence for a moderate rotation, some expansion, tidal stripping, and bursts of star formation.

But there has been a recurrent theme that its morphology, kinematics, and chemistry is not well replicated by a single population. For example, from a Gaia sample of red giant stars, Zivick et al. (2021) concluded with the need *'…to treat the SMC as a series of different populations with distinct kinematics'*.

A MAJOR MOVE in this direction has been suggested by Murray et al. (2024). They compared the structure of the interstellar medium of the Small Magellanic Cloud with the kinematics of young massive stars matched to the gaseous structures in which they probably formed.

In their study, the interstellar medium was traced by high-resolution observations of H I (neutral atomic hydrogen) from the Galactic Australian Square Kilometer Array Pathfinder survey (GASKAP–HI). The relevant stars were identified on the basis of their precise radial velocities, from both the APOGEE and Gaia surveys.



While the 21-cm H I emission map does not provide information on the line-of-sight distances, relative distances can be probed by means of the total dust extinction in the direction of individual stars, based on their infrared colours. Then if the average extinction towards stars with lower radial velocities is *less than* towards those with higher velocities, the lower velocity component must lie 'in front' (and vice versa).

The method is based on the use of the stellar radial velocities as a structural probe, although the Gaia DR3 astrometry was also used to help define both APOGEE and Gaia star samples by providing a robust membership probe. For their APOGEE sample, they started with 5938 stars within a 9° radius of the SMC centre, further restricting the sample to 2407 SMC members based on proper motion and radial velocity criteria.

The Gaia sample, selected through parallax and proper motion criteria, yielded just over 2 million initial SMC members. Further restricting the Gaia sources to those with accurate radial velocities from the Radial Velocity Spectrometer (and bearing in mind that the DR3 radial velocity catalogue is restricted to $G < 16$ mag), and within an appropriate radial velocity interval, resulted in a Gaia radial velocity sample of 3707 stars.

The colour–magnitude diagrams (G versus BP–RP) of the both the Gaia and APOGEE samples are shown in the figure above, with the red supergiants occupying the blue highlighted region. The effect of the magnitude limit of the Gaia radial velocity sample is evident in the left figure, but so is the very large number of brighter stars. For the 548 stars in common between the APOGEE and Gaia samples, the radial velocities agree to within $0.03^{+2.2}_{-1.6}$ km s^{-1}. The combined sample comprises 1947 stars (1165 from Gaia, 782 from APOGEE).

DETAILS ASIDE, their results are simply stated. By comparing the average dust extinction towards nearly 2000 stars with accurate radial velocities (from Gaia and APOGEE), and with membership rigorously established on the basis of the Gaia DR3 astrometry, they conclude that the inner ±4° of the Small Magellanic Cloud is composed of two structures with distinct stellar and gaseous chemical compositions. Specifically, the interstellar medium is organised into two, superimposed, star-forming systems with similar gas mass, separated by ~5 kpc along the line-of-sight.

The Gaia data have, in short, led to a radically new picture of this relatively nearby, well-studied, system.

172. The basic angle

IN THE LANGUAGE of space astrometry, and specifically for ESA's space astrometry missions Hipparcos and Gaia, the 'basic angle' is the angle between the instrument's two viewing directions on the sky. For Hipparcos, the basic angle was 58°. For Gaia it is 106.°5.

What is the reason for the two fields of view in the first place? How is the angle between them chosen? And why is it so different for Hipparcos and Gaia?

TO START at the beginning: the main reason for making astrometric measurements from space is to avoid the phase fluctuations caused by Earth's turbulent atmosphere. The effect is particularly important in the lowest 10 km or so, the troposphere, and is driven by convective heating rising from the Earth's surface. The turbulence affects light rays passing through the atmosphere, and causes the familiar twinkling of star light.

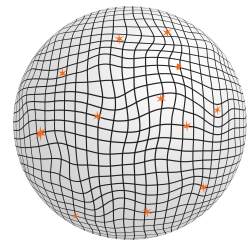
The thinner atmospheres of high mountain sites diminishes but cannot eliminate these turbulence-driven effects. Over the past 20–30 years, a combination of adaptive optics, interferometers and laser 'guide stars' has succeeded in combatting atmospheric effects over small angles of up to a few degrees, but the atmosphere still imposes an impenetrable barrier over the large angles needed for an all-sky celestial reference frame.

THERE ARE MANY technical challenges in reaching the positional accuracies that are targetted today. The 1 milli-arcsec of Hipparcos corresponds to the angular size of a golf ball viewed from across the Atlantic. The Gaia accuracies of 10 micro-arcsec correspond to the 'Bohr' radius of a hydrogen atom at a distance of 1 m.

Space offers three other big advantages in achieving these demanding accuracies. The 'weightless' environment eliminates effects introduced by the flexing of telescopes under their own weight as the supporting structures are steered to observe different parts of the sky.

And the stable thermal environment helps to minimise the tiny thermo-mechanical distortions as ground-based observatories go through their inevitable day and night cycles of warming and cooling.

A THIRD major complication is that any telescope on Earth can observe only part of the sky at any one time: a telescope in the northern hemisphere only ever sees the northern skies. Even so, it still requires a year to elapse for the entire region to be observable at night. A grid of star positions spanning the entire sky can only be constructed from a vast spider web of thousands of geometrical triangulations from separate telescopes observing accessible portions of the sky at different times.



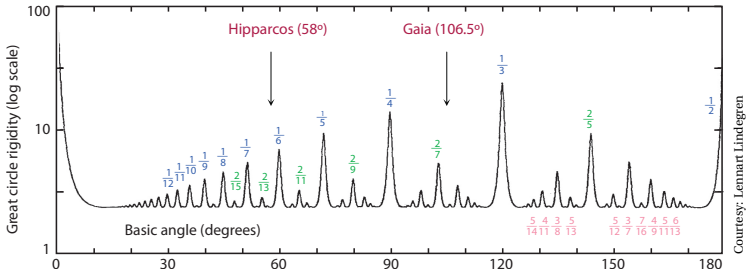
Like medieval surveys of the Earth made with early instruments, the result of centuries of celestial cartography pre-Hipparcos was a map of the sky, but one which was greatly warped and distorted.

Positions were quite unreliable much below a second of arc, and plagued by unknown errors. Within this distorted reference system, the determination of stellar parallaxes and proper motions at the angular accuracies demanded was simply impossible.

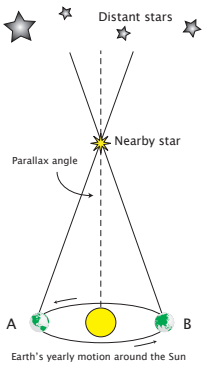
A SINGLE TELESCOPE operating above the atmosphere still cannot create a global reference system free of medium- and large-scale distortions. Imagine trying to create a map of the Earth with a 1-m ruler: local measurements might be tied together well, but the distance from Paris to Warsaw would be hopelessly inaccurate.

The solution, first proposed by French astronomer Pierre Lacroute in the 1960s, is a telescope with *two* widely separated fields of view, superimposed at the same focal plane. As I will outline below, being able to measure accurate relative star positions over wide separations on the sky unlocks the possibility of constructing a rigid global stellar reference system.

The same sort of wide-angle measurements cannot be performed from the ground. This is mainly a result of atmospheric turbulence, which becomes worse away from the zenith, and makes wide-angle measurements even more error prone. It is compounded by gravitational instrument flexure, and other complexities.



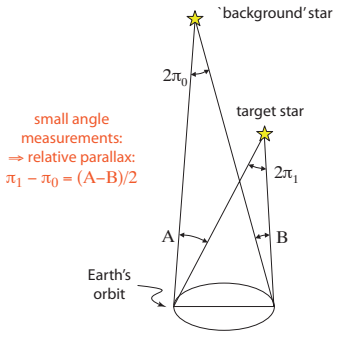
THE ABILITY TO MAKE accurate wide-angle measurements is also critical in overcoming one of the other major challenges in astrometry, tied to the measurement of a star's trigonometric parallax. Recall that a star's distance can only be determined from the relative displacement of its apparent position with respect to one or more background stars, as the Earth (and the satellite) moves in its annual orbit around the Sun.

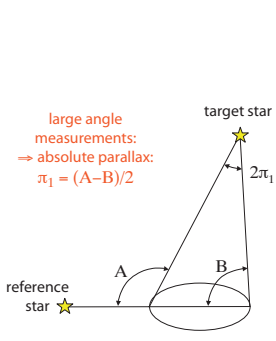


The problem is that we cannot rely on a star's parallax measured with respect to other stars nearby on the sky, but whose distances are themselves unknown (left figure below). Many methods have been tried, unsuccessfully, to convert such *relative* parallaxes to *absolute* parallaxes.

SUPERIMPOSING two widely separated fields allows the problem to be solved (right figure below). Imagine if we could superimpose an image of the target star with a reference star 90° away along the Earth's orbit. There is now no component of parallax for the reference star in this configuration, so that its parallax (or distance) is irrelevant. The situation is reversed 3 months later.

As I explain below, the superimposed fields should in fact *not* be orthogonal, but as long as they are separated by a reasonably large angle, the network of observations made from a continuous scanning of the celestial sphere permits *absolute* parallaxes to be determined.





WITH TWO widely separated fields providing absolute stellar parallaxes, the final question is: what should the angle between the viewing directions be?

The satellite's spin axis precesses around the Sun direction at a constant Sun-aspect angle (for thermal stability). At the same time, the satellite revolves around this spin axis – once every six hours for Gaia (more details are given in Gaia Collaboration et al., 2016).

Consider now just one of these great circle scans. If the viewing directions were separated by 90°, a little reflection will reveal that a given star will be connected to a star 90° away on the sky. That star, in turn would be connected to another star 90° away as the satellite spins, and so on. The end result is that the target star will be 'connected' to only three other fields, separated by 90°, 180°, and 270°. The resulting network of observations will have a poor degree of connectivity, or rigidity.

Fields separated by 60° perform a little better, but they are only connected to a total of five other regions along any great circle scan. It is not too difficult to see that small integer fractions of 360° (1/2, 1/3, 1/4, 1/5, etc.) are a poor choice if we want to create a network of star observations around the great circle with good rigidity. By extension, so too are small integer *multiples* of these angles, such as 2/3, 2/5, 3/7, etc. of 360°.

More formally, angles to be avoided are $360° \times m/n$, for small integer values of m and n.

DETAILED STUDIES in the early 1980s (Hoyer et al., 1981; see also Makarov, 1998), resulted in the kind of diagram shown above. Here, the ordinate is a measure of the rigidity of the system as a function of the 'basic angle' between the two viewing directions. Large values, and in particular the higher peaks, correspond to poorer great-circle rigidity. But between the peaks are small 'valley' regions providing good rigidity.

Any of these 'valleys' would have worked well for Hipparcos and Gaia, and the choice mainly rested on the physical accommodation of the optical hardware. Underpinned by detailed studies led by Lennart Lindegren, we chose 58° for Hipparcos, and 106.°5 for Gaia. Some more intricate subtleties aside, many other choices in the range 40 – 140° would have worked equally well.

In the earliest design studies for Gaia, we also considered a system with three viewing directions, the pairs separated by a carefully optimised 54° and 68°.

173. The breathing motion of spiral arms

I LOOKED AT spiral arms in essay 114, where I focussed on their structural features being revealed by Gaia. Here I will look at some remarkable and important insights that Gaia is revealing in the kinematics of the spiral features. As I will explain, Gaia is providing observational confirmation of so-called 'breathing modes' in the spiral arms which, in turn, support some of the theoretical and numerical models being developed to understand their origin and nature.

Let me first set the scene with a broad picture of what is known about spiral arms in general, and the spiral arms of our own Milky Way Galaxy in particular.

Spiral arms are common features of disk galaxies, both locally and at higher redshifts. The general consensus is that they are a manifestation of density waves, perhaps most generally arising from self-excited disk instabilities as seen in numerical simulations (e.g. recurrent 'groove modes', or 'swing amplification' of noise), although with some plausibly excited by other mechanisms such as the galaxy's central bar, or tidal encounters with other galaxies (Sellwood & Masters, 2022).

IN OUR OWN Galaxy, the disk comprises a central extended bar, beyond which are a series of spiral arms, delineated by a higher density of gas and dust, and more pronounced regions of ongoing star formation. Following the first identification of spiral structure by Morgan et al. (1952), subsequent observations and probes have been many and varied, including H I regions at 21 cm, H_2O masers with VLBA, molecular clouds in CO, H II regions probed by pulsars, red clump stars from 2MASS and Spitzer, cluster mapping by WISE, and others.

In essay 114, I described the advances being made in delineating the spiral morphology using Gaia data. Amongst these, Drimmel et al. (2023) used some 580 000 OB stars, together with 988 young open clusters, to map the spiral structure associated with star formation out to 4–5 kpc from the Sun. Further mapping of the spiral arms using Gaia DR3 combined with LAMOST and other large-scale surveys, has since been reported (e.g. Kounkel et al., 2020; Hawkins, 2023; Xu et al., 2023b).

AS I CONCLUDED in essay 114, and despite our 'close-up view' of the Milky Way, many observational details of the spiral structure of our own Galaxy remain unclear. Neither is there an unambiguous excitation mechanism known to be driving them. But for the rest of this essay, we can put the morphological details to one side, and look at some fascinating kinematic aspects being revealed by Gaia, and in particular evidence for what has been termed the 'breathing mode'.

Already pre-Gaia, the existence of large-scale non-zero mean *vertical* motions with respect to the disk had been identified in various surveys, notably from SEGUE (Widrow et al., 2012; Widrow et al., 2014), LAMOST (Carlin et al., 2013), and RAVE (Williams et al., 2013).

Observationally, these vertical motions (i.e. perpendicular to the disk) fall into two distinct classes: a 'bending motion' in which stars on either side of the mid-plane move together in the same direction, and a symmetric 'breathing motion' in which stars on either side move together in opposing directions, both sides either moving towards or away from the mid-plane.

Widrow et al. (2014) suggested that both could be excited by a passing satellite or dark matter subhalo, the bending motion dominating when the perpendicular component of the impactor's velocity is small compared with that of the stars, with the breathing modes excited at larger vertical velocities of the impactor.

CERTAIN PUZZLING features of these and other observations led Faure et al. (2014) and Debattista (2014) to investigate whether these vertical velocity features could arise in the absence of such an external perturber.

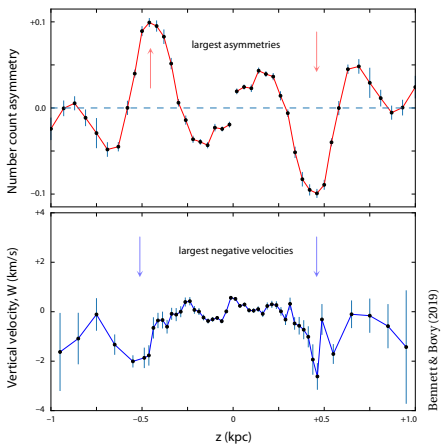
Faure et al. (2014) showed that in an equilibrium axisymmetric galactic disk, in which the mean galactocentric radial and vertical velocities might be expected to be zero everywhere, their 3D test-particle simulation showed instead that the global stellar response to a spiral perturbation induces both a radial velocity flow, and non-zero vertical motions, with a resulting mean velocity field qualitatively similar to what had been observed from SEGUE, LAMOST, and RAVE.

IN MORE detailed models, Debattista (2014) used three-dimensional N-body simulations to confirm that spiral structures indeed induce bulk vertical velocities, as large as 10–20 km s^{-1}. He found that the vertical motions are compressive (pointing towards the mid-plane) as stars enter the spiral, and are expanding (away from the mid-plane) as they leave it. As he explains: *'Since stars enter the spiral on the leading side outside the corotation radius, and on the trailing side within corotation, the relative phase of the expanding and compressive motions switches sides at corotation. Furthermore, because stars always enter the spiral on the shallow density gradient side and exit on the steeper side, the expanding motions are larger than the compressing motions'*.

Analytical models (Monari et al., 2016a), further N-body simulations (Monari et al., 2016b), and large-scale cosmological simulations (Grand et al., 2016) have confirmed these bending and breathing modes. More recently, Ghosh et al. (2022b) made high-resolution simulations of a system with prominent spiral structure, including effects of star formation. They identified two further *predicted* features: that the breathing motions induced by spiral structure have an increasing amplitude with mid-plane distance. And at any given height, the breathing motion amplitude decreases with age.

CAN ANY OF these detailed predictions be investigated with the Gaia data? Gaia has certainly revealed a rich and complex variety of kinematic substructure in the phase-space distribution (i.e. of positions and velocities) of the solar neighbourhood, amongst them imprints of the Hercules stream (essay 115), the Arcturus and HR 1614 streams (essay 116), and the Gaia 'phase-space spiral' (Antoja et al., 2018; essay 117).

Such bending and breathing modes in the solar neighbourhood were soon identified: in the RAVE–TGAS data (Carrillo et al., 2018), in a 3.2 million DR2 giant star sample (Katz et al., 2022), and in the number counts (2 million stars) and velocity distributions (865 000 stars) in DR2 by Bennett & Bovy (2019), as shown here.



IN AN ANALYSIS of 3.1 million Gaia DR2 stars within 10 kpc (with distances, and importantly ages, as given by Sanders & Das, 2018), Ghosh et al. (2022b) demonstrated that, at the location with the largest breathing motion ($x = 7.6 - 8.1$ kpc, $y = 0.9 - 1.4$ kpc) the amplitude increases monotonically with distance from the mid-plane, and decreases with age, i.e. the breathing motion is strongest for the youngest stellar populations.

With both these observational signatures consistent with their numerical simulations, they concluded that the observed breathing motions are indeed driven by spiral density waves, while the bending motions are not.

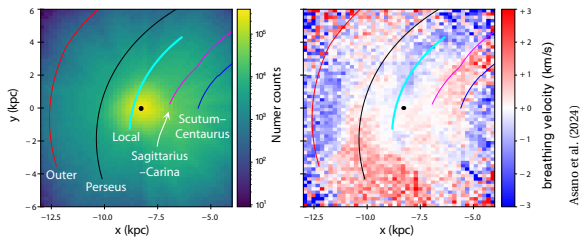
THESE FINDINGS became even more convincing with the improved accuracies of Gaia DR3, along with more robust estimates of the complex selection effects entering number density estimates (notably the satellite scanning law, the presence of open star clusters, and the effects of reddening and extinction).

Widmark et al. (2022) used several million DR3 stars to distinguish three distinct structures: the presence of the 'Gaia phase–space spiral', a large-scale bending mode seen in both number density and vertical velocity, and an elongated over-density in star counts with a corresponding breathing-mode compression in vertical velocity at the location of the Local Spiral Arm.

Asano et al. (2024) worked with a sample of 26 million DR3 stars to confirm a number of features of the now well-established breathing motion. They found a clear alignment of the compressing breathing motion with the Local Arm, similar to that seen in the growth phase of spiral arms in numerical simulations. They concluded that the Local Arm's compressing breathing motion can be explained by it being in the growth phase of a transient and dynamic spiral arm.

They also found tentative signatures of the expanding breathing motion associated with the Perseus arm, and a compressing breathing motion coinciding with the Outer arm, implying that the Perseus and Outer arms are in the disruption and growth phases respectively.

In their face-on map of number counts (left), the spiral arms are not evident due to the various selection biases; the locations shown are those of massive star-forming regions from Reid et al. (2019). But the alignment between the Local Spiral Arm and the compression mode in the Gaia data (right) is particularly striking. More insights will come from future Gaia data releases.



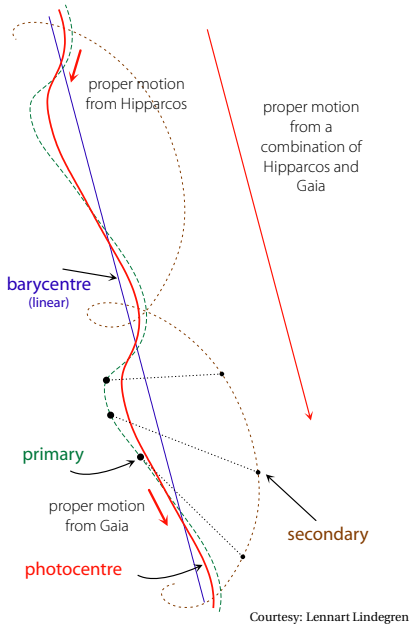
174. Proper motion anomalies

BINARY STARS are loosely classified as *visual* (where both components are resolved), *astrometric* (where the companion's presence is revealed by the photocentric motion), *spectroscopic* (showing orbit variations of spectral lines), or *eclipsing* (where light from one component is eclipsed by the orbital motion of the other).

I have discussed binary systems in essays 79 and 134 (I use the term 'binary' although higher multiplicities are usually also implied), where I explained how Gaia is contributing, profoundly, to their discovery and classification over this broad range of phenomenological type.

A subset of astrometric binaries effectively evaded identification as such over the observation period of the Hipparcos mission, but they can become recognisable from the difference in the proper motion determined by Hipparcos and that measured by Gaia.

These are frequently referred to as systems showing 'proper motion anomalies'. I will explain the phenomenon, the recent work which has underlined their ubiquity, and give some examples of their importance.



THE DIAGRAM illustrates the principles. It shows the sky-projected orbital motion of the primary and secondary components of a representative binary, the linear motion of the system's barycentre (blue), and the non-linear motion of the photocentre (red). It is evident that the proper motion (of the photocentre) measured by any series of observations is a function of the orbital period compared with the measurement duration.

If, for example, the observation interval covers only a small fraction of the orbital period, then the orbital motion may not be distinguishable from the underlying space motion. The measured proper motion vector can then deviate substantially from that measured by Gaia, or from the long-term barycentric motion over the time interval between the Hipparcos and Gaia missions.

For orbital periods of, say, 20–30 years or more, space astrometry can then exploit the near 30-year time interval between the Hipparcos and Gaia proper motions, because each samples the star's reflex motion at very different epochs in a long-period orbit.

THE HIPPARCOS CATALOGUE included a main part giving the single star (5-parameter) solutions, and a 'Double and Multiple Systems Annex' with five categories of astrometric solution (Lindegren et al., 1997):

- component solutions (C): with component separations in the range $\rho = 0.1-30$ arcsec, for which the 5 astrometric parameters are given for each component;
- acceleration solutions (G): apparently single stars with significantly non-linear motion, fitted by a quadratic or cubic polynomial of time (7- or 9-parameter solution);
- orbital solutions (O): with full orbital solution for the photocentre in addition to the 5-parameter astrometric solution for the barycentre;
- variability-induced movers, or VIMs (V): these are unresolved binaries in which one component is variable (I have touched on these already in essay 163);
- stochastic solutions (X): for those cases where no reasonable single or double star solution could be found.

The numbers of each are given in the following table, where it is clear that the single-star solutions dominated.

Hipparcos: solution type	Annex	No. of entries
Single-star solutions	–	100 038
Component solutions	C	13 211
Acceleration solutions	G	2 622
Orbital solutions	O	235
Variability-induced movers	V	288
Stochastic solutions	X	1 561
No valid astrometric solution	–	263
Total number of entries		118 218
Entries with valid astrometry		117 955

AVAILABLE to probe such proper motion anomalies are the Hipparcos observations, which extended over 3.3 years (1989.8–1993.2) with catalogue mean epoch J1991.25. Gaia DR2 spans an observation duration of 22 months with reference epoch J2015.5, and Gaia DR3 covers a 34 month interval with reference epoch J2016.0.

Estimates of the general fraction of stars that are binaries vary widely, from around 30–35%, with a lower fraction amongst low-mass stars, and perhaps up to 80% for the most massive O–B stars (e.g. Lada, 2006).

Naively, therefore, we might not be too surprised to find that some 30 000 of the Hipparcos single-star solutions display such a proper motion anomaly, albeit with different fractions amongst the various categories.

AS A SMALL aside, the Hipparcos 'stochastic solutions' replaced deterministic solutions with a quadratic increase in the standard error (typically 3–30 mas) such that the rms normalised residual was equal to unity.

The similar *re-normalised unit weight error*, RUWE, itself related to the reduced χ^2 of the single-star solution, is included in the Gaia catalogues to capture photocentric perturbations due to unmodelled orbital motion or calibration errors (Lindegren et al., 2018, §5.1).

Tested on a sample of known spectroscopic binaries, Belokurov et al. (2020a) showed that the implied amplitude of the centroid perturbation scales, as expected, with the binary period and the mass ratio, and illustrated how the binary fraction evolves across the HR diagram.

FIRST STEPS in deriving long-term proper motions using the Hipparcos observations as one epoch of a much longer time span was the 2.5-million star Tycho 2 Catalogue (Høg et al., 2000). This combined the Tycho 1 Catalogue ($\sigma_\mu \sim 10-25$ mas/yr) with the much earlier Astrographic Catalogue ($\Delta t \sim 100$ yr), and 143 other ground-based astrometric catalogues, to create Tycho 2 with long-term proper motions with $\sigma_\mu \sim 2.5$ mas/yr.

The first proper motions resulting from a combination of the Hipparcos and Gaia observations was the Tycho–Gaia astrometric solution (TGAS, Michalik et al., 2015), for which 2 million sources with a proper motion accuracy of ~1 milli-arcsec/yr (around 0.06 mas/yr for the subset of 94 000 Hipparcos stars) were included as part of the Gaia DR1 data release (Brown et al., 2016).

I HAVE OUTLINED the motives underlying the determination of long-term proper motions, and the implications of 'proper motion anomalies' when comparing the difference between their long-term Hipparcos–Gaia and short-term Gaia proper motion vectors. Let me now briefly survey the results of various investigations of the Hipparcos–Gaia 25-yr baseline proper motions.

A number of early studies used Gaia DR2 (Brandt, 2018; Kervella et al., 2019a; Makarov, 2020). Amongst these, Kervella et al. (2019b) determined a binary fraction of 80% for their sample of classical Cepheids.

Studies with EDR3 started with the Hipparcos–Gaia Catalogue of Accelerations by Brandt (2021). Kervella et al. (2022) derived tangential velocity anomalies with $\sigma(\Delta v_T) = 0.26\,\mathrm{m\,s^{-1}\,pc^{-1}}$, a factor 2.5 improvement compared to Gaia DR2. They identified 37 515 Hipparcos stars (32%) with a 3σ anomaly, and with detection limits extending into the exoplanet mass regime. Including the Gaia EDR3 re-normalised unit weight error (RUWE >1.4) as an additional measure, 50 720 Hipparcos stars (43%) exhibit at least one indicator of binarity. Of the Gaia EDR3 stars within 100 pc, they found common proper motion companions for 39 490 (7.3%).

Penoyre (2022b; 2022a) applied a similar approach to the Gaia Catalogue of Nearby Stars (Gaia Collaboration et al., 2021b). They found 22 699 binaries within 100 pc (~10%), with the fraction ranging from 20% for giants, 10% for main sequence, and ≲1% for white dwarfs.

ONE SPECIFIC APPLICATION for systems showing long-term proper motion anomalies has already been particularly well exploited. In essay 88, I explained how the implied accelerations can be used to determine companion masses, and specifically for previously imaged exoplanets, or as 'dynamical beacons' to pinpoint where on the sky a new planet must lie.

I will not go further into this topic here, other than to simply list some of the associated studies that have been made using the Hipparcos–DR2 temporal baseline (Brandt et al., 2019; Dupuy et al., 2019; Currie et al., 2020; Damasso et al., 2020; Kiefer et al., 2021), and subsequently based on Hipparcos–EDR3 (Brandt et al., 2021; Kammerer et al., 2021; Franson et al., 2022; Kuzuhara et al., 2022; Bonavita et al., 2022; Herz et al., 2024).

IN ANOTHER interesting application, Dodd et al. (2024) examined the nature of Be stars, the rapid rotating emission-line subset of B stars, for which the origin of rotation remains unclear. From the proper motion anomaly, and the Gaia-provided RUWE, they could identify unresolved binaries down to separations of 0.02 arcsec using the Hipparcos–DR3 temporal baseline.

They concluded that it is binary interactions that cause the Be phenomenon, with migration causing the dearth of Be binaries between 0.02–0.04 arcsec, and with triplicity playing a key role in the migration.

175. Black holes in open clusters

IT HAS recently become clear that the existence of stellar mass black holes should have observable consequences on the overall dynamics of open clusters and stellar streams, placing useful constraints on their formation. I will look at some early results for the Hyades open cluster in this essay, and at their effect on the morphology and kinematics of halo streams in the following.

Let me start with some context. The very existence of stellar mass black holes can be demonstrated in various ways. A black hole in a close binary system can be inferred from X-ray emission due to accreted material; through the astrometric motion of its visible companion; as ellipsoidal light variations due to tidal distortion of its companion; and, at the end of their orbital life, as a gravitational wave burst accompanying the system's final inspiral and merger. Isolated black holes may be revealed through the effects of gravitational microlensing.

I HAVE GIVEN an introduction to Gaia's contributions to gravitational wave searches in essays 131 and 136, and to the class of ellipsoidal variables in essay 133. In essay 101, I outlined some of the searches for black holes in X-ray binaries (a catalogue is maintained by Corral-Santana et al., 2016), from astrometric binary motions, and for isolated microlensing searches.

From Gaia DR3 astrometry, three non-interacting (X-ray quiescent) binary black hole candidates have been reported: Gaia BH1 with $M_{\rm BH} = 9.6 M_\odot$ at 480 pc (El-Badry et al., 2023b), and Gaia BH2 with $M_{\rm BH} = 8.9 M_\odot$ at 1.2 kpc (El-Badry et al., 2023a), making them the nearest known black hole candidates. The third, Gaia DR3 5870569352746779008, is a $> 5.6 M_\odot$ black hole in a 1350 d binary, discovered from a study of more than 60 000 DR3 binary solutions with both astrometric and spectroscopic data (Tanikawa et al., 2023).

Beyond Gaia, only one other non-interacting massive black hole candidate appears to be rather secure: the single-lined spectroscopic binary VFTS 243 in the Large Magellanic Cloud (Shenar et al., 2022). It comprises an O star of $25 M_\odot$ and an unseen companion of $> 9 M_\odot$, in a binary system with orbital period 10.4 d.

The most recent Gaia DR3-based photometric microlensing searches for isolated black holes have so far generated only a small number of unconfirmed *candidates* (Kruszyńska et al., 2024). This leaves OGLE–2011–BLG–0462 (aka MOA–2011–BLG–191) as the only microlensing black hole. Models indicate a distance of 1.5 kpc, and a lens (black hole) mass of $6 - 8 M_\odot$ (Sahu et al., 2022; Mróz et al., 2022; Lam & Lu, 2023).

WITH THE X-RAY black hole candidates lying at typical distances of 2 kpc or more, this leaves the non-interacting binary systems Gaia BH1 and Gaia BH2 as the nearest known black holes.

And this class of object is important for understanding the evolutionary pathways leading to the merging black holes observed by LIGO/VIRGO, and in particular the role of natal 'kicks' during supernova core collapse (e.g. Mandel & Farmer, 2022; Mandel & Broekgaarden, 2022). In interacting (X-ray) binaries, tidal forces circularise the binary orbit, suppressing information on these kicks previously encoded in the system's eccentricity. In contrast, weakly interacting (X-ray quiescent) binaries preserve the black-hole kick signatures in their orbits.

Stars expand dramatically beyond the main sequence, and early merging is only avoided if the initial orbit is rather wide. Models explaining the existence of these double compact binaries, which can then merge within a Hubble time, generally invoke a 'common-envelope' phase (e.g. Portegies Zwart et al., 1997; Tauris et al., 2017), allowing the system to shrink to a compact final configuration. The process is incompletely understood, and the formation of Gaia BH1 consequently uncertain (El-Badry et al., 2023b; Shikauchi et al., 2023).

Other models have considered a very different formation mechanism, viz. through dynamical interactions in a young star cluster (e.g. Portegies Zwart & McMillan, 2000; Shikauchi et al., 2020; Rastello et al., 2023). In these scenarios, the Gaia black hole binaries may be the inner binary of a triple star system in which Lidov–Kozai resonance oscillations also operate (Hayashi et al., 2023; Generozov & Perets, 2024; Tanikawa et al., 2024).

THE DISCOVERY of accreting black hole candidates in globular clusters, both extragalactic (Maccarone et al., 2007), and within the Milky Way (Strader et al., 2012; Chomiuk et al., 2013; Miller-Jones et al., 2015) has motivated specific studies of the formation of binary black holes in the centres of globular clusters (Portegies Zwart & McMillan, 2000; Antonini & Gieles, 2020), and indeed in open clusters (Di Carlo et al., 2019; Rastello et al., 2019; Kumamoto et al., 2020; Banerjee, 2021; Torniamenti et al., 2022). These discoveries already demonstrated that stellar interactions would not eject *all* of the black holes that might form in globular clusters (Kulkarni et al., 1993; Sigurdsson & Hernquist, 1993).

Other evidence for stellar mass black holes in globular clusters has been gathered from their large core radii, the lack of mass segregation, their mass-to-light ratio, and their tidal tails (for references, see Torniamenti et al., 2023). For example, and in a topic I will pick up in my next essay, N-body simulations of the halo globular cluster Palomar 5 can explain both its large extent and its extended tidal tails in models with 20% of its total mass in the form of stellar-mass black holes (Gieles et al., 2021).

ACCORDING TO Torniamenti et al. (2023), these kinds of dynamical searches for black hole populations in stellar clusters have focused on old ($\gtrsim 10$ Gyr) and relatively massive ($\gtrsim 10^4 M_\odot$) globular clusters in the Milky Way halo, with no such searches in young open clusters in the disk because of the limited phase-space (position and velocity) information hitherto available for them.

The situation has changed with Gaia, which has brought the discovery of hundreds of new open clusters, and a wealth of membership details along with accurate morphological and kinematic data now available for many (see, e.g. Cantat-Gaudin, 2022; and essays 74 and 144). This offers the prospects of quantifying their radial distributions out to their outermost regions (Tarricq et al., 2022), their extended halos (Meingast et al., 2021), and tracing the tidal tails of clusters including the Hyades (essay 20), Blanco 1, and Praesepe.

THE N-BODY CODE `PETAR`, combined with the `galpy` code for the Galaxy potential (Bovy, 2015), was used by Wang & Jerabkova (2021) to simulate the evolution of star clusters along with their tidal streams in various Galactic potentials. The most massive OB stars, and black holes, have a major effect, with clusters having the same initial conditions, but a different initial content of OB stars, following very different evolutionary paths.

As a consequence, the total initial mass and radius of an open star cluster cannot be unambiguously determined unless the initial content of OB stars is known. They showed that the stellar counts in the associated tidal tails, that can in principle be identified from the Gaia data, would help to resolve these uncertainties.

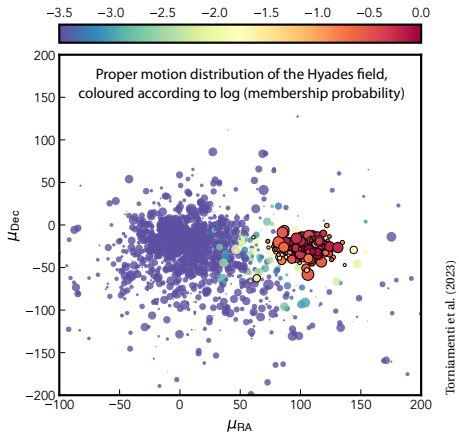
THE FIRST searches for such dynamical evidence for black holes in open clusters using this approach, and indeed following the same methods used in the Palomar 5 globular cluster study by Gieles et al. (2021), has been reported for the Hyades cluster, and using the Gaia DR3 data, by Torniamenti et al. (2023).

They generated a large suite of N-body models, as developed by Wang & Jerabkova (2021), and compared them with the radial profile of Hyades members, using masses derived from the Gaia DR3 data by Evans & Oh (2022). Their objective was to ascertain whether a residual black hole population is required to explain the observed structure. Models of binary black hole mergers, in which a significant fraction of stellar-mass black holes must receive negligible natal kicks to explain the gravitational wave detections, imply that black holes could and should be retained even in open clusters with low escape velocities of $\lesssim 1\,\mathrm{km\,s^{-1}}$.

Torniamenti et al. (2023) found that their Gaia DR3 observations are best reproduced by models with 2–3 black holes still in the Hyades cluster today. Models that never possessed black holes would have a half-mass radius significantly smaller than observed, while models in which the last remaining black holes were ejected recently, say 150 Myr ago, can still reproduce the observed radial density profile. In these model, the ejected (binary) black holes are at a typical distance of 60 pc from the cluster centre, and some 80 pc from the Sun.

In half of their models, the black holes were in binary systems with stellar companions. But their period distribution peaks at $\sim 10^3$ yr, and they are therefore unlikely to be found from orbital velocity variations. They identified 56 binary candidates on the basis of their large astrometric and spectroscopic errors, none of which were consistent with a massive compact companion.

Their models with 2–3 black holes had an elevated central velocity dispersion, which cannot yet be verified from the current observations. Nonetheless, one conclusion of their study is that the nearest stellar-mass black holes to the Sun are in, or near, the Hyades cluster!



176. Black holes in stellar streams

In my previous essay, I noted that recent work has suggested that the existence of stellar mass black holes should have observable consequences on the structure of open clusters and stellar streams. I provided some background to the formation and detection of black holes in interacting and quiescent binaries, then looked at the evidence for such systems in open clusters, and the suggestion that the Hyades possibly contains 2–3 such stellar mass black holes at the present time.

Here I will look at how the existence of stellar mass black holes might affect the morphology and kinematics of the stellar streams that are now known to exist in the inner and outer halo of our Galaxy, and to what extent Gaia can help to distinguish between those that are rich in, or devoid of, stellar mass black holes.

Although the main development that I will report on here – their influence on the structure of the tidal tails of the globular cluster Palomar 5 – is not directly based on Gaia data, the results have benefitted from, and will be more widely applicable to, many of the other stellar streams that Gaia is discovering and characterising.

These developing ideas, still in their infancy, throw light on topics such as the distinction between streams originating from globular clusters or disrupted galaxies, and why some systems being captured by our Galaxy have tidal tails and others do not.

Let me first gather a few key points by way of context, for the story has a number of facets.

The 'hierarchical merger' scenario for our Galaxy, in which it grew through mergers of smaller galaxies, developed from the ideas of Searle & Zinn (1978), and began to take shape with the first observational evidence for halo tidal streams (Majewski et al., 1996), and the discovery of the Sagittarius dwarf galaxy which appears to be in the process of tidal disruption (Ibata et al., 1994).

Today, this picture is supported by improved knowledge of the orbits of the Local Group galaxies, the presence of many more stellar streams, and by the large-scale N-body simulations of ΛCDM cosmology. The pre-Gaia development of the field of stellar streams is nicely detailed by Newberg (2016).

In this process of hierarchical galaxy formation, tidal forces slowly disrupt the accreted systems, forming the tidal tails of extant progenitors, or residual stellar streams, roughly aligned with the progenitor's orbit. Stars near the system's L1 and L2 Lagrange points are lost preferentially, and the differences in orbital frequency between stripped stars and the progenitor produce the leading and trailing tails that grow in extent with time.

Some 100 halo streams are known today, many from Gaia (e.g. Mateu, 2023), some representing the tidal debris of captured dwarf galaxies (where they can be of considerable width, that of Sagittarius being ~6 kpc), with thin streams inferred to be the debris of dissolved globular clusters. Some can be identified as stellar overdensities, others as stars with similar locations and velocities or angular momenta (essays 15, 71, and 156).

Discussions of whether individual progenitors were globular clusters or dwarf galaxies is complicated not least because of the incomplete understanding of the origin and nature of globular clusters, as well as some overlap as well as inconsistencies in their properties.

Simply stated, globular clusters were originally believed to be dense stellar systems which resulted from one generation of stars formed from a single giant molecular cloud (and thus with roughly the same age and metallicity, although the present picture is more complex), and generally considered free of gas and dark matter. Dwarf galaxies contain stars, sometimes gas, and (generally) significant amounts of dark matter.

I will not say more on the complex dynamical processes that occur in globular clusters, which include various mass-loss mechanisms (e.g. Weatherford et al., 2023; Weatherford et al., 2024), the dynamical heating of binary stars, mass segregation and, in some cases, 'core collapse' (e.g. Gürkan et al., 2004).

But let me emphasise a key point made in essay 175: that stellar mass black holes must exist in at least some globular clusters (rather than all being ejected), as evidenced by the presence of active X-ray binaries and gravitational wave generating binary mergers, and guided by simulation results (e.g. Morscher et al., 2015).

Of the several dozen *thin* stellar tidal streams now known in the Milky Way halo, none has a known progenitor. But their narrow widths ($\lesssim 100$ pc) imply that their progenitors must have had a small velocity dispersion, suggesting that they originated from dark matter-free globular clusters rather than dwarf galaxies. Streams of width 100 pc or more show an overlap between the two populations (Patrick et al., 2022).

But unambiguously associating these with disrupted globular clusters encounters some problems. One was the finding by de Boer et al. (2020) that the implied mass-loss rate of the GD–1 stream was a factor 4 higher than in early models of cluster evolution (Baumgardt & Makino, 2003). They suggested that their high stellar escape rate could specfically imply low-density progenitors, or the presence of black holes within the globular cluster.

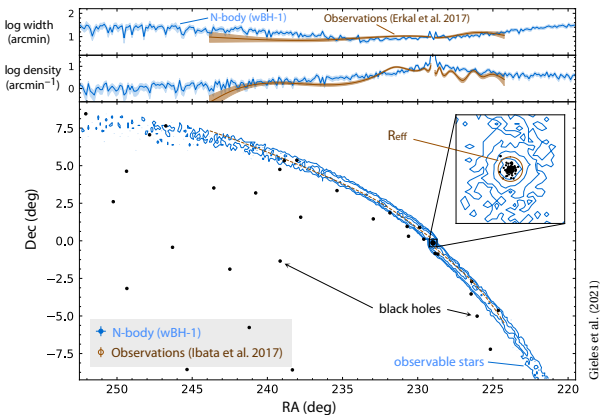
That the presence of black holes provides an abrupt cluster dissolution mechanism was demonstrated in the `MOCCA` N-body simulation of globular clusters by Giersz et al. (2019), who also found that this could lead to a strong increase of black hole and black hole–black hole binaries in the Galactic halo and bulge. Some further implications are detailed by Gieles & Gnedin (2023).

Let me now turn to Palomar 5, a distant (~20 kpc) globular cluster with an unusually low central density and two prominent tidal tails, discovered with SDSS, and extending over more than 2° on the sky (Odenkirchen et al., 2001). More recent studies, some based on Gaia data, have extended knowledge of the spatial extent of these tidal tails to around 30°, and demonstrate how its morphology and kinematics provide constraints on the enclosed Milky Way mass, and on the oblateness of its dark matter halo (essay 109).

For example, Pearson et al. (2015) found that Galaxy potential models which assume a spherical dark matter halo match the observed morphology, while no plausible model could be found in the triaxial potential of Law & Majewski (2010b), proposed to explain the properties, and 'fanning', of the Sagittarius stream.

In another pre-Gaia study, Ibata et al. (2017) showed that the paucity of low-mass stars in the cluster itself also extended to the stellar populations along the trailing arm out to 6°, implying that the ejection of the low-mass stars occurred before the formation of the stream.

Most relevant for the subject of this essay, Gieles et al. (2021) used N-body simulations, matched to the cluster members from Erkal et al. (2017), to show that both the cluster's sparseness, and its tidal tails, can be explained by a stellar-mass black hole population comprising 20% of its present mass. Their model corresponds to the presence of 124 black holes with an average mass of $17.2 M_\odot$, currently residing within its half-light radius, $R_{\rm eff}$, of 3.21 arcmin (18.7 pc).



Their models followed the cluster for 11.5 Gyr in a three-component Milky Way (bulge, disk, and halo), including the effects of stellar and binary evolution, and with two different prescriptions for the black hole natal kicks.

They concluded that Pal 5 could have formed with a 'normal' black hole mass fraction (of a few per cent). But with stars being lost at a higher rate, the black hole fraction gradually increased, further enhancing tidal stripping and tail formation. They also predicted that Pal 5 will dissolve, in ~1 Gyr, to leave a 'black hole cluster', and suggested that such black hole dominated clusters may be the progenitors of these thin stellar halo streams. Whether clusters evolve towards black hole-free clusters, or 100% black hole clusters depends, they found, on their initial density relative to the tidal density.

They also investigated whether the black hole population could be detected from the cluster's kinematics, along the lines attempted by Wan et al. (2021) for NGC 3201, or from the binary population, and the (limited) prospects of discovering them by microlensing.

Subsequent N-body simulations have suggested that the inclusion of primordial binaries has a noticeable but not drastic effect on the cluster's dynamical evolution, and that observations focussing on the cluster's velocity dispersion, and on binaries with periods of $10^4 - 10^5$ days in its inner and tail regions, will best constrain the black hole existence (Wang et al., 2024).

Further numerical simulations to quantify the differences between streams originating from star clusters with and without black holes have been made by Roberts et al. (2025). They found that, compared to streams from black hole-free clusters, those from black hole-rich clusters are some five times more massive; have a peak density three times closer to the cluster after 1 Gyr; and have narrower peaks and more extended wings in their density profile.

With various caveats, they also concluded that if the tails suggest a progenitor having a high mass-loss rate, the most likely interpretation is that the progenitor was rich in black holes.

177. An intermediate-mass BH in M4?

MY TWO PREVIOUS ESSAYS looked at how stellar mass black holes can leave an observable imprint on the morphology and kinematics of open clusters (essay 175) and stellar streams (essay 176).

Here, I look at Gaia's contribution in probing the existence of *intermediate-mass* black holes in globular clusters, and specifically in the globular cluster M4. At 6 kpc and age 12.2 Gyr (Caputo et al., 1985), M4 (NGC 6121) is the closest such system to the Sun.

STELLAR MASS black holes ($\lesssim 100 M_\odot$) are predicted to form in the late-evolutionary collapse of a single massive star, while supermassive black holes ($\gtrsim 10^5 M_\odot$) form in the high-density environment of galaxy centres. Intermediate-mass black holes, if they exist, sit between these two extremes: too massive to have formed by single star collapse, but lacking the environment necessary to form a supermassive black hole.

Unlike the observational situation for stellar mass or supermassive black holes, no intermediate-mass objects are definitively known. But interest in them lies in the fact that various secure mechanisms are expected to lead to their formation, e.g. via the merging of stellar mass black holes or other compact objects, or via the runaway collision of massive stars in dense globular star clusters. They may also exist as primordial objects formed in the Big Bang (e.g. Bernal et al., 2018).

THE DENSE STELLAR environment of globular clusters, which are compact gravitationally-bound spheroidal systems comprising tens of thousands to many millions of stars, provide environments rich in many complex dynamical processes.

N-body simulations show, for example, less-massive stars migrating outwards with the core region becoming more crowded, complex interactions and orbital modifications due to the the dynamical heating of binary stars, effects of mass segregation, stellar escape through a variety of mechanisms (e.g. Weatherford et al., 2023) and, in some cases, 'core collapse' (e.g. Gürkan et al., 2004), all further influenced by the varying tidal field in their orbits around their host galaxy (e.g. Gnedin et al., 1999).

ALONGSIDE THE FORMATION of other 'exotic' objects such as black hole–luminous star binaries, Type Ia supernovae, blue stragglers, young neutron stars and fast radio bursts, one of the potential outcomes of dynamical processes within these dense environments are such intermediate-mass black holes (Miller & Hamilton, 2002; Portegies Zwart & McMillan, 2002; Portegies Zwart et al., 2004; Giersz et al., 2015; González et al., 2021).

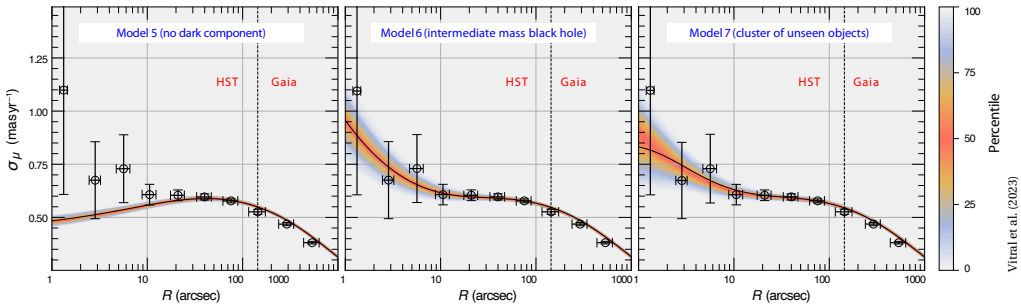
In the early N-body simulations of Miller & Hamilton (2002), for example, a massive black hole of $\sim 50 M_\odot$ is formed, either by a single massive star collapse or through collisions of a number of smaller mass objects. Importantly, such an object would be too massive to be ejected from the cluster by subsequent dynamical recoil. It then sinks to the centre of the cluster (in $\lesssim 10^6$ yr), where it continues to accrete lighter black holes to reach $\sim 1000 M_\odot$ over the $\sim 10^{10}$ yr cluster lifetime.

Predicted observable signatures (Giersz et al., 2015) include the spatial and kinematic structure of the host star cluster, possible radio, X-ray and gravitational wave emission due to dynamical collisions or mass transfer, and the creation of hypervelocity main-sequence escapers during strong dynamical interactions between binaries and such an intermediate-mass black hole.

THERE ARE numerous intermediate-mass black hole *candidates*. Chilingarian et al. (2018) identified 305 objects of mass $3 \times 10^4 - 2 \times 10^5 M_\odot$ in a sample of SDSS active galaxies, with X-ray emission inferred to originate from their accretion disks in 10 of them.

Ultra-luminous X-ray sources in globular clusters may also signal their presence but lack dynamical confirmation, with examples being NGC 4472 in the Virgo cluster (Maccarone et al., 2007), and the X-ray/optical outburst source 3XMM J215022.4–055108 (Lin et al., 2020). Other globular cluster candidates include G1 in M31 (Baumgardt et al., 2003), and MGG 11 in M82 (Portegies Zwart et al., 2004).

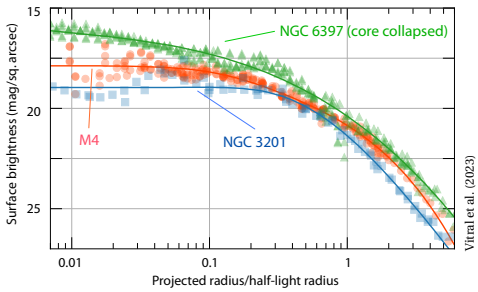
A very different candidate is the LIGO/Virgo gravitational wave event GW 190521 of 21 May 2019, in which the merger of two black holes ($85 M_\odot$ and $65 M_\odot$) probably led to a $142 M_\odot$ black hole (González et al., 2021).



THE WORK that I will now describe also refers only to an intermediate-mass black hole *candidate*, one in the globular cluster M4, and resting on Hubble Space Telescope proper motion measurements as well as Gaia. But it demonstrates the value and fidelity of the Gaia results, and underlines the future prospects especially when the next data release, DR4, becomes available.

Previous to their work on M4, Vitral et al. (2022) used HST and Gaia EDR3 proper motions to model the internal kinematics of two other globular clusters: NGC 6397 (core-collapsed, as evidenced by the central luminosity profile) and NGC 3201 (non core-collapsed). In both clusters, they found velocities consistent with isotropy and evidence for a 'dark central mass' of $\sim 1000 M_\odot$. But their models did not suggest the existence of an intermediate-mass black hole in either.

Rather, for NGC 6397 they found a strong preference for an *extended* dark central mass of size 1% of the cluster's effective radius, and with the internal kinematics consistent with a population of hundreds of massive white dwarfs. For NGC 3201 they found a mild preference for an extended dark central mass, consistent with some 100 segregated stellar-mass black holes extending over a few per cent of the cluster's effective radius.



SIMILAR MODELLING of M4 (NGC 6121) using HST and Gaia EDR3 data was then performed by Vitral et al. (2023), which in turn built on earlier work modelling the cluster's internal dynamics using both ground-based radial velocity data (e.g. Baumgardt & Hilker, 2018), N-body simulations (Hénault-Brunet et al., 2019), as well as Hubble Space Telescope imaging data to access the most crowded core region (e.g., Bedin et al., 2013; Malavolta et al., 2015; Baumgardt et al., 2022).

M4 is generally considered to be non core-collapsed, although possibly close to core-collapse (McLaughlin & van der Marel, 2005). A comparison of its surface brightness profile compared with those of NGC 6397 and NGC 3201 is shown in the figure opposite. The velocity data used by Vitral et al. (2023) provides coverage from the cluster's interior (the Hubble Space Telescope data extending from 149 arcsec down to 0.9 arcsec) out to its outermost radii, and comprises 4365 stars from HST, and 6158 stars (beyond 149 arcsec) from Gaia EDR3.

The median HST proper motion baseline is 10.6 yr, and the cluster's proximity yields high-precision motions with errors a factor 10 smaller than the cluster's velocity dispersion. This in turn provides a sufficient number of stars within ~1 arcsec from its centre to probe the possible presence of an inner dark mass.

As in Vitral et al. (2022), their fits used the `MAMPOSSt` code (Mamon et al., 2013), based on solving the Jeans equations to derive the global mass profile from the density and velocity dispersions of their tracer population (Cappellari, 2008; Watkins et al., 2013). Such models can include anisotropy and rotation (possible implications of a residual cluster rotation or non-sphericity are discussed), and can also allow for radially varying mass-to-light ratios as proxies for any differences in the spatial distributions of different mass populations.

THEY FOUND isotropic motions in the core, with tangential motions in the outer parts. As shown above for their different models, the velocity dispersion in the outer parts is particularly well-defined by the Gaia data. They infer the existence of a dark central mass of some $800 \pm 300 M_\odot$, but were unable to distinguish between a point-like intermediate-mass black hole, and a dark population of stellar remnants extending out to 0.016 pc (or occupying a region a factor of two larger when removing one high-velocity star from the cluster centre).

They conclude that the dark central mass may be either an intermediate-mass black hole, or a very compact black hole population.

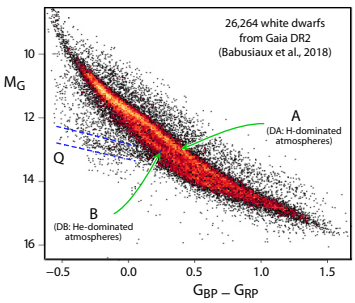
Future modelling should benefit from the data from Gaia DR4, or by making use of additional central objects from the dedicated 'Service Interface Function' scans already obtained for ω Cen, M4 and other high-surface density regions (Weingrill et al., 2023; see also essay 157).

178. Bifurcation in the white dwarf HRD

THE POSSIBILITY to position hundreds of millions of stars accurately in the observational Hertzsprung–Russell diagram was one of the central goals of Gaia, and is proving to be one of its many successes.

In this context, an early discovery was a bifurcation in the white dwarf colour–magnitude diagram, the 'Gaia gap', which I first described in essay 42 (October 2021). Revised analyses of this feature, which I discuss here, further illuminate the insights that Gaia is providing into the complex physical processes at play in these objects.

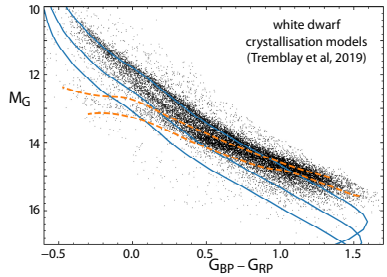
IN THEIR original paper, Babusiaux et al. (2018) used Gaia DR2 to select 26 264 nearby white dwarfs with relative parallax uncertainties better than 5%. The accurate parallaxes, combined with simultaneous accurate Gaia photometry, then yield accurate absolute magnitudes, which allow them to be precisely located in the HR diagram. Occupying the lower-left part of the diagram, the resulting white dwarf sequence showed several prominent features which they commented on.



First is the concentration of stars running continuously from upper-left to lower-right (A), coinciding with the evolutionary tracks for $0.6 M_\odot$ DA white dwarfs (whose envelopes are dominated by hydrogen). Below the main band is a second, distinct concentration (B), which they favoured attributing to DB white dwarfs (whose atmospheres are dominated by helium). The feature corresponds to those seen in Sloan Digital Sky Survey colour–colour diagrams (Harris et al., 2003).

The second prominent feature is the objects lying *above* the main DA sequence, considered to be white dwarfs in binary systems. A third, weaker structure is the rising transverse feature 'Q', a faint band of stars visible between the two dashed lines which, as Tremblay et al. (2019) and others have shown, results from core crystallisation as the white dwarfs cool (see essay 108).

I will say no more about these last two features, other than to show this figure from Tremblay et al. (2019). It illustrates (orange lines) the region over which crystallisation is relevant, which should therefore not affect the bifurcation, A–B, and is not considered further here.



TO GO FURTHER into the explanation for the bifurcation or gap, some background on white dwarfs is useful (see also the broad review by Saumon et al., 2022).

White dwarfs are the endpoint of stellar evolution for ~97% of stars, specifically those below $8 - 10 M_\odot$, depending on metallicity (Ibeling & Heger, 2013; Woosley & Heger, 2015). They result from the star fusing H into He (on the main sequence, and later on the red giant branch), then fusing He into C and O (first on the horizontal branch, and later on the asymptotic giant branch), thereafter losing its outer layers to leave only a dense C–O core, representing ~99% of its total mass.

The remaining 1% is made of any H/He not fused during earlier evolutionary phases, arranged in a He envelope surrounding the core, and an outer H envelope, with their thicknesses depending on details of the nuclear burning. The upper layers of the H/He envelope constitute the atmosphere, the only region observable.

Some 20–25% of white dwarfs consume virtually all their H during previous evolutionary phases, resulting in H-deficient (and hence He-rich) atmospheres (Althaus et al., 2005). Furthermore, if the progenitor is close to the $10 M_\odot$ upper mass limit, it may reach temperatures high enough to fuse C, leading to an O–Ne white dwarf.

White dwarfs are classified accordingly: an initial D, and a letter describing the dominant (atmospheric) spectral feature: DA dwarfs have *atmospheres* dominated by H I, DB by He I, DC have a continuous spectrum, DO are dominated by He II, DQ by carbon, and DZ by metal lines. DA and DB are the most common.

GAIA DATA are used to define samples to 100 pc or so (essay 29), of which those to 20 pc (Hollands et al., 2018), 40 pc (Gentile Fusillo et al., 2021), and 100 pc (Vincent et al., 2020) are widely used. Gaia astrometry and photometry are then used to construct the HR diagram, to derive mass estimates, and so on. Let me also recall these points to frame the subsequent studies:

• masses of field white dwarfs are obtained only indirectly by combining surface fluxes (using distance and radius), spectroscopic $T_{\rm eff}$ and $\log g$, evolutionary models, and by appeal to the appropriate mass–radius relationship (e.g. Saumon et al., 2022, Eq. 1; essay 107);

• derived masses follow a narrow distribution peaking at $0.55-0.58M_\odot$, similar for DA and DB dwarfs (Bergeron et al., 2019, Fig. 13; Ourique et al., 2019, Fig. 2; McCleery et al., 2020, Fig. 7). 60% are in the range $0.5-0.7M_\odot$, but a few extend to $0.2-1.35M_\odot$ (Pelisoli & Vos, 2019);

• very low-mass white dwarfs, $M \lesssim 0.3M_\odot$ are believed to have formed by mass loss in binary systems (Iben & Livio, 1993). The 'ultra-massive' objects ($\gtrsim 1.1M_\odot$) are (generally) thought to have O–Ne cores (Siess, 2007);

• cooling is dominated by core neutrino emission for the first 20 Myr. Thereafter, over many Gyr, it is controlled by the thermal conductivity of the core and envelope, the radiative opacity of its atmosphere, element transport in its interior, and the passage through phase transitions;

• the relation between progenitor mass and final white dwarf mass (the 'initial-to-final mass relation', IFMR), must account for the various mass-loss processes, and hence provides a probe of the local star-formation and evolutionary history. It can be inferred, by population synthesis methods, from an initial stellar population matched to the observed white dwarf mass distribution (e.g. El-Badry et al., 2018; Cunningham et al., 2024b).

SUBSEQUENT studies agree that the gap cannot simply arise from different evolutionary tracks for DA and DB objects. The subsequent debate illustrates the complexity of the physics, and how Gaia is contributing.

El-Badry et al. (2018) noted that the distribution remains bimodal even if only spectroscopically-confirmed DA white dwarfs are considered. They found that an initial-to-final mass relation that flattens at around $3.5M_\odot$ (and which they attributed to a mixed-age stellar population) results in a bimodal white dwarf mass distribution with the normal peak at $0.58M_\odot$ and a secondary peak at $0.8M_\odot$. In this explanation, it is the large number of massive white dwarfs which leads to a secondary DA sequence offset below the primary DA sequence.

Kilic et al. (2018) also used a population synthesis approach to infer a significant contribution from more massive white dwarfs, but argued that while the El-Badry et al. (2018) hypothesis of a modified IFMR could explain the gap, it could not explain the different binary fractions found on the main-sequence and for white dwarfs. They argued instead that their significant contribution from relatively massive white dwarfs likely arose through mergers. Studies with somewhat similar conclusions were also reported by Jiménez-Esteban et al. (2018) and Gentile Fusillo et al. (2019).

A DIFFERENT explanation was given by Bergeron et al. (2019). They found that the pure He atmospheric models previously used do not properly describe the white dwarfs collectively assigned to the 'non DA' class.

Specifically, their theoretical models for the DB stars agreed with the observed sequence for $T_{\rm eff} > 11\,000$ K, but not at lower temperatures where non-DA stars exist both as He-dominated (DB) stars, but also as DZ, DQ, or DC objects. Assuming a pure He compositions for *all* non-DA stars gave a shift in absolute magnitude corresponding to a shift in mass in the $M_{\rm WD} - T_{\rm eff}$ diagram.

However, these high inferred masses would instead be interpreted (via evolutionary models) as more normal masses, of $\sim 0.6M_\odot$, with the inclusion of a small amount of hydrogen, and/or metals, in their atmosphere.

Similar arguments were presented by Serenelli et al. (2019) and Ourique et al. (2020), where the term 'spectral evolution' describes the evolution from H to He envelopes (Chen & Hansen, 2011; Chen & Hansen, 2012; Rolland et al., 2018; Ourique et al., 2019).

IF THIS INDEED represents a consensus interpretation, it can now be simply stated: to reproduce the observed split in the cooling sequence, and specifically to replicate the B branch, trace amounts of hydrogen and/or metals must be present in the He-dominated atmospheres of hydrogen-deficient (DB) white dwarfs.

Blouin et al. (2023) found that neither the convective mixing of residual hydrogen, nor the accretion of hydrogen (or metals) can be the dominant cause of the bifurcation. Rather, convective dredge-up, from the deep interior, of small quantities of carbon, below the threshold of optical detection, can account for the observations.

New evolutionary models including this effect in cold He-rich white dwarfs have recently been given by Camisassa et al. (2023). The presence of carbon in their atmospheres produces a continuum absorption favouring the emission at bluer wavelengths, so creating the B branch. The resulting mass distribution peaks around $0.6M_\odot$, consistent with standard evolutionary channels.

AND SO we can now conclude: Gaia's accurate placing of solar neighbourhood white dwarfs in the HR diagram has revealed structure not predicted by previous evolutionary models. But trace amounts of carbon, added by convective dredge-up, can explain the gap. And it makes a specific prediction: that traces of carbon should be seen, perhaps in ultraviolet spectra.

179. Stellar masses from SB2 binaries

MASSES ARE, of course, one of the most fundamental of stellar properties, crucial in determining their structure and evolution… and everything that follows from them. Yet ways of determining accurate star masses are strictly limited and, even today, only a couple of hundred are *measured* to better than 1–2%.

The few accurate masses that are available underpin the calibration of the model-dependent methods of isochrone or stellar track fitting, which employ models of stellar structure and evolution matched to the observed star properties (e.g. Lebreton & Reese, 2020).

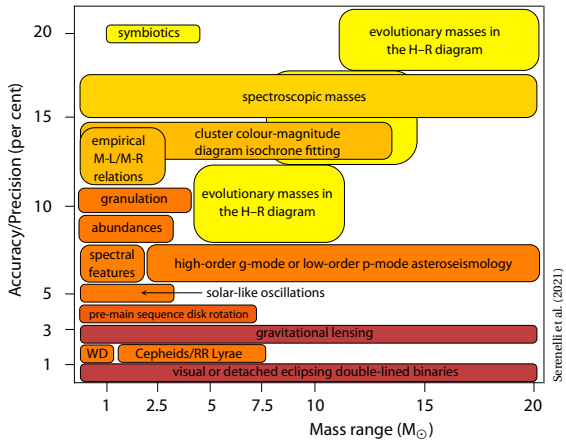
In their recent review, Serenelli et al. (2021) describe the various methods that can be used to estimate stellar masses across the Hertzsprung–Russell diagram. The most accurate, widely applicable and model-independent method employs Kepler's third law applied to detached (non-interacting) binary systems. In the absence of accurate astrometry to establish the orbit inclination, these have been mostly eclipsing systems.

Asteroseismology has opened a new, powerful, but nonetheless still model-dependent approach, while other methods can be applied to microlensed systems, or to stars in specific evolutionary stages, ranging from pre-main sequence to evolved giants and white dwarfs.

THE FIGURE shows their synopsis of these various methods, along with the applicable mass ranges and current accuracies (Serenelli et al., 2021). Methods are colour-coded, with a darker colour for increasingly model-*independent* methods, such that the darkest red regions deliver model-independent masses, which are not only precise, but importantly also *accurate*.

They list just 200 or so stars with relative mass accuracies between 0.3–2% over the mass range $0.1-16M_{\odot}$, 75% of which are main-sequence core H-burning, and the remainder cover all the other stages. This is somewhat better than the situation presented 30 years ago by Andersen (1991), at least in accuracy if not in number.

Here, I will look only at the most accurate (and model-independent) method of mass determination, viz. the use of double-lined spectroscopic binaries, and explain how Gaia is contributing.



IN THE ABSENCE OF MORE restricted techniques applicable to single stars (notably microlensing and asteroseismology) the determination of stellar masses relies on their gravitational effects in binary orbits, and it is useful to recall some basics.

In a non-interacting binary, each star moves in a closed elliptical orbit in inertial space, with the centre of mass at one focus. Such a Keplerian orbit in three dimensions is described by 7 parameters: $a, e, P, t_{\rm p}, i, \Omega, \omega$, where a and e specify the size and shape of the orbit, P is related to a and the component masses through Kepler's third law, and $t_{\rm p}$ is the position of the object along its orbit at a particular reference time. The three angles (i, Ω, ω) represent the projection of the true orbit into the observed (apparent) orbit; they depend solely on the orientation of the observer with respect to the orbit.

For single-lined spectroscopic binaries, in which only one star's spectrum is seen, radial velocity measures provide the mass function, $f = (M_2 \sin i)^3/(M_1+M_2)^2$. For double-lined spectroscopic binaries (i.e. with two distinct spectra), the mass function for both can be established, and hence the mass ratio, but still neither the individual masses, nor the orbit inclination.

From radial velocities alone, masses can only be determined unambiguously if the system is eclipsing, imposing an explicit value for the inclination, viz. $i \simeq 0$.

ECLIPSING SYSTEMS, of course, represent only a small subset of all binaries, and bias inferences to shorter periods. They can, in turn, be affected by mutual interactions including tidal distortion and mass transfer, so that only a minority of eclipsing binary components are truly representative of single stars (e.g. Popper, 1967).

The addition of astrometric measurements changes prospects significantly, because all seven orbit elements become accessible in principle. In practice, different considerations apply according to whether the system is a 'visual binary' or an unresolved 'astrometric binary', whether the measurements probe all or part of the orbital motion of one or both components, or only that of the photocentre, and whether the orbital solutions also take account of constraints from radial velocity data (e.g., Torres, 2004; Anguita-Aguero et al., 2022).

For example, an orbit solution from the astrometry of (resolved) visual binaries yields the orbit inclination, and hence individual masses (e.g. Serenelli et al., 2021, §2.3). Suitable astrometric measurements have not been easily obtained in the past, often calling for speckle or interferometric measurements over many years.

Over its 3-year mission, Hipparcos provided just 235 orbital solutions for unresolved astrometric binaries (Lindegren et al., 1997), of which mass ratios could be obtained for about 25 (Söderhjelm, 1999).

WITH THAT background, let me summarise the progress in mass determinations over the past four decades as exemplified by four major reviews.

Popper (1980) gave masses for about 100 eclipsing and visual binary components known to better than 20% (requiring parallaxes to better than 7%). He commented that *'It is rather sad, in view of the very great amount of difficult observing for more than 150 years, that the number of visual binaries for which masses are known to an accuracy of about 20% is not more than a dozen or so.'*

Andersen (1991) gave masses, again better than 20%, for just 45 detached double-lined eclipsing binary systems (90 single stars), covering spectral types O8–M1 on the main sequence, and with two red giants. Compared to Popper (1980), data for only 6 systems remained unchanged; improved data were given for 18 systems, and 21 systems were new additions.

Torres et al. (2010) listed 95 detached binaries containing 190 stars (comprising 94 eclipsing systems, and the astrometric binary α Cen) for which masses *and* radii were estimated to better than 2–3%, along with 23 other systems with accurate masses but less accurate radii. Their sample more than doubled that of Andersen (1991), and extended the mass range to $0.2 - 30 M_\odot$.

Serenelli et al. (2021) provide the most recent review of mass-determination methods. Their resulting compilation includes masses for 40 detached eclipsing binaries (80 stars) better than 2% (their Table 2), and for 36 visual binaries (72 stars) better than 3% (their Table 3).

NOW TO GAIA, where the binary and multiple star processing is highly complex, dependent on the wide range of systems (orbital period, magnitude difference, variability, etc.), and on the various combinations of data that can be used in the orbit solution (astrometry, photometry, and RVS velocities). Arenou et al. (2023) identified 800 000 binaries in DR3 with orbit or trend parameters, classified as astrometric, spectroscopic, and eclipsing, in various combinations (their Table 1).

Of these, 165 500 are astrometric solutions characterising the orbit of the *photocentre*, yielding the system's parallax and proper motion, the orbit inclination and its standard error (Halbwachs et al., 2023c, Eqs A6 and A20), and the 'astrometric mass function', which depends on the component fluxes, F_1 and F_2 (Arenou et al., 2023 Eq. 2; Halbwachs et al., 2023c Eq. 14). Again, the degeneracies can be broken for eclipsing binaries, or for double-lined (SB2) spectroscopic binaries.

I am not aware of any synthesis of the best masses available from DR3, but results include several astrometric orbits with $\sin^3 i$ better than 1%, with the best masses at the level of around 0.3% (Halbwachs et al., 2023b).

RADIAL VELOCITIES from Gaia's RVS spectrometer are not at the accuracies required for the best mass determinations, and state-of-the-art masses exploiting the Gaia data still largely also rely on ground-based radial velocity measurements. Thus Chevalier et al. (2023) combined DR3 astrometry with SB2 data from the Ninth Catalogue of Spectroscopic Binary Orbits (SB9), and APOGEE, to determine masses for 56 systems (43 from SB9, and 13 from APOGEE), and provided an empirical mass–luminosity relation down to $0.12 M_\odot$.

Similarly, ground-based campaigns are ongoing to acquire radial velocities (along with interferometric or speckle data in some cases) targetting masses at 1% accuracy when eventually combined with future Gaia astrometry. Amongst these are 70 binaries being observed with OHP–SOPHIE (Halbwachs et al., 2014; Kiefer et al., 2016; Halbwachs et al., 2020), and the follow-up of other suspected SB2s from Gaia (Halbwachs et al., 2023a).

Other works are using these fundamental masses, combined with theoretical stellar evolution models, to estimate stellar masses from observed luminosities, based on Gaia G-band magnitudes and stellar distances (e.g. Lebreton & Reese, 2020; Malkov et al., 2022; Chevalier et al., 2023; Eker et al., 2024; Pérez-Couto et al., 2024).

TO MY KNOWLEDGE, no results have been published yet from Gaia's resolved orbital binaries, for which masses should also be available, although I have no feeling for the numbers of objects involved. Data Release 4, in 2025, will also cover a longer time interval, and will include astrometry at each measurement epoch. Again, I can offer no useful insight into the expected state of stellar mass determination at the end of the Gaia mission.

180. The spectra of solar system objects

I OUTLINED Gaia's contribution to the study of solar system objects in two earlier essays. In essay 64 (March 2022) I described the main goals of these observations, the associated data processing, and the first haul of just 14 000 asteroids included in the DR2 release (covering observations from July 2014–May 2016).

Data Release 3, in June 2022 (see essay 76) was based on 34 months of observations (July 2014–May 2017), and provided astrometry for 157 000 solar system objects, orbits for 154 787, and BP/RP reflectance spectra (the focus of this essay) for 60 518. Orbit determination was described by Tanga et al. (2023), who also gave a summary of the associated photometric and spectral data.

In essay 159 (Jan 2024) I described the special solar system objects 'Focused Product Release'. This was an intermediate product between DR3 (June 2022) and DR4 (expected in late 2025), treating the same 157 000 asteroids as in DR3, but now providing the epoch astrometry and orbit reconstruction based on the much longer data interval of DR4, viz. 66 months compared to the 34 months of DR3 (David et al., 2023).

AS FORESEEN in the science case for Gaia in 2000, the observation of all objects brighter than 20–21 mag would yield a deep, uniform detection of minor planets and other solar system bodies, enabling detailed studies of their dynamics and taxonomy.

These small bodies preserve relatively unaltered materials that date back to the formation of the solar system from the proto-solar nebula 4.567 Gyr ago (e.g. Amelin et al., 2002; Johansen et al., 2015). And they provide insights into the early accretion of primordial material from the protoplanetary disk, the transport of water and other organic material to Earth (e.g. Morbidelli et al., 2000), and collisional events leading to asteroid 'families' (e.g. Nesvorný et al., 2002; Nesvorný et al., 2015).

Gaia's accurate multi-epoch astrometry also allows characterisation of binarity, e.g. for (4337) Arecibo and the Pluto–Charon system (Tanga et al., 2023), measurement of the Yarkovsky effect, the secular orbital drift due to the anisotropic emission of thermal photons, which is important in their migration (Tanga et al., 2023; Dziadura et al., 2023), and the identification of more complex dynamical resonances (Carruba et al., 2024).

WHILE THE EMPHASIS of my earlier essays was on the astrometry and resulting orbits, I will look here at the photometry obtained from Gaia's blue and red photometers, and the derived reflectance spectra published as part of DR3 (Galluccio et al., 2023). The multi-epoch multi-colour photometry had always been expected to provide information about their physical properties, including their shape and rotation (Cellino & Dell'Oro, 2012), and their composition and taxonomic classification (Delbo et al., 2012; Klimczak et al., 2022).

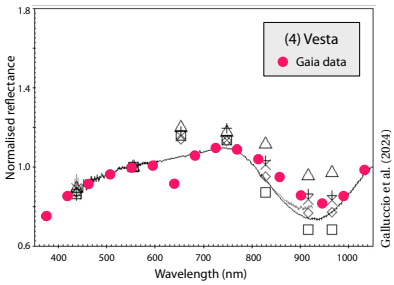
One of the reasons why these sorts of observations are of great interest is that the asteroid population does not remain unchanged over time, and various physical processes play a role in their evolution (e.g. Galluccio et al., 2023). These include collisional evolution (which affects the number and size distribution of the main belt asteroids as well as their surface structure); space weathering (due to surface irradiation from cosmic rays, the solar wind, and micro-meteorites, which all modify their reflectance spectra); and the evolution of the regolith structure and properties due to the continuous thermal cycling of their surfaces.

SUCH CONSIDERATIONS have driven a dozen asteroid flyby missions, as well as numerous ground-based spectrophotometric surveys. Amongst the earliest of these were an 8-colour survey by Zellner et al. (1985), a 52-colour survey by Bell et al. (1988), a 7-colour infrared survey by Clark et al. (1993), and numerous others since. Galluccio et al. (2023) estimate that more than 1.5 million spectrophotometric observations of asteroids exist, covering more than 7600 asteroid spectra.

Amongst Gaia's advantages are its accurate multi-epoch multi-colour photometry, its extensive sky coverage (solar elongations $45 - 135°$), the large numbers of objects (expected to total around 350 000), and its broad wavelength range (covering 400–1000 nm).

GAIA'S MULTI-COLOUR photometry, designated BP and RP, is achieved by two fused-silica prisms which disperse the spectra over ~45 pixels in the along-scan direction (see essay 68). The downstream processing of the resulting low-resolution spectra is detailed by De Angeli et al. (2023).

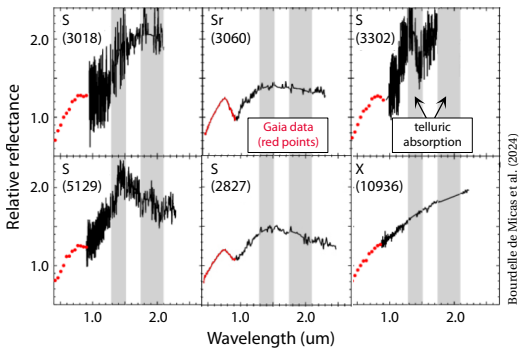
The reflectance spectrum at each epoch was derived by dividing the object spectrum by a solar-type composite (Galluccio et al., 2023). A mean reflectance spectrum was then calculated in 16 spectral bands. Comparison with ground-based observations under similar illumination geometry show very good agreement across taxonomic class, their examples including (1) Ceres, (4) Vesta (shown here), (21) Lutetia and (433) Eros.



NUMEROUS STUDIES are making use of the Gaia reflectance spectra, and I will give a few examples.

Asteroid 'families' result from collisions, some of which are ancient, others being much more recent (e.g. Nesvorný et al., 2002). Collisional fragments can be ejected at moderate velocity (~1 km s^{-1}) and, at least in the main asteroid belt, can remain clustered in orbital space and with similar physical properties.

Bourdelle de Micas et al. (2024) studied 263 fragments of an S-type asteroid recently discovered in the inner main belt, with an age of 4.4 Gyr, and comprising both planetesimals and collisional fragments (Ferrone et al., 2023). The figure below shows six of their break-up candidates. The combination of existing and newly-measured spectra extending to the near infrared, combined with Gaia reflectance spectra below 1 μm (shown here in red), allowed identification of 71 (non-S-type) interlopers, and a final list of 190 S-type members.



THE LOW ALBEDO inner belt C-type (carbonaceous) asteroids, of which there are 11 known families, have been suggested as the sources of the near-Earth asteroids Ryugu (visited by Hayabusa 2 in 2018) and Bennu (by OSIRIS-REx the same year). Delbo et al. (2023) used the Gaia reflectance spectra to characterise these families, showing that the Polana and Eulalia groups can be distinguished in the range 370–500 nm, and that the average spectra of the Eulalia and Polana families are most similar to those of Bennu and Ryugu respectively.

IN STUDIES of space weathering, Galluccio et al. (2023) found 21 909 S-type (siliceous or 'stony') asteroids in the Gaia reflectance sample, of which 9225 belong to known families (Nesvorný et al., 2015). Some have ages estimated by backward orbit integration to identify their convergent origin. They showed that the spectral slope increases, and the depth of the 1 μm absorption band decreases, with increasing age (their Fig. 20), which they attribute to the progressive effects of space 'weather'.

The rare inner-belt T-type asteroid (596) Scheila underwent a collision in December 2010 (Ishiguro et al., 2011). Hasegawa et al. (2022) used the Gaia reflectance spectrum (amongst others) to place limits on the effects of surface weathering over a time span of just 10 years.

ANDESITE IS a type of volcanic rock, named after the Andes where it was first identified. The andesitic meteorite Erg Chech 002 was found in the Sahara desert in 2020. With a crystallisation age of 4.565 Gyr, 2.25 Myr after the solar system origin (Barrat et al., 2021), it preserves a record of volcanic events and crust formation in primordial planetesimals, although no asteroid analogue had been identified. Galinier et al. (2023) searched for analogues in the Gaia DR3 reflectance spectra, finding 51 similar main-belt asteroids, and 91 others with the spectrum of a space-weathered analogue. Near-infrared spectra would provide a more definitive association.

The V-type (basaltic) asteroids are also important in understanding planetesimal formation and evolution in the early solar system. Oszkiewicz et al. (2023) identified some 2000 possible V-type asteroids from their Gaia reflectance spectra, increasing the known number by more than a factor three.

Other studies relate to the understanding of small binary asteroids, believed to form through fission induced by Yarkovsky-driven spin-up, depending on their structural strength and therefore composition (Minker & Carry, 2023); to planet-crossing asteroids (Sergeyev et al., 2023); and discovery of the first olivine-dominated A-type asteroid family (Galinier et al., 2024).

THE DR3 results give an indication of what is to come. Gaia Data Release 4 will contain more objects, and with twice the temporal coverage than for DR3.

181. The Yarkovsky effect

THE YARKOVSKY EFFECT is an important phenomenon in solar system dynamics. It appears in discussions of the astrometry and photometry of the small solar system objects which are part of Gaia's global observing programme (essays 64 and 159). I touched on it in my previous essay on the reflectance spectra of more than 60 000 objects released as part of Gaia DR3 (essay 180).

Here, I will say more about the effect, why it is important, and how Gaia is contributing to its understanding.

THE ORBITS of all solar system bodies are controlled by gravitational forces, and in particular by the Sun. Today's state-of-the-art positions and velocities of the Sun, Earth, Moon, and planets (INPOP10e prepared by the Observatoire de Paris, and DE430 by Caltech's JPL) are based on a numerically integrated dynamical model which, for DE430, also includes the perturbative effects of 343 asteroids, viz. 90% of the mass of the main belt.

Solar radiation also affects an object's orbital motion. Solar radiation *pressure*, imparted by photon momentum, and quantified by Petr Lebedev in 1899, is a tiny force, but one with a large cumulative effect over long periods of time, affecting the motion of small solar system bodies (including spacecraft such as Gaia!).

The Poynting–Robertson effect, described by J. H. Poynting in 1903 and H. P. Robertson in 1937, only affects particles smaller than about 1 mm. A particle's motion around the Sun, combined vectorially with the velocity of the incident radiation (as in stellar aberration), results in the particle losing angular momentum and, for small particle sizes, spiralling inwards towards the Sun.

The Yarkovsky effect is a force acting on a *rotating* body, was described by Polish engineer Ivan Yarkovsky in 1901, and by Ernst Öpik in 1951 (see Beekman, 2005). It results from the re-radiated thermal emission, which carries momentum, lagging behind the incident radiation in time, thus (perhaps non-intuitively) contributing a component of force tangent to the orbital motion.

The Yarkovsky–O'Keefe–Radzievskii–Paddack/YORP is a second-order effect, mainly affecting the body's spin. I will look at its importance for Gaia in essay 182.

THE MAGNITUDE of the Yarkovsky effect is dependent on the orbit size and eccentricity, and on its mass, size, shape, spin and composition (e.g. Bottke et al., 2006). But qualitatively, the Yarkovsky force is in the direction of orbital motion for a prograde rotator, causing the semi-major axis to increase steadily. Conversely, a retrograde rotator spirals inwards, with the Yarkovsky and Poynting–Robertson 'drags' working together.

Models supplement the standard dynamical model with solar radiation pressure as an extra radial force, A_1, and a transverse 'Yarkovsky' force, A_2, both inversely proportional to the heliocentric radius. Coefficients are determined by a model fit to the observations (Vokrouhlický, 1999; Farnocchia et al., 2013; Vokrouhlický et al., 2015a; Del Vigna et al., 2018; Fenucci et al., 2024).

For an object 100–1000 m in size, the effect is all-but-negligible even over many years, but it is relentless: over millions of years an object can be transported from the asteroid belt to the inner solar system. As a result, though tiny and difficult to detect, it is important in understanding various aspects of solar system dynamics.

FIRST observational confirmation of the Yarkovsky effect came from radar tracking of the half-kilometer size asteroid (6489) Golevka, from Arecibo, between 1991–2003 (Chesley et al., 2003). Over that time, and as predicted qualitatively by Vokrouhlický et al. (2000), it was 15 km from the position predicted from gravitational interactions with other solar system bodies alone.

Chesley et al. (2003) estimated a Yarkovsky force of 0.25 N, and acceleration $10^{-12}\,\mathrm{m\,s^{-2}}$. The accompanying press release described this as the *'equivalent of one ounce of thrust pushing against a 0.5 km, 210 million-ton tumbling "mountain".'*

The use of radar tracking in this context is noteworthy because, pre-Gaia, it was the most powerful technique for determining a 'nearby' asteroid's orbit.

The second detection was the 400-m diameter asteroid (152563) 1992 BF from astrometric measurements over 50 years, including precovery observations from 1953 (Chesley et al., 2006; Vokrouhlický et al., 2008).

THE YARKOVSKY EFFECT complicates orbit determination for near-Earth asteroids, occasionally frustrating their identification (e.g. Vokrouhlický et al., 2008), and affects the prediction of near-Earth approaches and impact probabilities (e.g. Giorgini et al., 2002; Milani et al., 2009; Vokrouhlický et al., 2015b).

But it also appears to play a much deeper role in solar system dynamics. It could explain the origin of high-eccentricity meteorites (Peterson, 1976; Afonso et al., 1995), and perhaps NEOs more generally, by delivering main belt asteroids up to 20 km in size into Earth-crossing orbits (Morbidelli & Vokrouhlický, 2003).

It also drives an additional orbital dispersion of asteroid 'families' beyond the effects of collisions alone, leading to their characteristic V-shape in diagrams of semi-major axis versus absolute magnitude (or $a - 1/D$), which is important for their age estimation (e.g. Bottke et al., 2001; Bottke et al., 2002; Milani et al., 2014; Spoto et al., 2015; Milić Žitnik, 2020; Novaković et al., 2022).

As a result of its role in the YORP effect, where it changes the spin rates and axes of the smaller irregular asteroids, it could also explain the large number with very high and very low rotation rates (Rubincam, 2000; Kaasalainen et al., 2007; Lowry et al., 2007).

Finally, NEAs are generally of relatively low mass, and the Yarkovsky effect provides the possibility of mass determinations, providing an important constraint on the body's internal structure and composition (Chesley et al., 2003). This method is distinct from masses derived from mutual orbit perturbations of some of the more massive objects (e.g. Bange, 1998; Viateau & Rapaport, 1998; Mouret et al., 2007; Podlewska-Gaca et al., 2020).

ACCORDING TO Tanga et al. (2023), the JPL Small-Body Database includes measurements of the Yarkovsky effect for 234 asteroids. All are NEOs, for which small objects and accurate orbits dominate. Some estimates predate Gaia (e.g. Greenberg et al., 2020), while others are from Gaia alone or combined with other observations.

EARLY GAIA studies suggested that the effect might be detectable in several tens of NEOs (Tanga et al., 2007; Delbo et al., 2008; Dziadura et al., 2022), and 64 promising candidates were listed by Mouret & Mignard (2011). Detection models are given by Tsiganis et al. (2012), Desmars (2015), and Fenucci et al. (2024).

Results reported from Gaia are often in combination with other observations, mainly long-term astrometry or radar ranging. Thus Hanuš et al. (2018) used Arecibo and DR2 astrometry of (3200) Phaethon to constrain the secular drift of its orbital semi-major axis, yielding the bulk density assuming that the drift is Yarkovsky dominated. Results are typical of large C-complex asteroids, and support its association with asteroid (2) Pallas, as suggested from their dynamics (Todorović, 2018).

With a target list chosen with prospects for detectability in mind, Greenstreet et al. (2019) used Gaia DR1 with the Las Cumbres ground network to detect the Yarkovsky effect in 18 out of 36 observed asteroids.

Gaia DR3 contains orbits for 154 787 solar system objects. Orbit post-fit residuals have a standard deviation of some 5 milli-arcsec along scan, a factor 100 better than ground-based values. Tanga et al. (2023) identified 447 NEOs in DR3, of which 24 had a previous measurement of the Yarkovsky effect, mostly from radar data.

They discussed two specific cases. For (3200) Phaethon, the parent body of the Geminid meteorite shower, they found a value for the Yarkovsky term when combining ground-based with Gaia DR3 astrometry within 1σ of the JPL Small-Body Database value derived from ground-based astrometry combined with radar observations. For (1620) Geographos they reported a first 3σ detection of the Yarkovsky effect using 5242 ground-based optical observations, 7 radar observations, and 105 observations from Gaia DR3. More generally, combining the accurate astrometry from Gaia DR3 with the existing observations from the Minor Planet Center significantly improves the asteroid orbit, and enhances the detectability of the non-gravitational terms.

Dziadura et al. (2023) used the Minor Planet Center's OrbFit software, along with their complete astrometric data set (which includes all radar and Gaia DR3 data) to derive the orbits of 446 Near-Earth Asteroids (including 93 Potentially Hazardous Asteroids, 54 094 inner main belt asteroids, and various Mars crossing asteroids. They determined a significant and improved Yarkovsky term (and associated bulk densities) for 49 Near-Earth Asteroids, including 10 new detections.

No Yarkovsky acceleration term was detected in the main-belt asteroid study by Dziadura et al. (2023), with a similar non-result for 134 MBA objects observed with the University of Hawai'i 88-inch telescope, supplemented with the observations from the Minor Planet Center, including Gaia DR3 (Hung et al., 2023).

DETECTION DIFFICULTY for main-belt asteroids notwithstanding, one notable asteroid without a measured Yarkovsky term today is (35334) Yarkovsky!

THE FUTURE will see even deeper asteroid surveys, including ESA's Flyeye, the Vera Rubin Observatory, and possibly the space-based NEO Surveyor and NEOMIR, together likely increasing the known asteroids by an order of magnitude (Jones et al., 2018).

While improved Gaia astrometry will come with DR4 in 2025, I believe that the expected orbit improvements already appear in the Focused Product Release reported by David et al. (2023). This covers 66-months of Gaia mission data, compared with the 34 months of DR3. I am not aware of any Yarkovsky studies based on it so far.

182. The YORP effect

In essay 181, I looked at the effect of solar radiation on the orbits of minor solar system bodies. To recap, depending on orbit, size, spin rate and composition, solar radiation can 'push' an object away from the Sun through radiation *pressure* or, for particles less than a few mm, act as a 'headwind', slowing its orbital speed, and sending it inwards (the Poynting–Robertson effect).

For a rotating body, the incident solar flux can lead to re-radiated but delayed thermal emission, sending the object outwards or inwards depending on whether its rotation is prograde or retrograde (the Yarkovsky effect). The momentum transfer is minute, but can result in significant orbit evolution over tens of millions of years.

The YORP effect (Yarkovsky–O'Keefe–Radzievskii–Paddack) is a second-order effect, influencing the spin rate and spin axis orientation of small *irregular* asteroids. Like the Yarkovsky force, its effects accumulate over long periods, resulting in some rather remarkable changes in the properties of the asteroid population.

I will describe what YORP is, expand more on its scientific importance, and what Gaia has to say about it.

The term acknowledges four contributions to the understanding of how solar electromagnetic radiation, through diffuse reflection, absorption, and thermal re-emission, changes a body's angular momentum relative to its centre of mass, and dependent on its detailed properties such as shape and albedo (Rubincam, 2000).

Specifically, the term encapsulates Yarkovsky's realisation (1901) that the thermal radiation carries momentum as well as heat, Radzievskii's suggestion (1954) of rotation driven by changes in albedo, Paddack's realisation that shape was more effective in altering a body's spin rate, and Paddack and O'Keefe's suggestion that the various effects can lead to rotational 'bursting', with small asymmetric bodies eventually reduced to dust (Paddack, 1969; Paddack & Rhee, 1975.

With asteroids being irregular in shape, and having rotation periods of order days, being much shorter than their orbital periods, the YORP effect is the secular change in its rotation state after averaging the solar radiation torques over the body's spin and orbital periods.

The YORP effect is presumed responsible for the many asteroids with very high and very low spin. Those with diameter $\gtrsim 125$ km follow a Maxwell distribution, while 50–125 km sizes show an excess of fast rotators. The smallest, $\lesssim 50$ km, have an excess of very fast and slow rotators, becoming more pronounced with decreasing size (Rubincam, 2000; Kaasalainen et al., 2007).

Direct confirmation of the YORP effect came with optical observations of the 100-m diameter asteroid 2000 PH5, later designated (54509) YORP. Over 4 years, the 12-min rotation period showed a continuously increasing spin rate of $2.0 \pm 0.2 \times 10^{-4}$ deg d^{-2} which could not be explained by gravitational torques (Lowry et al., 2007; Taylor et al., 2007). A radar image and 3d model is given at the object's wiki page.

Similar results were found for the 1-km sized asteroid (1862) Apollo, where the change is clearly visible in the long-term light curve, amounting to an extra rotation in just 40 yr (Kaasalainen et al., 2007; Durech et al., 2024). Predictions for (25143) Itokawa were made by Vokrouhlický et al. (2004) and Scheeres et al. (2007), and duly confirmed by Kitazato et al. (2007), although sensitive to its assumed surface topography (Breiter et al., 2009; Statler, 2009; Lowry et al., 2014; Walsh, 2018).

Today, the YORP effect has been identified individually for just 12 asteroids (Durech et al., 2008; Durech et al., 2022; Durech et al., 2024, Table 1). Gaia DR3 has contributed no others, and DR4 is also likely to be of too short a duration for additional detections.

But there is *indirect* evidence for its role in the orbital evolution of asteroids over long periods, most prominently in the clustering of the directions of rotation axes in asteroid families (Vokrouhlický & Čapek, 2002; Bottke et al., 2006; Vokrouhlický et al., 2006).

Extreme YORP-driven spin-up may also explain asteroid fragmentation (Paddack, 1969; Paddack & Rhee, 1975; Veras & Scheeres, 2020), including the observed breakup of asteroid P/2013 R3 (Jewitt et al., 2014).

It may also be important in the formation of binary asteroids (Walsh et al., 2008), and asteroid pairs with very similar orbits (Vokrouhlický & Nesvorný, 2008).

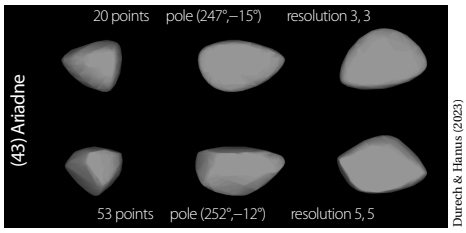
ALL THE WORK that I have described so far pre-dates any use of the Gaia data. I will now look at what insights Gaia is contributing in the area of the YORP effect, over and above the results on the Yarkovsky effect which I considered in essay 181. That is, I will focus on the use of Gaia photometry for studies of asteroid rotation.

In fact, the following results come mainly from a study by Durech & Hanuš (2023) on the 'Reconstruction of Asteroid Spin States from Gaia DR3 Photometry'.

I WILL FIRST recall the contents of Gaia DR3 most relevant for this particular study. DR3 includes accurate photometry for 154 787 solar system bodies, covering a time interval of 34 months, between July 2014–May 2017 (essay 76). With some 3 million measurements in total, the number of photometric observations per asteroid ranges from a few to several tens. And, remarkably, their spins can be determined in many cases.

To estimate the spin state and shape of each asteroid, Durech & Hanuš (2023) computed the illumination geometry for each observation, and used light-curve inversion to find the best-fit physical model. This was parameterised by the sidereal rotation period, the spin axis direction, and a low-resolution convex shape.

To find the best-fit model, they tried tens of thousands of trial periods in the range 2–10 000 h, with tens of trial pole directions. To establish the correct rotation period, they also used a triaxial ellipsoid model as an approximation to the asteroid shape. Some examples for asteroid (43) Ariadne are shown here (from their Fig. 3).



For more than 8600 of the original 150 000 objects, they were able to determine the spin state, together with a low-resolution convex shape model. Their results are limited to objects for which they found just two solutions for the rotation poles, (λ_1, β_1) and (λ_2, β_2) in ecliptic coordinates, and for which the two differed by less than some (small) selected values of $\Delta\lambda$ and $\Delta\beta$.

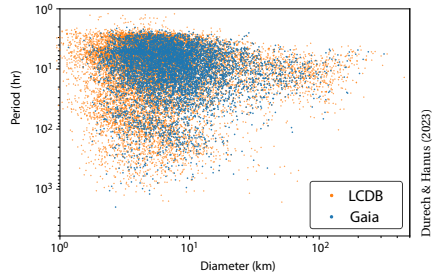
Their first 10 objects are tabulated here (an extract of their Table 3). It gives the two pole solutions, the identified period, and the number of observations used in the solution. Remarkably, the uncertainty on the rotation period is on the order of the last decimal place.

They also include a column, omitted here, giving the method used for computing their periodograms: C (convex shape models), E (ellipsoidal models), and CE (where both methods gave the same period).

	Asteroid	λ_1 [deg]	β_1 [deg]	λ_2 [deg]	β_2 [deg]	P [h]	N_{obs}
5	Astraea	122	37	312	40	16.8008	36
26	Proserpina	83	–46	255	–56	13.1092	58
32	Pomona	103	39	279	41	9.4475	38
33	Polyhymnia	20	–29	197	–31	18.6088	36
41	Daphne	207	–35	355	–39	5.98807	35
43	Ariadne	69	–10	252	–12	5.76182	53
44	Nysa	96	45	287	54	6.4214	25
45	Eugenia	123	–32	299	–20	5.69918	30
48	Doris	107	32	291	45	11.8900	50
50	Virginia	103	12	281	26	14.3107	31

THIS LARGE SAMPLE of new asteroid models compares very favourably with previous results based on a similar modelling approach. Such results include 18 solutions, of which 10 were new, from BlueEye600 (Durech et al., 2018b); 173 solutions, 129 new, from Gaia DR2 (Durech & Hanuš, 2018); 1100 solutions, 762 new, from Gaia DR2 and Lowell (Durech et al., 2019); 900 solutions, 662 new, from Lowell and WISE (Durech et al., 2018a); and 2750 solutions, of which there were about 1800 new, from ATLAS photometry (Asteroid Terrestrial-impact Last Alert System; Durech et al., 2020).

Gaia DR3 has increased by more than a factor two the numbers of asteroids with known spin states and rotation poles (viz. 8600 compared with 3500 previously listed in the DAMIT data base; Durech et al., 2010), with a significant advance in the study of their spin properties. One of their figures, included here, shows the distribution of period versus diameter for the Gaia DR3 solutions, and those from the Asteroid Lightcurve Database (LCDB, Warner et al., 2009), as of December 2021.



THE RESULTS of Durech & Hanuš (2023) all go in the direction of confirming previous findings, but with much greater confidence. And they give a foretaste for the larger samples that will be available with Gaia DR4.

First, small asteroids have poles clustered toward the ecliptic poles, an effect attributed to, and well-explained by, the YORP-induced spin evolution. Second, asteroid migration due to the Yarkovsky effect depends on the spin orientation. Third, the sense of rotation of members of asteroid families are correlated with their semi-major axis. Thus, over the age of the family, orbits of prograde rotators have evolved, due to the Yarkovsky effect, to larger semi-major axes, while those of retrograde rotators have drifted in the opposite direction.

183. Gaia CCDs, CTE, and solar activity

THE DETECTORS used across Gaia's astrometric, photometric (BP/RP), and radial velocity spectrometer (RVS) instruments are CCDs. And while CCDs are ubiquitous in consumer and professional imaging systems, and widely used in space, Gaia presented a number of new and demanding requirements, which in turn required an extensive development program pre-launch.

In 'normal' CCD imaging, data readout is performed *after* terminating an exposure. Then, all CCD rows are successively shifted down one row, with a pause after each shift as the latest row is 'clocked out' through a horizontal shift register and associated output amplifier.

In the 'time-delay integration' (TDI, or 'drift scanning') mode of operation, there are no discrete exposures: star images drift continuously across the focal plane of a telescope. The CCD rows are shifted to match the image drift rate as the exposure builds up. This type of scanning has been used, e.g. by Spacewatch between 1984–2011, and by the Sloan Digital Sky Survey Telescope since 1999. With Gaia, instead, the telescope scans the sky, and rows are shifted to match the spin rate.

THERE ARE 106 CCDS in Gaia's combined focal plane (including 4 for metrology). The TDI read-out rate is fixed, and the fields of view follow the pre-defined scanning law by continuous adjustment of the satellite's rotation using micro-Newton thrusters. This avoids the overheads of the usual 'point-and-stare' type operation, and results in a fixed exposure time (of 4.4 s per CCD).

A substantial challenge in implementing this observation mode for Gaia was in optimising the fraction of the full CCD that could be read out and sent to the ground. Reading out all CCD pixels (as is done for Gaia's 'sky mappers') would have demanded a much higher readout rate (and associated higher readout 'noise'), and a data volume (and resulting data rate) far too large to be telemetered from its L2 operational location.

The adopted solution is a complex 'windowing strategy', where only small sky regions surrounding each detected object image are read out and sent to the ground (others are flushed at high speed in the readout register).

A FAR BIGGER challenge was the stringent requirement on the accuracy of the estimation of the location of each stellar image in the along-scan direction. For example, a parallax accuracy of 30 μas for a 15 mag G2V star after 5 years requires a highly demanding location accuracy of 0.005 pixels at each observation epoch.

Accordingly, the pre-launch CCD development program addressed issues such as their metrology and thermal stability, optimisation of the readout noise, minimisation of amplifier cross-talk, and so on. But there were other highly critical issues that required specific developments of the CCD themselves.

Indeed, CCD91–72 was designed and manufactured specifically for Gaia by the UK company e2v (since 2017, Teledyne e2v), and they considered it their most complex to date. I will touch here only on a few key design features and properties, out of an enormous body of industrial specifications, laboratory tests, in-orbit measurements, and on-ground software calibration efforts.

THE GAIA CCDS are 4-phase, back-illuminated devices, with 4500 × 1966 pixels, each $10 \times 30\,\mu\mathrm{m}^2$. The full-well capacity is ~190 000 electrons, the readout noise is 3–5 e^- rms, and the operational temperature of 163 K was selected as a compromise to minimise both dark current and the effects of radiation damage (Prusti et al., 2016; §3.3.2). While all used the identical architecture, different anti-reflection coatings and resistivities were used for the BP and RP instruments.

Charge overflow in the case of bright stars has been a well-known problem for CCDs since their first use in astronomy in the late 1970s. 'Anti-blooming' techniques were further developed for Gaia and, in particular, included the use of a dedicated anti-blooming 'drain'. This prevents excess charge bleeding down columns (i.e. along scan) from bright stars, thus allowing the simultaneous measurement of faint stars in their vicinity.

Observations of the brightest stars ($G \lesssim 13$) make use of 'gates' (at 4500, 2900, 2048… 16, 8, 4, and 2 TDI lines) to restrict integration times, according to the sky mapper magnitude estimates (not all are used in practice).

RADIATION DAMAGE is a phenomenon which has adversely affected satellites since the dawn of the space age. It was responsible for many problems and the ultimate demise of Hipparcos in 1993, stranded in its geostationary transfer orbit and passing through the Earth's Van Allen radiation belts twice per day. Today, satellites routinely make use of radiation-hardened electronics, radiation shielding, and redundant units.

In its operational L2 orbit, Gaia is exposed to Galactic cosmic rays, but it is the sub-atomic particles of the solar wind (electrons, neutrons, and in particular solar protons) that dominate the radiation environment, especially from coronal mass ejections or flares during periods of enhanced solar activity. With energies up to several MeV, early predictions of the effects for Gaia were based on data from the IMP and OGO spacecraft between 1963–91 (Feynman et al., 1993).

HIGH-ENERGY PARTICLES also have a potentially disastrous effect on the CCDs themselves. Here, non-ionising processes, quantified by the Non-Ionising Energy Loss (or NIEL), can result in long-term cumulative effects by creating 'displacement damage' due to the displacement of atoms in the CCD crystalline substrate. The physics is complex, with the effects of vacancies, diffusion, and both doping and impurity atoms contributing to the creation of bulk 'traps', of different energies. These traps lead to a decrease in the 'charge transfer efficiency' (CTE), or equivalently to an increase in charge transfer *inefficiency* (CTI), as the CCD is read out. Some early analyses were made, for example, in the case of the HST–WFPC2 (e.g. Holtzman et al., 1995).

The problem for Gaia is that the signal carriers (electrons) can be captured stochastically, then re-emitted at later times (on time-scales of μs to several seconds), resulting in distortion in the shape of the charge packet, with an associated positional bias and potentially significant loss of accuracy. Effects depend on the location and type of traps (which accumulate over time), on the charge packet size (i.e. star magnitude), and on the previous illumination history of that part of the CCD. Modelling efforts are founded on the Shockley–Read–Hall formalism (Shockley & Read, 1952; Hall, 1952).

The effect was considered to be potentially catastrophic for Gaia, and a substantial effort was devoted to its characterisation and mitigation.

THE LAUNCH of Gaia was originally planned for 2012, and eventually took place at the end of 2013. With solar cycle 24 expected to start in 2008–10, and with a predicted maximum expected in 2013, radiation effects on all satellite subsystems were taken very seriously. Unhelpfully, pre-2006 predictions varied between *'the smallest solar cycle in 100 years'*, and *'the most intense cycles since record-keeping began'*.

PROGRESS INVOLVED many studies and laboratory tests, numerous meetings with the CCD manufacturer e2v, radiation studies by the electro-optical company Sira UK (e.g. Hopkinson et al., 2010), and in-depth commitment by the industrial prime contractor EADS Astrium (now Airbus; e.g. Laborie et al., 2007).

One specific feature built into the Gaia CCDs was the inclusion of a 'supplementary buried channel', in addition to the standard 'buried channel' structure below the silicon surface (where the charges collect). This confines the charge packets to a smaller silicon volume, reducing the number of traps with which the signal interacts. The design goal was for a supplementary buried channel accommodating up to 3000 electrons, although the reality was considerably more complex (Kohley et al., 2009; Seabroke et al., 2010; Seabroke et al., 2013).

Radiation damage effects on the astrometric accuracies were eventually mitigated through a combination of extra spacecraft shielding; the periodic injection of 'sacrificial charges' to occupy traps and so avoid them capturing signal charges; and innovative treatment in the on-ground processing (Short et al., 2010; Prod'homme et al., 2011; Prod'homme et al., 2012; Holl et al., 2012; Short et al., 2013; Lindegren et al., 2021, §3.3).

BUT FORTUNE WAS ALSO on our side. Solar activity in cycle 24 was minimal until early 2010, reaching a maximum in April 2014 at a value substantially lower than other recent solar cycles, and *'unseen since cycles 12–15'* (1878–1923). Solar cycle 25 began in 2019, with early indications suggesting low activity comparable to cycle 24, together perhaps indicative of a possible 'modern Gleissberg minimum' (Upton & Hathaway, 2018).

As part of the detailed calibration of the astrometric solution (Lindegren et al., 2012), the degradation in charge transfer efficiency in the scan direction is evident (e.g. Lindegren et al., 2021, §6.2), but at an order of magnitude less than predicted pre-launch. The CTI in the serial register is still dominated by traps inherent in the manufacturing process, with radiation-induced degradation of only a few per cent after 3 years (Crowley et al., 2016). The situation had not degraded significantly further even after 6 years in orbit (Ahmed et al., 2022).

Incidentally, two 'prompt particle events' in Gaia's sky mappers have also been identified with astronomical sources: Cyg X–1, and the γ-ray burst GRB221009A (see Gaia's Image of the Week, 9 November 2022).

AFTER 10 years in orbit, all CCDs still continue to operate flawlessly (although with a failure of the AF1 VPU/PEM electronics on 15 May 2024). These industrial achievements are a tribute both to e2v (led by David Morris), and to the prime contractor EADS Astrium, where the focal plane development was led by Anouk Laborie, and the CCD development by Cyril Vetel.

184. Stellar streams and sub-halos

AN INCREASING NUMBER OF accreted stellar structures are being identified in our Galaxy's halo. Some are attributed to captured dwarf galaxies, others to disrupted globular clusters, some of still uncertain origin. With 20 stellar streams tabulated by Grillmair & Carlin (2016), to around 100 known today (e.g. Mateu, 2023), some were found from the Sloan Digital Sky Survey, others from LAMOST and the Dark Energy Survey, with the majority of the most recent being identified by Gaia.

Members of these stellar streams are characterised by their common orbital motion, and generally very low metallicity. Search algorithms include `Streamfinder`, which determines stream membership probabilities based on the similarity of their orbits with those of their neighbours (Malhan & Ibata, 2018), and `StarGo`, which searches for streams and sub-structures clustered in dynamical space using neural-networks (Yuan et al., 2019).

As I noted in essay 156, analyses of these streams is enabling substantial progress in reconstructing the assembly history of the Milky Way, and in inferring the shape and mass of its dark matter halo.

A specific application that I will look at here is the use of these stellar streams in placing constraints on the existence and nature of the numerous dark matter sub-halos (halos within halos) that are predicted, in the standard ΛCDM cosmology, to exist surrounding the Milky Way (e.g. Springel et al., 2008; Zavala & Frenk, 2019).

An absence of such sub-halo driven perturbations in the Gaia data might be a challenge for standard ΛCDM.

IT IS NOW well established, observationally as well as in N-body simulations, that satellite galaxies interacting and merging with the Milky Way can result in tidal heating of the disk, in tilts and warps, and can trigger the growth of asymmetric structures such as the central bar, and the Galaxy phase-space spiral (Antoja et al., 2018).

Lower mass sub-halos, $\lesssim 10^9 M_\odot$, being largely devoid of gas and dust, are far more challenging to detect. The morphology and flux-ratios of strongly-lensed quasars currently suggests consistency with ΛCDM, albeit at low statistical significance (Ritondale et al., 2019; Hsueh et al., 2020; O'Riordan & Vegetti, 2024).

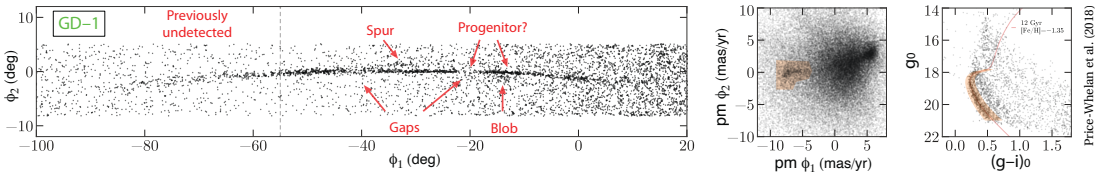
VARIOUS OTHER WAYS OF demonstrating the existence of dark sub-halos have been suggested: these include searches for γ-ray emission as a result of dark matter annihilation (Lake, 1990; Calcáneo-Roldán & Moore, 2000; Diemand et al., 2007), or through detecting the coherent vertical velocities of disk stars attributable to the passage of such sub-structures through the Galactic disk (Feldmann & Spolyar, 2015), or from the gravitational scattering of stars in tidal streams (Ibata et al., 2002b; Johnston et al., 2002; Siegal-Gaskins & Valluri, 2008).

In the latter, the underlying idea is that a narrow, dynamically cold stellar stream is susceptible to heating by repeated close encounters with the massive dark sub-halos, resulting in characteristic features such as gaps whose details depend on the sub-halo mass and distance from the progenitor (e.g. Sanders et al., 2016; Bovy et al., 2017), and which were considered detectable with Gaia (Ibata et al., 2002b; Feldmann & Spolyar, 2015).

As graphically described by Erkal & Belokurov (2015): *'Around a Milky Way-like galaxy, more than a thousand of these sub-halos will not be able to form stars but are dense enough to survive even deep down in the potential well of their host. There, within the stellar halo, these dark pellets will bombard tidal streams as they travel around the Galaxy, causing small but recognisable damage to the stream density distribution.'*

The importance of such searches was underlined by Bullock & Boylan-Kolchin (2017): *'Observational programs to constrain or discover and characterise the number of truly dark low-mass halos are among the most important, and achievable, goals in this field over the next decade. These efforts will either further verify the ΛCDM paradigm or demand a substantial revision in our understanding of the nature of dark matter.'*

TODAY, some of the most discussed features are those in the GD–1 stream. This was discovered, from SDSS, to span 63°, with a retrograde orbit with pericentre 14 kpc and apocentre 26 kpc (Grillmair & Dionatos, 2006). Based on their SDSS data, the discoverers considered that there was 'no evidence of perturbations by large mass concentrations in the nearby halo'.



FOLLOW-UP ground-based observations, including radial velocities, provided the first evidence for gaps along the stream (Koposov et al., 2010; Carlberg & Grillmair, 2013). The distant retrograde orbit suggested that interactions with disk sub-structure were unlikely, confirming it as an excellent candidate for the study of gaps induced by dark sub-halos (Amorisco et al., 2016; de Boer et al., 2018; Koppelman & Helmi, 2020).

Further insights came with the availability of the Gaia DR2 proper motions, from which the GD–1 stream was detected as one of the highest contrast features in the Galaxy halo (Malhan et al., 2018b). Additional stars surrounding the stream were detected by Price-Whelan & Bonaca (2018), including an off-track 'spur', perhaps suggesting the presence of massive perturbers (Bonaca et al., 2019), or that the progenitor originated within a larger system (Malhan & Ibata, 2019).

The 'exquisite astrometry from Gaia' allowed a clean separation of the stream from Milky Way stars (Price-Whelan & Bonaca, 2018), and showed clear evidence for high-contrast gaps along the stream (see the figure above). Here, ϕ_1 and ϕ_2 are angles on the sky in the coordinate system based on the stream itself.

Also working with the DR2 proper motions, and with improved filtering, de Boer et al. (2020) confirmed these various 'gaps' and 'wiggles'. They argued that a striking sinusoidal wiggle at high ϕ_1, straddling a gap feature at $\phi_1 \sim -3°$, cannot be the characteristic S-shape signature of stellar debris torn off the stream's progenitor, since it has the wrong orientation with respect to the stream's orbit. They concluded that this feature must instead come from a perturbation to the stream.

AGAIN BASING their analysis on DR2, but adding Pan-STARRS photometry, and new ground-based radial velocities, Ibata et al. (2020b) reached a different conclusion. They found that the density profile exhibits high contrast periodic peaks separated by 2.64 ± 0.18 kpc. Their N-body simulations suggested that this morphology could be modelled with simple epicyclic motion in a smooth Galactic potential, partly compounded by incompleteness in Gaia's sky-scanning pattern in DR2.

Such epicyclic *overdensities* arise because tidal stripping mainly occurs near pericentric passages, leading to bursts of debris along the stream (Küpper et al., 2010; Küpper et al., 2012; Sanders et al., 2016; Bovy et al., 2017). Ibata et al. (2020b) concluded that massive dark sub-halos do *not* appear to be required to explain the density clumping along the stream.

BUT, IN TURN, Banik et al. (2021a) argued that the power induced by episodic tidal stripping is far below that induced by dark matter sub-structures. They argued that the stellar density variations cannot be due to known baryonic structures, such as giant molecular clouds, globular clusters, or the Milky Way's bar or spiral arms (a similar conclusion was drawn by Doke & Hattori, 2022), and instead (in a joint analysis of the GD–1 and Pal 5 streams) requires a population of dark sub-structures with masses in the range $10^7 - 10^9\,M_\odot$.

They went on to infer a total abundance of dark sub-halos corresponding to a mass fraction in the sub-halos $f_{\rm sub} = 0.14^{+0.11}_{-0.07}$ percent, compatible with hydrodynamical simulations of cold dark matter with baryons.

MORE IS BEING inferred about the likely progenitor of the GD–1 stream. Although its location remains unknown, the N-body simulations by Webb & Bovy (2019) suggest that it probably lies between $-45° < \phi_1 < -30°$, and that it either completely disrupted ~2.5 Gyr ago, or it disrupted only 500 Myr ago, resulting in the underdensity at $\phi_1 \sim -40°$.

From 43 spectroscopically confirmed stream members, Gialluca et al. (2021) measured a radial velocity dispersion of 2.1 ± 0.3 km s^{-1}, constant over the 15° region surveyed. Compared with an unperturbed model of the GD–1 stream having a velocity dispersion of 0.5 km s^{-1}, the observed dispersion implies that the stream has undergone dynamical heating. They hypothesise that GD–1 originated from a globular cluster which, prior to its accretion by the Milky Way, orbited a dwarf galaxy with a cored density profile. They infer that imprints of its original host galaxy, including the inner slope of its dark matter halo, remain observable in the stream today.

In further modelling using Gaia DR3, Ibata et al. (2024) insist on the periodicity of the high contrast peaks, and suggest a two-component stream model (a kinematically cold part with dispersion 7 km s^{-1}, and a hot part with dispersion 29 km s^{-1}), consistent with simulations where globular clusters form at random locations within dark matter sub-halos and are subsequently accreted onto large galaxies (Carlberg, 2020; Malhan et al., 2021; Carlberg & Agler, 2023).

SIMILAR STUDIES for the Pal 5 stream (essay 109) suggest the presence of two gaps due to sub-halos of $10^6 - 10^7\,M_\odot$ and $10^7 - 10^8\,M_\odot$ (Erkal et al., 2017; Bonaca et al., 2020; Banik et al., 2021a; Banik et al., 2021b). Again, much more can be expected with Gaia DR4.

185. Mira variables

As evolved stars cool and expand on the red giant or asymptotic giant branch (AGB), they become pulsationally unstable. Such stars are classified, according to the amplitude and regularity of their light curves, as Mira (M), semi-regular (SR), and irregular (L) variables.

Miras are AGB stars of initial mass $1.5-4\,M_\odot$, with high mass-loss rates, and high luminosities ($\sim 10^3\,L_\odot$) due to their large envelopes. They are characterised (and defined) by large variability (≥ 2.5 mag in V, 0.3–1.0 in K_S), and periods of 100–1000 d. They occupy a key stage of stellar evolution, contribute to the heavy-element enrichment of the interstellar medium and, being luminous, are important tracers of Galactic structure. One of the nearest is Mira itself, at ~90 pc.

I have mentioned them, in passing, in my essays on Galactic tracers (6), the Large Magellanic Cloud (38), discovering variability (61), the Andromeda survey (84), the distance to the Galactic centre (111), the distance scale (122), non-radial pulsators (148), and radial velocity time-series of long-period variables (158).

Here, I will look at some Gaia results regarding their use as distance indicators and probes of Galactic structure. Many other surveys (including ASAS, LAMOST, VISTA and ZTF) are also contributing to new advances.

The chemistry of Mira variables is dominated by either C-rich or O-rich species according to the strength of 'dredge-up' episodes during the asymptotic giant branch phase (essay 167), largely reflecting their initial mass and metallicity (Höfner & Olofsson, 2018).

Both types satisfy their own period–luminosity relations (e.g. Iwanek et al., 2021), with the O-rich relations typically being tighter in the near-infrared due to effects of significant circumstellar dust in the C-rich variables. This makes the O-rich Mira variables particularly useful distance indicators, both within our Galaxy and beyond.

They are contributing to the 'Hubble tension' debate (to reconcile the early Universe expansion rate based on the CMB with the local value derived from Type Ia supernovae in nearby galaxies) as an independent Population I calibrator of the supernova luminosities, currently best served by the classical Cepheids (essay 44).

With a limit of $V \sim 12.5$ mag, Hipparcos observed some 900 long-period variables, including Miras, semi-regular or irregular variables, O-rich, C-rich, and S stars. Hipparcos contributed, for example, to improvements in the period–luminosity relations and to suggestions that the semi-regular variables are Mira progenitors (Bedding & Zijlstra, 1998), to the physics of the dredge-up episodes (Barthès et al., 1999), and whether they are controlled by fundamental or first overtone pulsation modes (van Leeuwen et al., 1997).

The Gaia sample is, of course, vastly larger. In 2018, DR2 provided an all-sky catalogue of 550 737 variable stars (with G, G_{BP}, and G_{RP} photometric time-series), of which 151 761 are long-period variable candidates having $\Delta G > 0.2$ mag, with one-fifth of these considered to be Mira candidates (Mowlavi et al., 2018).

The 34-month DR3 (essay 76) contains 1.72 million long-period variable candidates, including 392 240 with derived periods ranging from 35–1000 d, of which more than 40 000 are identified as Mira variables confirmed by OGLE-IV (Lebzelter et al., 2023). Robin et al. (2012a, Table 3.7) actually predicted some 40 000 with $G < 20$ mag, and around 18 000 with $G < 12$ mag.

Segregation into O-rich and C-rich Miras, using the Gaia BP/RP spectra to distinguish the different TiO bandheads, is discussed by Lebzelter et al. (2023, §2.4), Sanders (2023, §2.2), and Zhang & Sanders (2023, Fig. 3).

Let me turn to the use of Mira variables as distance indicators. The period–luminosity relation first established for Galactic Miras (Robertson & Feast, 1981) was soon shown to be much tighter for those in the LMC (Glass & Evans, 1981). Multiple sequences and other refinements were later revealed by the microlensing surveys MACHO and OGLE (Wood et al., 1999; Wood, 2000; Ita et al., 2004; Soszyński et al., 2013 and references).

Parallel work made use of the Hipparcos parallaxes (Whitelock & Feast, 2000). Whitelock et al. (2008) adopted a distance modulus of 18.39 ± 0.05 for the LMC to derive an infrared period–luminosity relation for O-rich Miras of the form $M_K = \rho(\log P - 2.38) + \delta$, with slope $\rho = -3.51 \pm 0.20$, and zero-point $\delta = -7.15 \pm 0.06$.

THE GAIA DATA have been used in several studies to revisit the Mira period–luminosity relation (including dependencies on colour and metallicity), both in our own Galaxy (Sun et al., 2023[DR2]; Sanders, 2023[DR3]), and in the Magellanic Clouds (Bhardwaj et al., 2019[DR2]).

While a complete picture remains unclear, the general findings are that uncertainties in the Galactic period–luminosity relation (e.g. Sanders, 2023, Fig. 5) remain larger than those found in the LMC (Sun et al., 2023) and, using HST–WFC3 infrared data, in M101 (Huang et al., 2024b). They may also be steeper at short periods than the LMC relations, perhaps suggesting the existence of population effects (Sanders, 2023).

Nonetheless, using these relations as anchors for Mira variables in other galaxies, Sanders (2023)[DR3] derived $H_0 = 73.7 \pm 4.4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ for the Type Ia host galaxy NGC 1559, while Huang et al. (2024b)[HST] found $H_0 = 72.37 \pm 2.97\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ for M101. As with recent Cepheid determinations, the Mira results suggest a discrepancy in H_0 between early and late Universe values.

A FURTHER, but exploitable, complication is that the pulsation period of Miras is correlated with their scale height and/or velocity dispersion (Feast, 1963). Interpreted as a period–age correlation, this allows them to be used as age indicators within the Galaxy and beyond (Grady et al., 2020). Mira variables in clusters, from Gaia-based membership determinations, confirm this connection (Grady et al., 2019; Marigo et al., 2022). Zhang & Sanders (2023) used 46 107 O-rich Mira candidates from Gaia DR3 to derive a period–age relation (with P in d) $\tau \approx (6.9 \pm 0.3)(1 + \tanh[(330 - P)/(400 \pm 90)])$ Gyr.

COMBINING THESE properties, Miras are being used as key probes of Galactic structure and evolution. As stated by Grady et al. (2020): *'Owing to their impressive brightness in the near infrared, we are able to trace the Miras right across the disk and through the bulge'.*

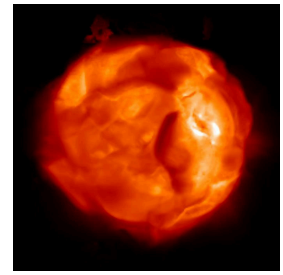
Using 21 149 O-rich Galactic Miras from Gaia DR2, they showed that the morphologies of both disk and bulge evolve as a function of age/chemistry. The disk is 'stubby' at the oldest ages (9–10 Gyr), but becomes thinner and radially extended ('peanut-like') at younger ages, consistent with the 'inside-out' and 'upside-down' formation of the Milky Way disk, and suggesting that the bar formation and buckling took place 8–9 Gyr ago. Similar results were found by Semczuk et al. (2022).

Such studies are starting to provide some profound insights into the formation of our Galaxy's central bar, and the nature of the associated nuclear stellar disk (Sormani et al., 2022b) and nuclear star cluster (Neumayer et al., 2020). Sanders et al. (2024) used Miras from VVV, placed on the Gaia reference frame, to elucidate one possibility: that the bar formed ~8 Gyr ago, close to the time of the Gaia–Sausage–Enceladus infall merger, potentially implying that the bar was tidally-induced.

THERE ARE, I should stress, theoretical explanations for these correlations: a period–luminosity relation follows from the fact that stars of a given mass, excited by convection, only begin pulsating in the fundamental mode over a narrow range of radii (Trabucchi et al., 2019). That between period and age is traced to the dependency on mass (Trabucchi & Mowlavi, 2022).

But a recurrent topic in the consideration of Mira variables is the finding that the scatter in the Galactic period–luminosity relation is larger than for the Large Magellanic Cloud and, at the same time, that the Gaia parallax standard errors for Miras are often significantly underestimated, by some 30–80%. This has been attributed in part to (presently uncalibrated) time-dependent chromatic displacements. But it is probably compounded by their large physical sizes (e.g. Mira, with $R \sim 400 R_\odot$, subtends ~40 mas). This means that many are resolved by Gaia, resulting in significant photocentric motions due to their large convection cells (Andriantsaralaza et al., 2022; Maíz Apellániz, 2022).

This has been further demonstrated by El-Badry et al. (2021, §5) who invoked the presumably identical parallaxes of some binary companions, and by Andriantsaralaza et al. (2022) who compared Gaia and VLBI parallaxes for a sample with masing circumstellar envelopes, finding significant differences for many.


Chiavassa et al. (2018, Fig 1a)

Models by Chiavassa et al. (2011) already predicted this sort of effect. Chiavassa et al. (2018) used 3D hydrodynamic simulations of convection to construct intensity maps in the Gaia G band (325–1030 nm). Comparison with solar neighbourhood semi-regular variables from Gaia DR2 suggested time-dependent photocentric excursions of 5–10% of the stellar radius, with convection-related variability accounting for a substantial part of the Gaia DR2 parallax error.

This was further confirmed with Gaia EDR3 by Chiavassa et al. (2022) for a sample of red supergiants in the young cluster χ Per (at 2.260 ± 0.020 kpc). An implication, further discussed by Kochanek (2023), is that these inferred photocentric displacements might be used as a probe of the star's surface dynamics. Woodland & Montez (2022) showed that the same phenomenon also affects the interpretation of 'proper motion anomalies' (essay 174) when attempting to identify potential long-period binary companions.

SUCH CONSIDERATIONS are clearly complicating the use of Galactic Mira variables in constructing a robust local distance ladder. The extent to which this continues to be the case with future Gaia releases remains to be seen.

186. Moving groups and traceback ages

YOUNG LOCAL ASSOCIATIONS and moving groups are small nearby aggregates of typically a few dozen stars, assumed to have been born within the same molecular cloud, and therefore at the same time and place, with similar initial chemical composition, and sharing a common global space motion. Their uniform initial properties underpins their importance for studies of star formation and their subsequent evolution (e.g. de Zeeuw et al., 1999; Jayawardhana, 2000).

Their age is key in modelling the early stages of star and planet formation, and various methods based on stellar evolution models can be used for their estimation, most notably based on theoretical isochrones or models of lithium depletion.

An independent method, yielding what are referred to as dynamical or traceback ages, makes use of their expanding space motions. Assuming that the stars were formed at a time when the association was most spatially concentrated, tracing back their space motions can establish their birth epoch. In most cases, this yields ages broadly consistent with those from stellar evolution models, but not always precisely so.

In practice, dynamical traceback models have variously employed linear trajectories, epicyclic approximations, or orbit integration within a specified Galactic potential. One crucial limitation has been knowledge of the stars' space motions, due to uncertainties in their distances, proper motions, and radial velocities (e.g. Ortega et al., 2002; Song et al., 2003; Ortega et al., 2004).

Today, traceback studies of several moving groups are making use of the Gaia astrometry and radial velocities. I will start with the important case of β Pictoris.

ONE OF THE NEAREST, richest and today most intensively studied of these nearby associations or moving groups, β Pic was discovered through the identification of two companions to the A star β Pic by Barrado y Navascués et al. (1999). Subsequent identification of other co-moving stars has led to several hundred possible members known today (Zuckerman et al., 2001; Torres et al., 2006; Malo et al., 2013; Binks & Jeffries, 2016).

AT A DISTANCE of only 40 pc, some members host disks (Kalas & Jewitt, 1995; Kalas et al., 2004); exoplanets (Lagrange et al., 2010; Lagrange et al., 2019; Chauvin et al., 2012), and exocomets (Kiefer et al., 2014).

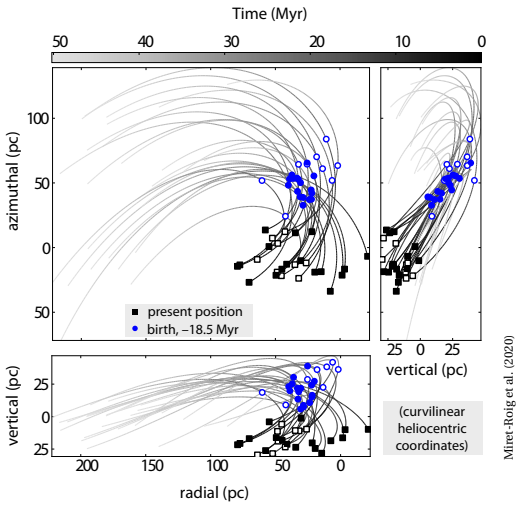
Of some two dozen ages for the β Pic moving group tabulated by Miret-Roig et al. (2020, their Table 6), pre-Gaia estimates have centred around 20 Myr (e.g. Mamajek & Bell, 2014), but with Li depletion estimates as high as 25 ± 3 Myr (Messina et al., 2016), and more uncertain dynamical traceback ages ranging from 11.5 Myr (Ortega et al., 2002) to 31 ± 21 Myr (Makarov, 2007).

Gaia astrometry has now been used to refine group membership, based on DR2 (Ujjwal et al., 2020), and subsequently DR3 (Lee et al., 2024). The latter resulted in 106 single and resolved companions, and 47 unresolved binaries, but still with a wide model-dependent age range: 23 ± 8 Myr from Dartmouth magnetic models fit to the lithium depletion boundary, 33 ± 10 Myr fit to the Gaia M_G versus $B_P - R_P$ colour–magnitude diagram, and 11 ± 4 Myr as best fit to the 2MASS–Gaia M_{K_S} versus $B_P - R_P$ colour–magnitude diagram.

BUT I WILL focus here on the association's Gaia-based dynamical traceback age. Miret-Roig et al. (2020) used DR2 astrometry, supplemented by new ground-based as well as Gaia-determined radial velocities, to determine the accurate space motions for their selected 26 *bona fide* members.

They used a specific axisymmetric (bulge/disk/halo) Galaxy potential to integrate the equations of motion, propagated backward in time to –50 Myr. They defined the dynamical age as the time at which the members of the association were most concentrated in space.

Their orbital projections in the Galaxy (actually in curvilinear heliocentric coordinates, centred on the Sun's position, and rotating around the Galactic centre) are shown below, colour-coded according to this backward time. Black squares are the present positions, with blue circles representing positions at the inferred birth epoch, $t = -18.5$ Myr. They also identified a particular concentration of 17 'core' stars, shown as filled symbols.



They found a dynamical traceback age of $18.5^{+2.0}_{-2.4}$ Myr, broadly consistent with the most robust estimates from isochrone or lithium depletion models, and a size (defined by the trace covariance matrix) of 7 pc at birth. And they concluded that, with Gaia, the observational uncertainties no longer dominate the uncertainties in the age, although it does remain sensitive to the definition and selection of association members. Further DR2-based studies were made by Crundall et al. (2019).

A subsequent study based on Gaia DR3 by Couture et al. (2023) concluded that the radial velocities include biases, due to gravitational redshift and convective blueshift, of order 0.6 km s^{-1}. Their chosen sample of 25 stars (out of their full sample of 76 members) then yields a corrected age of 20.4 ± 2.5 Myr.

LET ME BRIEFLY MENTION some of the other Gaia-based studies that are being made on nearby associations. Some have been focussed on refining association membership, some on improved isochrone ages, and others on estimating dynamical traceback ages.

At around 400 pc, the Orion star-forming region is of complex morphology, with multiple stellar populations, and star formation having taken place over an extended period of some 10 Myr. It is too distant for precise parallax or proper motion estimates by Hipparcos, such that much of its structure and dynamics has remained uncertain. Studies have been reported using Gaia DR2 (Kounkel et al., 2018), and EDR3 (Swiggum et al., 2021), the latter providing evidence for radial expansion of two of its stellar groups from a common centre.

Other Gaia-based membership and traceback studies have been reported for 32 Ori (Luhman, 2022), TW Hya (Luhman, 2023), Upper Scorpius (Squicciarini et al., 2021), Ophiuchus (Miret-Roig et al., 2022), Cepheus Far North (Klutsch et al., 2020; Kerr et al., 2022a), and Fornax–Horologium (Kerr et al., 2022b).

INTERESTING insights into the star-formation process are emerging from detailed consideration of these expansion velocities. Kuhn et al. (2019) used Gaia DR2 data for a sample of 28 clusters and associations with ages from 1–5 Myr, to show that at least 75% are expanding, but with rotation detected in only one. Typical expansion velocities are on the order of 0.5 km s^{-1}, with a positive radial gradient in expansion velocity in some.

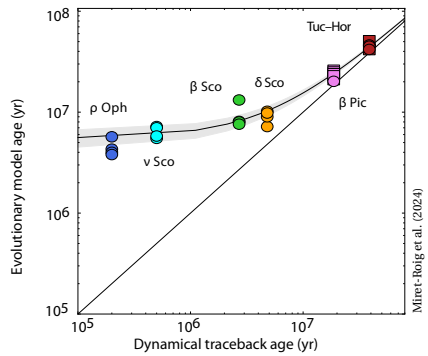
Systems still embedded in molecular clouds are less likely to be expanding than those that are partially or fully revealed. In star-forming regions containing multiple clusters or sub-clusters, they found no evidence that the groups are coalescing, implying that hierarchical cluster assembly, if it occurs, must happen rapidly during the embedded stage.

In related work, a variation of the dynamical traceback age based on the association's 'evaporation age', is also being advanced by Gaia (Pelkonen et al., 2024).

DEVELOPING THESE IDEAS, and based on a comparison of recent dynamical traceback and isochrone ages, Miret-Roig et al. (2024) showed that there is a systematic difference between the two, with the traceback ages consistently younger than the isochrone ages by an average of 5.5 ± 1.1 Myr. They concluded that the discrepancy arises because the two methods have a different time origin: if the star cluster is gravitationally bound before the dispersion of the parent gas cloud, the zero-point of the expansion time scale is a few Myr after that of the colour–magnitude diagram method.

In other words, the dynamical traceback 'clock' starts when a stellar cluster or association begins to expand after expelling most of the gas, whereas the isochronal 'clock' starts earlier when most stars form; for example, when most of the material in the envelope has collapsed onto the disk, and the central protostar becomes observable at infrared wavelengths (e.g. Wuchterl & Tscharnuter, 2003). As a result, the age difference between the two methods is providing important clues about the cluster formation and gas dispersal processes.

In particular, the age difference appears to represent an observational measurement of the duration of the embedded phase, and the timescale of gas dissipation.



187. Gaia synthetic photometry

A RECENT PAPER by Montegriffo et al. (2023b) is titled *'Gaia DR3. The Galaxy in your preferred colours: Synthetic photometry from Gaia low-resolution spectra'*. It is an important contribution to Gaia's science. It runs to 59 pages, and containing many details. But it is well summarised by the first two sentences of the abstract.

> *'Gaia DR3 provides novel flux-calibrated low-resolution spectrophotometry for 220 million sources in the wavelength range 330–1050 nm. Synthetic photometry directly tied to a flux in physical units can be obtained from these spectra for any passband fully enclosed in this wavelength range.'*

ANYONE who has worked with photometric standards, or has had to convert observations to absolute flux, will know how important, yet how subtle and complex, such work can be. Montegriffo et al. (2023b) provide the tools to convert Gaia photometry to any other photometric system, subject to a couple of obvious caveats.

Suppose that you have several years of Sloan Digital Sky Survey photometry in its five photometric bands (u, g, r, i, z), and you want to extend the time series using the Gaia data. Or you want to transform the Gaia photometry to the systems used by Pan-STARRS or the Vera C. Rubin Observatory. This can now be done!

GAIA OBSERVES in three bands. The main astrometric field, G, maximises photon throughput, and spans 330–1050 nm, defined by mirror reflectivities and the CCDs. The BP and RP data (together referred to as XP) are acquired as low-resolution spectra ($\lambda/\Delta\lambda \simeq 25-100$) over 330–680 nm and 640–1050 nm respectively. They have been subject to an extensive validation (De Angeli et al., 2023), and internal/external calibration (Carrasco et al., 2021; Montegriffo et al., 2023a).

Gaia's all-sky, multi-epoch measurements yield photometry which rivals the best available. Gaia DR3, from 2022, provides mean G magnitudes for 1.8 billion sources, and mean $G_{\rm BP}$ and $G_{\rm RP}$ photometry for 1.5 billion, typically at sub-millimag precision. Mean BP and RP *spectra* are provided for 220 million sources, mostly with $G < 17.5$ mag. The final data release, DR5, foreseen around 2030, will extend this to all 2–3 billion sources.

LET ME GIVE some background to the design of the low-resolution spectra for reasons that will become clear later. The acquisition of multi-colour photometry in parallel with the astrometric observations was a design goal from the earliest days of Gaia, and was originally foreseen (and extensively optimised) as a set of (eventually 19) broad- and narrow-band filters (Jordi et al., 2006). As I described in essay 68, and at a rather late satellite design stage, the dedicated photometric telescope with its associated filters was exchanged for two fused silica prisms which disperse the light entering the main telescope's two fields of view.

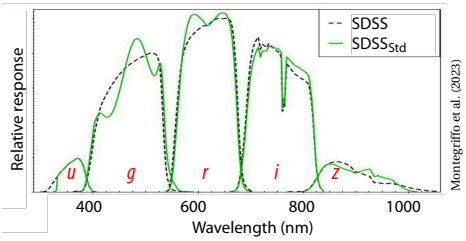
Their resolution, which matched the earlier filter design, results from the intrinsic dispersion of fused silica, and varies from 3–27 nm per pixel over 330–680 nm for BP, and from 7–15 nm per pixel over 640–1050 nm for RP.

SYNTHETIC PHOTOMETRY, as derived from observed spectrophotometry, is based on the computation of a normalised mean flux obtained by integrating the product of a transmission curve, $S(\lambda)$, and a spectral energy distribution over a given wavelength (or frequency) interval, depending on the adopted photometric system (e.g. Bessell, 2005). It can be used to provide colours within standard systems, for insights into the spectra or colours of unusual stars or stellar systems, for the validation or re-calibration of existing photometric surveys, and *'the opportunity to experiment with the performance of a photometric system on a huge data set of real data on real sources before its actual realisation'*.

In the Gaia archive, the XP spectra are stored as the coefficients (and their covariances) of a set of orthogonal basis functions, from which the spectral energy distributions can be reconstructed (De Angeli et al., 2023; Montegriffo et al., 2023a).

A KEY STEP in the derivation of the XP synthetic photometry, described in detail by Montegriffo et al. (2023b, §2.2), is 'standardisation'. For Gaia, this compensates any systematics present in the externally calibrated XP spectra by 'tweaking' the Gaia transmission curve to minimise any residual magnitude or colour differences.

The example below shows the Sloan Digital Sky Survey transmission curves (Doi et al., 2010, black dashed line), and their 'tweaked' version, obtained with the standardisation process, as the green continuous lines. I stress that the green lines do *not* indicate some improved estimation of the SDSS filter profiles, but rather reflect empirical adjustments which correct for the systematic errors that still affect the externally calibrated XP spectra.



In principle, such synthetic photometry can be obtained from the calibrated XP spectra in any photometric system subject to two constraints: that the target passband must (of course) be enclosed within the spectral range covered by Gaia's XP spectra (330–1050 nm), and that the passband's characteristic width exceeds the 'line-spread function' of the XP spectra at the relevant wavelength.

THEY THEN detail a number of important photometric systems which they have reconstructed from the Gaia XP-derived synthetic photometry. Their general conclusions are that existing high-quality photometry can be reproduced within a few per cent over a wide range of magnitudes and colour, for wide and medium bands, and with around milli-mag accuracy.

For example, in their earlier paper detailing the external calibration of the XP spectra, Montegriffo et al. (2023a, §8.4.2) showed that the Hipparcos H_p, B_T, and V_T photometry, recognised as a benchmark of excellent precision by Bessell (2005), are all reproduced by the XP synthetic photometry to better than 2.5 milli-mag.

Montegriffo et al. (2023b) then detail and demonstrate performances for various wide-band systems, including the Sloan Digital Sky Survey, the Johnson–Kron–Cousins system, and the PanSTARRS and HST systems.

Amongst narrow-band photometric systems, they detail results for the Strömgren system (widely used for the determination of effective temperature and surface gravity), the Javalambre J–PAS and J–PLUS surveys, and the IPHAS Hα emission-line survey.

WITH MY LONG involvement in the preparation of the Gaia photometric system, I was particularly fascinated by one of their specific examples: replicating the proposed photometric *filter* system which I referred to earlier: the set of broad- (C1B) and medium-band (C1M) filters, carefully designed by Jordi et al. (2006) to maximise the scientific return in terms of $T_{\rm eff}$, metallicity, gravity, reddening, and even α-element abundances.

To investigate how the system would have performed, Montegriffo et al. (2023b, §4.4) reconstructed the C1 colour indices, and showed that they could, indeed, clearly separate giant and main sequence stars in the $\log g$ versus $T_{\rm eff}$ plane (their Fig. 22), and could pick out both the white dwarf sequences, and the 'Jao gap' (essay 152) in the various colour–magnitude diagrams.

THEY ILLUSTRATE several other scientific applications, including the detection of multiple populations in globular clusters, metallicity derived in the Strömgren system, metallicity estimation for very metal-poor stars, and the classification of emission-line sources.

And to make this synthetic photometry more readily accessible, they provide two derived catalogues.

The first is the Gaia Synthetic Photometry Catalogue (GSPC), which includes the majority of the 220 million stars with XP spectra released in Gaia DR3, in 13 passbands, including UBVRI in the JKC system, *ugriz* in the SDSS system, and two in the HST–ACS/WFC system.

The other is the Gaia Synthetic Photometry Catalogue for White Dwarfs (GSPC–WD), comprising 100 000 white dwarfs with DA/non-DA classification, obtained with a Random Forest (machine-learning) algorithm.

PRE-DATING the synthetic photometry, Gaia photometry was already being used to correct other photometric surveys. The 6-colour zero-points of ANU's SkyMapper Southern Survey (SMSS), had already been anchored to Gaia DR2 (Huang et al., 2021). Pancino et al. (2022) used Gaia EDR3 to recalibrate 200 000 secondary standards in the Johnson–Kron–Cousins system.

Huang et al. (2024a) describe independent efforts to correct residual systematics in the Gaia DR3 XP spectra using external spectral libraries including CALSPEC, Hubble's NGSL, and LAMOST DR7.

Amongst work already making use of the GSPC, Zhou et al. (2023a) used the DR3 synthetic photometry to correct patterns in the DESI (Dark Energy Spectroscopic Instrument) Legacy Imaging Surveys. Anderson et al. (2024) derived corrections for their OGLE variables in their measurement of TRGB magnitudes.

López-Sanjuan et al. (2024) used the catalogue to recalibrate the 12 passbands of the Javalambre Photometric Local Universe Survey (J–PLUS) third data release.

Xiao et al. (2024) undertook a comprehensive recalibration of narrow-, medium- and broad-band photometry from the Southern Photometric Local Universe Survey (S–PLUS) using the Gaia synthetic photometry.

THE LOW-RESOLUTION BP and RP spectral data are described by De Angeli et al. (2023) as *'one of the exciting new products in Gaia Data Release 3'*. Montegriffo et al. (2023b) suggest that the availability of the derived synthetic photometry may constitute *'a true revolution in optical photometry'*. We can let future users decide!

188. The tip of the red giant branch

SOME IMPORTANT CONTRIBUTIONS are being made by Gaia in furthering the use of the 'tip of the red giant branch' as a robust distance indicator. These sorts of new insights are becoming ever more crucial in the context of the ongoing 'Hubble tension' debate.

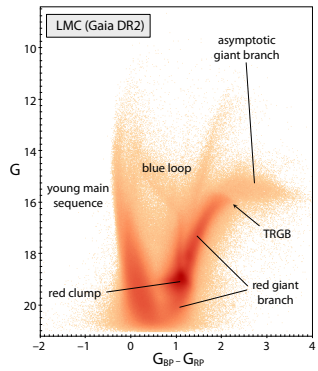
VARIOUS STAGES OF stellar evolution are characterised by well-defined luminosities. This has led to the construction of 'distance ladders' which go far beyond the reach of direct trigonometric parallaxes, and which allow distances to be estimated across our Galaxy, and into the Local Group of galaxies and beyond. Several potential 'standard candles' have been identified over recent decades, amongst them the RR Lyrae, Cepheid and Mira variables, and Type 1a supernovae.

Today, Cepheids lead efforts to calibrate the extragalactic distance scale (essay 122). But it is the discrepancy in the values of H_0 characterising the local ('late Universe') cosmic expansion, inferred from the combined HST and Gaia EDR3 measurements of Cepheids, $H_0 = 73.2 \pm 1.3$ (Riess et al., 2021), compared to that inferred in the 'early Universe' from the Planck satellite measurements of the Cosmic Microwave Background, $H_0 = 67.4 \pm 0.5\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ (Aghanim et al., 2020), that sits at the heart of the 'Hubble tension' (essay 44).

Both values are very precise, and both are also considered to be very *accurate*, such that the disagreement, although small, is considered statistically significant.

THE 'tip of the red giant branch' (TRGB) is arguably becoming *the* other leading distance indicator. Unlike the Cepheid scale, which only applies to (Pop I) systems with recent or ongoing star formation, the TRGB method can be used wherever metal-poor red giant branch stars are sufficiently abundant to allow its location to be defined precisely (Beaton et al., 2018).

It provides distance indicators for evolved populations, and can reach the galaxy hosts of SN Type Ia (e.g. M101, M106, NGC 1448, IC 1613), and so can contribute to determining the Hubble constant (e.g. Beaton et al., 2019; Hoyt et al., 2021; Jang et al., 2021; Hoyt, 2023).



The Gaia DR2-based G versus $G_{BP} - G_{RP}$ colour–magnitude diagram for the central region (8.°8 radius) of the Large Magellanic Cloud (Clementini et al., 2020). It shows the TRGB location, along with the various stellar populations described by Luri et al. (2021).

Sakai (1999) traces the method to the realisation by Baade (1944) that the brightest resolved red stars in M31 and two companion ellipticals all had the same brightness and colour. Sandage (1971) later proposed that the brightest are at the tip of the first-ascent red giant branch. Here low-mass stars, evolving up the red giant branch, abruptly halt their increasing luminosity at the moment of He-ignition within their core, resulting in a sharp discontinuity in the star's evolutionary track.

Iben & Renzini (1983) showed that the bolometric luminosity at this core He 'flash' for low-mass stars varies by only ~0.1 mag over ages of 2–15 Gyr, and for metallicities most relevant for Galactic globular clusters, $-2.2 \leq$ [Fe/H] ≤ -0.7. More recent models are given by Serenelli et al. (2017), Saltas & Tognelli (2022), and others. A considerable and growing body of observational work continues to confirm the method's potential.

The 'tip' manifests itself as a discontinuity in the population's luminosity function, or colour–magnitude diagram. In the Gaia data it can be seen, for example, as a prominent feature of the G versus $G_{BP} - G_{RP}$ colour–magnitude diagram for the central region of the Large Magellanic Cloud (e.g. Luri et al., 2021; Fig. 2). Quantitative methods to estimate the tip's location include edge-detection (Sobel) filters to measure the first derivative of the RGB luminosity function (Lee et al., 1993; Freedman et al., 2019; Scolnic et al., 2023; Anderson et al., 2024).

Determining H_0 using the TRGB then involves determining absolute distances to galaxies that host SN Ia events, but which are also close enough to have their distances measured (whether by TRGB or Cepheids), then using the SN Ia luminosities to infer distances for a sample of galaxies far enough into the Hubble flow that their peculiar velocities are a small fraction of the cosmological recessional velocities.

The Carnegie–Chicago Hubble Program (CCHP) bases its current determination of H_0 on a sample of 10 galaxies over distances 7 Mpc (M101) to 20 Mpc (NGC 1316), which together embrace 11 supernovae from the Carnegie Supernova Project (Krisciunas et al., 2017). HST observations (typically using the F814W filter) establish the TRGB location in the I-band.

The final critical step is to establish the absolute zero-point of the TRGB colour–magnitude diagram using some more fundamental distance measure. The latest CCHP determination uses a distance modulus of 18.477 mag for the LMC based on 20 detached eclipsing binaries (Pietrzyński et al., 2019), to yield an I-band TRGB absolute magnitude of $M = -4.049$ mag. A similar value is found using the Megamaser-anchored distance to NGC 4258 (Jang et al., 2021).

The latest CCHP value, $H_0 = 69.8 \pm 0.8\,(\pm 1.1\%\ \mathrm{stat}) \pm 1.7\,(\pm 2.4\%\ \mathrm{sys})\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$, appears compatible with both Cepheid and Planck values (Freedman et al., 2019; Freedman, 2021). But the recent determination by Scolnic et al. (2023), anchored to NGC 4258, gives $H_0 = 73.22 \pm 2.06\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$, favouring the Cepheid value.

Freedman et al. (2019) note that their *'ultimate goal for the absolute calibration is the geometric parallax measurements for Milky Way red giant branch stars being obtained by Gaia'*, although they were still waiting for a more robust estimate of the Gaia parallax zero point.

I will not go deeper into this topic, which includes effects of age, metallicity, and extinction, but refer the reader to recent more in-depth reviews (Freedman, 2021; Freedman & Madore, 2023; Lee, 2024).

One of the main ways that Gaia is contributing is by determining the TRGB absolute magnitude using geometrical parallaxes of Milky Way stars. Working with Gaia DR2, Mould et al. (2019) showed that the high Galactic latitude colour–magnitude diagram, drawn from our Galaxy's thick disk and inner halo, is consistent with the CCHP calibrations of the TRGB.

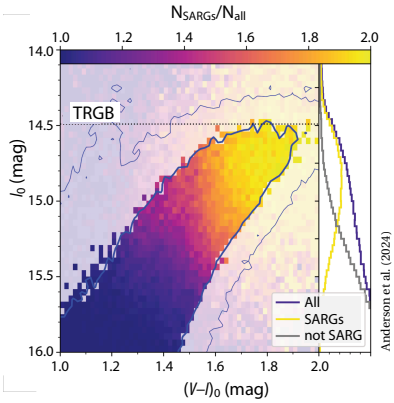
Soltis et al. (2021), in their wider Gaia EDR3 study of the Milky Way globular cluster ω Cen (essay 40), emphasised its merits for anchoring the TRGB calibration: the direct use of trigonometric parallaxes, well-calibrated extinction, and with nearly 200 stars within a magnitude of the tip. They estimated an I-band TRGB magnitude $M_I = -3.97 \pm 0.06$ mag, fainter by 0.07 mag than that used by Freedman et al. (2019), and yielding $H_0 = 72.1 \pm 2.0\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$, closer to the Cepheid value.

Li et al. (2022c) used a maximum likelihood method to calibrate the brightness of the TRGB using Gaia EDR3 parallaxes of Milky Way field giants at high Galactic latitude, finding $M_I = -3.91 \pm 0.05 \pm 0.09$ mag. Li et al. (2023c), used DR3 and Gaia synthetic photometry (essay 187) to better constrain the luminosity function, yielding $M_I = -3.970^{+0.042}_{-0.024} \pm 0.062$ mag.

Dixon et al. (2023) used high Galactic latitude halo stars to minimise effects of metallicity, dust, and crowding, and used PARSEC isochrones (instead of Sobel edge detection) to find $M_I = -4.042 \pm 0.041 \pm 0.031$ mag.

Two other contributions illuminate some additional complications in using the TRGB as distance indicator. Anderson et al. (2024) noted that stars near the TRGB are typically regarded as non-variable. But, in a detailed study of the LMC, they show that all stars near the TRGB are small-amplitude red giants (SARG) that follow several period–luminosity sequences.

While their variability data is (presently) only from OGLE, the contributions of Gaia (DR3) were to remove foreground stars by astrometric cuts, to remove blended stars, and to provide synthetic photometry in the HST–ACS/F814W passband. They concluded that this variability population diversity affects the TRGB at a level exceeding the stated precision, and that both luminosity function smoothing and edge detection weighting can further bias the measurements. They derived $M_I = -4.025 \pm 0.014 \pm 0.033$ mag, assuming the geometric distance to the LMC given by Pietrzyński et al. (2019).



Another possible complication is that the location of the tip of the red-giant branch would be affected by any energy loss leading to a larger core mass at He-ignition, and thus to a brighter luminosity than predicted by standard models. From the Gaia DR2 distance of ω Cen, Capozzi & Raffelt (2020) gave a limit on the neutrino dipole moment of $\mu_\nu < 1.2 \times 10^{-12} \mu_\mathrm{B}$, and on the axion–electron coupling of $g_\mathrm{ae} < 1.3 \times 10^{-13}$.

There are many challenges in using the TRGB. Gaia is certain to contribute much more in the future.

189. More metallicities and gravities

SINCE Data Release 3 in 2022 (which included a catalogue of metallicities, temperatures and gravities for 470 million sources), four community-generated catalogues of [M/H], $T_{\rm eff}$, and $\log g$ have been made available, which I will describe here. To understand their context, let me start with some background.

ALONG WITH THE main astrometric field, and the radial velocity spectrometer (RVS) field, the Gaia focal plane includes two low-resolution objective prism fields, BP (blue photometer) and RP (red photometer), together generally denoted XP (essay 68). The prisms disperse the spectra over ~45 pixels in the along-scan direction, with the 60-pixel window allowing for sky subtraction. The spectral dispersion varies from 3–27 nm per pixel over 330–680 nm for BP, and from 7–15 nm per pixel over 640–1050 nm for RP.

For the majority of objects, $G \gtrsim 11.5$ mag, the BP and RP spectra are binned on-chip in the across-scan direction, over 12 pixels, to form one-dimensional, along-scan spectra. The dispersed spectra overlap in crowded regions, and BP/RP acquisition is accordingly limited to about 750 000 objects per square degree. The spectra are wavelength and flux calibrated, while their integrated fluxes provide calibrated multi-epoch photometry, designated $G_{\rm BP}$ and $G_{\rm RP}$.

The 34-month Data Release 3 (essay 76) then comprises 1.812 billion sources with astrometric solutions to 21 mag, 1.806 billion with mean G magnitudes, 1.5 billion with mean $G_{\rm BP}$ and $G_{\rm RP}$ photometry, along with 219 million mean BP/RP spectra.

The task of source classification (as star, white dwarf, physical binary, quasar, or galaxy), and the 'astrophysical parameters inference system' (Apsis) is carried out by Coordination Unit 8 of DPAC, and I have given some details (and references) in essay 89. The 13 Apsis modules take relevant combinations of the RVS spectra (GSP–Spec) and BP/RP spectra (GSP–Phot), fit a number of astrophysical parameters to the data, and pass the results to the FLAME module which derives the evolutionary parameters radius, luminosity, mass, and age.

GSP–Phot uses the XP spectra to estimate $T_{\rm eff}$, $\log g$, [M/H], $M_{\rm G}$, radius R, distance, and extinctions by modelling the BP/RP spectra, G magnitude, and parallax (Andrae et al., 2023a; Creevey et al., 2023, §3.5). To match the XP spectra, Andrae et al. (2023a) used four stellar atmospheric models to cover different ranges of $T_{\rm eff}$, the Fitzpatrick (1999) mean extinction law, and a grid of (PARSEC) isochrones to fix the absolute magnitudes.

As a result, astrophysical parameters in DR3 include $T_{\rm eff}$, $\log g$, and [M/H] for 470 million sources using BP/RP (and 6 million using RVS); radius (470 million), mass (140 million), age (120 million), and spectral types (220 million), along with smaller catalogues of chemical abundances, diffuse interstellar bands, activity indices, Hα equivalent widths, and emission-line stars.

IN AN IMPORTANT application of the XP spectral database, Montegriffo et al. (2023) provided the tools to generate synthetic photometry, for all 220 million sources with mean BP and RP spectra, in any user-specified photometric system (see essay 187).

They also made available an associated Gaia Synthetic Photometry Catalogue (GSPC) which contains, for the majority of the 220 million stars with XP spectra from DR3, colours in 13 passbands. These include UBVRI in the Johnson–Kron–Cousins system, *ugriz* in the SDSS system, and two in the HST–ACS/WFC system.

LET ME EMPHASISE the size of the Gaia XP spectral database: there are more than 20 times as many as in the largest ground-based spectral survey, LAMOST, although at only 1/20th of the spectral resolution.

And the metallicities published with DR3 can still be improved, in part through improved calibration, more so through additional observations as planned for DR4, and also through better modelling of some key spectral lines given the low resolution of the XP spectra. Furthermore, the use of synthetic model spectra (to match the observed XP spectra) is most likely sub-optimal, leaving parameter estimates sensitive to inaccuracies and omissions in the underlying models.

THE ASTROPHYSICAL parameters generated by DPAC CU8 were never intended be the final word in classification or parameter estimation, and improvements in calibration, and in training sets and algorithms were considered inevitable.

The motivation to construct better metallicities is, of course, based on the pursuit of accurate chemical abundances across all stellar populations throughout the Galaxy, which are crucial inputs for studies of star formation, detailed nucleosynthesis modelling, Galactic chemical and dynamical evolution, and so on.

And already, since 2023, four other large catalogues of astrophysical parameters, focusing on [M/H], have been derived from the XP spectra, as follows:

ZHANG ET AL. (2023) used forward modelling to estimate $T_{\rm eff}$, $\log g$, and [Fe/H] for all 220 million DR3 stars with XP spectra. Their training set used atmospheric parameters from LAMOST, augmented by 2MASS and WISE photometry to reduce degeneracies and yield more precise estimates of $T_{\rm eff}$ and reddening. Their catalogue includes $T_{\rm eff}$, $\log g$, and [Fe/H], along with revised parallaxes and extinctions. It ignores binary stars, and does not cover all parts of the Hertzsprung–Russell diagram, notably white dwarfs.

ANDRAE ET AL. (2023b) building on previous work by Rix et al. (2022), employed a specific machine-learning algorithm, XGBoost. It was trained on 500 000 stars with stellar parameters from APOGEE, including those with CatWISE 3.4 μm and 4.6 μm infrared photometry to reduce the degeneracy between $T_{\rm eff}$ and reddening. The training set was augmented by some 300 very metal-poor stars from LAMOST (Li et al., 2022a), and they included the Gaia parallaxes to assist constraints on $\log g$ and [M/H].

Although therefore tied to the parameter scale of the APOGEE survey, the resulting catalogue of 175 million stars has a mean precision of 0.1 dex in [M/H] and, obtained as by-products, 50 K in $T_{\rm eff}$, and 0.08 dex in $\log g$. They also provide a catalogue of 17 million bright ($G < 16$) red giants using more conservative cuts to ensure a higher data quality.

Chandra et al. (2024) describe an application of this catalogue in identifying the evolution of angular momentum in the Galaxy with metallicity.

HATTORI (2024) used tree-based machine-learning to estimate [M/H] and [α/Fe] for 48 million giants and dwarfs in low-extinction regions from the DR3 XP spectra. Again, the training set used APOGEE DR17 and the metal-poor stars of Li et al. (2022a). It resulted in a mean precision of 0.09 dex for [M/H] and 0.04 dex for [α/Fe], with the most reliable values being for giants and metal-rich stars which dominate the training set.

YAO ET AL. (2024) focused on the much rarer very metal-poor stars, [Fe/H] < -2, also using XGBoost. For $G_{\rm BP} < 16$, they developed classifiers optimised for turn-off stars and for giant stars, finding 11 000 metal-poor turn-off stars, and 111 000 or 44 000 bright metal-poor giants depending on the target purity. For $G_{\rm BP} > 16$, they identified 38 000 additional turn-off candidates, and 41 000 additional metal-poor giant candidates.

Investigations by Witten et al. (2022), meanwhile suggest that metal-poor stars can be identified for $G < 16$ using the XP spectra, but that true detections will be overwhelmed by false positives at fainter magnitudes.

THE XP DATA will provide a magnificent resource for chemical abundance investigations. But it is also a substantial database for research into classification algorithms, and the trade-offs between 'physics-driven' (relying on synthetic stellar spectra) and 'data-driven' (based on machine-learning) classification.

Some of this path has already been trodden, for example by APOGEE (e.g. Ness et al., 2015; Leung & Bovy, 2019; Ting et al., 2019).

As a cautionary example, at least some algorithms which estimate [α/Fe] from the XP spectra do so by exploiting known correlations between [α/Fe] and other elements, rather than the direct effect of [α/Fe] on the spectrum (e.g. Gavel et al., 2021; Hattori, 2025).

THOSE INVOLVED in these kinds of studies will appreciate the contribution of Laroche & Speagle (2025). They argue that physics-driven models inevitably suffer from a 'synthetic gap', viz. a combination of theoretical and instrumental effects which together produce unresolvable differences between synthetic and observed spectra. At the same time, data-driven models which depend on 'labels' (a generic term here covering $T_{\rm eff}$, $\log g$, [M/H], and [α/Fe]) themselves suffer from 'label systematics' which decrease any model's performance.

They demonstrate this by applying a variational auto-encoder (unsupervised learning) to the XP spectra which (they argue) learns stellar properties directly from the data. They also show that the spectra *do* contain meaningful [α/Fe] information, by identifying α-bimodality in the absence of stellar label correlations.

They conclude: *'Label-dependent models are incapable of exploiting the entire astrophysical information in the XP data, because they are limited by the availability of stellar labels to train on. Novel data-driven techniques must be developed to tackle this big data problem'.*

Using the same techniques used by Large Language Models for AI, and applied to the XP spectra, Leung & Bovy (2024) argue that *'building and training a single foundation model without fine-tuning using data and parameters from multiple surveys to predict unmeasured observations and parameters is well within reach'.*

190. More phase-space features

In the early days of the Gaia studies, Adriaan Blauuw, ESO's second Director General, a supporter of space astrometry, and chair of the scientific selection committee for Hipparcos, asked me: was I concerned that Gaia would generate *too much* data… too big a step after Hipparcos… more than astronomers could handle?

It's an interesting viewpoint. The main Hipparcos catalogue of 120 000 stars in five printed volumes occupies half a meter of shelf space. The same format for Gaia would extend to 10 km. Nobody can ever look at every Gaia light curve: just 1 s each would require 60 years.

But with today's computing power, the huge stellar content is one of Gaia's great strengths: a vast net that can catch fleeting stages of stellar evolution, define statistically useful samples of rare star types, find tiny features in the HR diagram that point to exotic physics, and sift through millions of stars to seize on a few that are the fossil record of a galaxy captured billions of years ago.

And it is the number and variety of complex phase-space features in our Galaxy that has been one of Gaia's greatest contributions to studies of its structure and evolution to date. Recent studies have uncovered still more.

It is easy to feel lost in the labyrinth of Gaia science, so let me start by listing some of the dynamical features that Gaia has discovered (marked *) or substantially advanced, where numbers refer to my own essays.

These include: the Hercules stream attributable to dynamical resonances (115); the origins for the Arcturus and HR 1614 streams (116); discovering some 100 halo streams* (156); the breathing motion of spiral arms* (173); detecting the bar's deceleration due to the dark matter halo* (112); the disk warp (72); the Gaia phase-space spiral, attributed to a massive collision* (117); characterising globular cluster and open cluster tidal tails (109); identifying the primordial heart of the Milky Way* (102); measuring aberration due to Galaxy rotation* (32); discovering the Radcliffe Wave as an instability or perturbation structure* (127).

Here I look at four more recent discoveries: one each related to the inner and outer halo, one to the central disk, and one to our Galaxy's cosmological evolution.

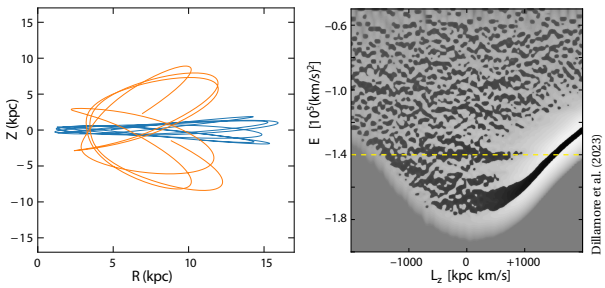
The Galaxy's central bar drives complex dynamical behaviour in the stellar disk, responsible for some of the clumps of space velocities in the solar neighbourhood, including the Hercules stream (essay 115).

The Gaia data have already identified specific resonances in the local disk driven by the bar (essay 112). Local kinematics are consistent with a bar pattern speed $\Omega_b = 35.5\,\mathrm{km\,s^{-1}\,kpc^{-1}}$, and possibly a deceleration due to the dark matter halo of $-4.5\,\mathrm{km\,s^{-1}\,kpc^{-1}}$ per Gyr (Chiba et al., 2021; Chiba & Schönrich, 2021).

A recent Gaia result by Dillamore et al. (2023b) suggests that these resonances also extend to halo stars, consistent with some earlier suggestions (Hattori et al., 2016; Myeong et al., 2018b; Schuster et al., 2019).

Dillamore et al. (2023b) used DR3 astrometry and RVS radial velocities (for 22 million stars with 6D phase space and [Fe/H] data) to identify a prominent feature in the phase-space distribution of halo stars: a 'ridge' at constant energy and positive angular momentum, and consistent with $\Omega_b \simeq 35-40\,\mathrm{km\,s^{-1}\,kpc^{-1}}$. Their simulations associate this with stars trapped in resonance orbits with the bar, particularly at the corotation resonance, many with large vertical (halo-like) excursions.

The resonances give the inner stellar halo a net spin in the direction of the bar's rotation (cf. Weinberg, 1985; Athanassoula, 2002). And the variation of rotational velocity with radius is similar to that seen for metal-poor stars in the APOGEE survey (Abdurro'uf et al., 2022).



Left: two simulated orbits in the $R-Z$ plane, some of which are disk-like (blue), and some are halo-like (orange). Right: the Gaia data in the $E-L_z$ plane, where dark regions correspond to overdensities. The dashed line marks the ridge-like resonance.

Let me mention that an important contribution to the use of Gaia's astrometry, and exploited below, are the low-resolution spectra, BP and RP. Mean spectra were made available with DR3 for 220 million stars, along with derived estimates of [M/H], $T_{\rm eff}$, and $\log g$.

The spectra have been re-calibrated by a number of groups (essay 189). As an example, Andrae et al. (2023b), provide 175 million stars with a precision of 0.1 dex in [M/H], 50 K in $T_{\rm eff}$, 0.08 dex in $\log g$, with a higher quality subset of 17 million bright ($G < 16$) red giants.

Growing evidence that the halo comprises substantial debris from an early massive accretion event has been brought into sharp focus by Gaia, through its ability to identify stars, mixed in phase space, but sharing the same metallicity and the same conserved integrals of motion. The present picture is that the Gaia–Sausage–Enceladus (GSE) merger, 8–10 Gyr ago and [Fe/H] $\simeq -1.2$, contributed a large fraction of the stellar halo, at least out to ~30 kpc (essays 15 and 71).

Whether the merger was prograde or retrograde is still debated. Evidence for the latter includes the Arjuna stream, of similar metallicity, considered as a possible remnant of its early-stripped tail (Naidu et al., 2020).

Beyond 10–20 kpc lie several 'overdense' regions. The Hercules–Aquila Cloud, and the Inner Virgo and the Eridanus–Phoenix overdensities, have been suggested to be its apocentric pile-ups. The Pisces Overdensity at 70 kpc (Sesar et al., 2010; Carlin et al., 2012), and the Outer Virgo Overdensity at 80 kpc (Sesar et al., 2017), might be other distant remnants of the same merger.

Chandra et al. (2023) have made a major step in disentangling and interpreting these structures by making use of Gaia DR3 astrometry, supplemented by metallicities from Gaia's low-resolution prism spectra. They used 200 000 red giant branch stars out to 100 kpc, with full 5D (and some with 6D) kinematics, using isochrone-based distances constrained by their measured metallicity.

They identified a large population of retrograde debris representing most distant 'echoes' of the GSE merger, linking the more distant northern Outer Virgo Overdensity and the southern Pisces Overdensity to successive apocentres. The majority beyond 40 kpc follow a great-circle track consistent with the GSE orbit, and distinct from the Sagittarius stream which occupies the same plane, but orbits in the opposite sense.

Their findings also match the N-body simulations of the GSE merger by Naidu et al. (2021), which consisted of a $5 \times 10^8 M_\odot$ infall, starting at $z \sim 2$, and with a tilted and retrograde orientation before rapidly 'radialising'.

Of course in giving these, and the two following, short summaries of important advances in understanding our Galaxy's structure and evolution, I have omitted many details and associated studies.

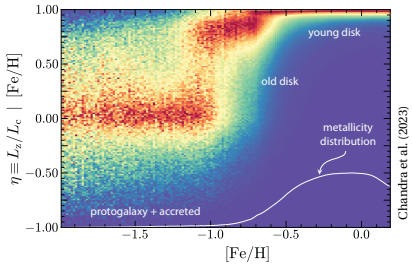
Using the Gaia DR3 astrometry and spectroscopy of 5.8 million stars, along with metallicities from Andrae et al. (2023b), Malhan & Rix (2024) identified two high-contrast overdensities in the energy–angular momentum ($E - L_z$) plane of bright ($G < 16$) and metal-poor ($-2.5 <$ [M/H] < -1.0) stars, which they designated Shakti and Shiva. Both have $M_\star \gtrsim 10^7 M_\odot$, and both follow prograde orbits inside the solar circle. But while their orbits suggest an accreted origin, their metal abundances are more typical of an *in situ* population.

They reconciled these properties by inferring either that they may have resulted from resonant orbit trapping of the field stars by the rotating bar, along the lines described (above) by Dillamore et al. (2023b). Or that they were protogalactic fragments that formed stars rapidly and coalesced early. In this case, they might be similar to those of the 'poor old heart' of the Milky Way (Rix et al., 2022, see essay 102), although not as deep in the Galactic potential, and with still discernible orbits.

Finally, Chandra et al. (2024) used the [Fe/H] and [α/Fe] for 9.9 million red giants derived by Andrae et al. (2023b), to characterise their angular momentum as a function of metallicity. Taking this as a proxy for age, they identified three distinct evolutionary phases: a disordered/chaotic protogalaxy, a (kinematically) hot old disk, and a cold young disk. In their interpretation, the old high-α disk starts at [Fe/H] $\simeq -1.0$, 'spinning up' from the nascent protogalaxy (as also inferred by Belokurov & Kravtsov, 2022), and then exhibiting a smooth cooling down toward more ordered and circular orbits.

The overlap between the protogalaxy and hot disk, $-1.2 \lesssim$ [Fe/H] $\lesssim -0.9$, includes stars with intermediate orbits, extending to [Fe/H] ~ -0.5, and coinciding with a previously identified 'in-situ halo' or 'splash' population (Bonaca et al., 2017; Belokurov et al., 2020b).

They also identified an analogue from the cosmological simulations TNG50, in which the protogalaxy spins up into a thin high-α disk, before being heated and torqued by a major gas-rich merger (as with the GSE merger). This adds a large amount of low-metallicity gas and angular momentum, from which the kinematically cold low-α stellar disk is subsequently born.



As the authors state, we are steadily gaining a coherent picture of our Galaxy's three-phase formation… involving spin-up, merger, and cooldown.

191. Extinction

A RECURRENT THEME in the Gaia literature is interstellar extinction. This is important in the determination of stellar parameters through spectral modelling (e.g. [M/H], $T_{\rm eff}$, and $\log g$), in the interpretation of some key evolutionary states, and in describing the distribution of gas and dust in the solar neighbourhood.

My goal here is to summarise *only* what has been measured with Gaia, and I will not touch on any scientific interpretation. Nor will I discuss the extra contribution of circumstellar dust around specific objects, nor atmospheric effects relevant for ground-based observers.

LIGHT is absorbed and scattered by gas and dust along the sight line from star to observer, and most stellar spectra are affected by at least some interstellar dust. The details and spatial variation of interstellar extinction depend on the properties of dust grains along a line-of-sight, and measurements also convey information about their composition and size distribution.

Interstellar extinction is rather a smooth function of wavelength, albeit with superimposed absorption features due to particular chemical species. These include the ultraviolet (217 nm) bump, diffuse interstellar bands (some seen in the Gaia RVS spectra; essays 92, 160), and others beyond Gaia's response in the infrared.

Extinction in a given direction can be inferred by comparing an observed spectrum with its closest atmosphere model. In the solar neighbourhood, extinction in the V-band averages some 0.7–1.0 mag kpc^{-1}, but it is much higher in specific regions, notably in the Galactic plane, and especially towards the Galactic centre.

Since blue light is more strongly attenuated than red, extinction also causes objects to appear redder as well as dimmer. This 'reddening' is often simply characterised by an object's *colour excess* in some specified photometric system, e.g. $E_{B-V} = (B-V)_{\rm obs} - (B-V)_0$.

The parameter $R \equiv A_V/E(B-V)$ characterises the ratio of total to selective extinction. Ranging between 2.2–5.8 for sight lines where ultraviolet extinction has been measured, a mean relationship $A_V = 3.1\,E(B-V)$ is frequently used (e.g. Cardelli et al., 1989; Fitzpatrick, 1999; Fitzpatrick & Massa, 2007).

THERE IS A HUGE literature on extinction, and historical advances have also followed from ultraviolet and (far-)infrared observations. Early studies were made by Trumpler (1930), van Rhijn (1949), Schatzman (1950), Münch (1952), Chandrasekhar & Münch (1952), Gottlieb & Upson (1969), Milne & Aller (1980), and Neckel & Klare (1980), Arenou et al. (1992) and many others.

For extragalactic astronomy, and as a foreground in interpreting the Cosmic Microwave Background, a two-dimensional map of *integrated* dust extinction and reddening is usually sufficient. The maps of Schlegel et al. (1998), based on IRAS 100 μm and COBE–DIRBE 100 μm and 240 μm data, are still some of the most widely used corrections (e.g. Malhan & Rix, 2024).

THE IMPROVED HIPPARCOS distances advanced the determination of spatial variations within the Galaxy. Vergely et al. (1998) used extinctions determined from Strömgren photometry and Hipparcos parallaxes for 11 837 stars with $r < 400$ pc, from which they derived an average opacity $A_V = 1.5$ mag kpc^{-1}, and a scale-height of 70 pc. Hipparcos distances to open clusters were also used to construct extinction maps in the Galactic plane, $|b| < 10^\circ$ (Chen et al., 1998; Joshi, 2005).

To extend the mapping horizon pre-Gaia, star distances were generally estimated from stellar spectral energy distribution models.

Drimmel et al. (2003) presented a model based on the COBE–DIRBE dust maps. The dust distribution was mapped in greater detail using near-infrared colours from 2MASS (Marshall et al., 2006), and using optical and near-infrared colours from SDSS (Berry et al., 2012).

Probabilistic (Bayesian) models, inferring both the distribution of dust and the stellar types and distances, were derived from the INT Photometric Hα Survey of the Northern Galactic Plane (Sale et al., 2014), and from Pan-STARRS 1 and 2MASS photometry of 800 million stars over three-quarters of the sky, which extended the mapping to several kpc (Green et al., 2015; Green et al., 2018a). Lallement et al. (2014) used 23 000 sight lines to trace the dust within 2.5 kpc, adopting a Gaussian weighting to enforce smoothness on small spatial scales.

AROUND THE TIME of the Gaia launch, Lallement et al. (2014) noted that *'3D maps of the Galactic interstellar matter are a potential tool of wide use, but accurate and detailed maps are still lacking'*.

Gaia is changing this. Accurate parallaxes provide the means of quantifying the dependence of extinction on distance for millions of sight lines in the Galaxy. And in addition, reddening and extinction *for each star* can be estimated from the low-resolution BP/RP spectra, and from the RVS spectra. Because successive Gaia data releases supersede the previous, I will summarise the early results only briefly. The referenced papers generally also provide links to their derived extinction maps.

USING DR2, Andrae et al. (2018) derived G-band extinctions (A_G) and $E(B_P - R_P)$, for 88 million sources, but based only on integrated photometry (G_{BP}, G_{RP}), and broad assumptions. Leike & Enßlin (2019) used a Gaussian process model using these extinctions to derive reddening maps within a few hundred pc.

Chen et al. (2019a) used DR2 parallaxes, (G_{BP}, G_{RP}) photometry, and infrared photometry from 2MASS and WISE, to derive $E(G-K_S)$, $E(B_P-R_P)$, and $E(H-K_S)$ for 56 million stars in the Galactic plane, with an angular resolution of 6 arcmin. The dust shows the disk warp, and complex structures associated with the Sagittarius, Local, and Perseus arms.

Lallement et al. (2019) used DR2 parallaxes with DR2 and 2MASS photometry to derive the dust distribution in a $6\times6\times0.8$ kpc^3 volume centred on the Sun. They discussed various features evident in the Carina–Sagittarius arm, the Local arm/Cygnus Rift, and the Perseus arm (including a wavy pattern in some regions), and the link between the dust concentrations and the locations of molecular clouds, H II regions, O stars, and masers.

Green et al. (2019a) derived reddening maps from 799 million stars with $\delta > -30°$ and out to a few kpc, using DR2 parallaxes combined with Pan-STARRS1 and 2MASS photometry.

Anders et al. (2019) used DR2 parallaxes and photometry, along with Pan-STARRS1, 2MASS, and AllWISE photometry for 137 million stars with $G < 18$, yielding a median precision of 0.20 mag in A_V for $G \leq 14$. Amongst their extinction-corrected colour–magnitude diagrams, there is a clear manifestation of the Galactic bar in the stellar density distributions.

Sun et al. (2022) used Gaia DR2 photometry combined with LAMOST photometry to determine reddening maps for 2 million stars, comparing them favourably with the widely used maps of Schlegel et al. (1998).

WITH EDR3, several improved and updated maps, based on the same general methods, were constructed and made available (Anders et al., 2022; Lallement et al., 2022; Vergely et al., 2022).

SUBSTANTIAL IMPROVEMENTS came with the availability of Gaia DR3, in part from the improved parallaxes, but also with the availability of the mean BP/RP (aka XP) spectra for 220 million stars (see essay 189). For their determination of atmospheric parameters using GSP–Phot to match the XP spectra, Andrae et al. (2023a) actually used the Fitzpatrick (1999) mean extinction law.

Kordopatis et al. (2023) used parallaxes, atmospheric parameters, 2MASS and Gaia EDR3 photometry to compute ages, masses, and reddenings for 5 million stars with RVS spectra. Another model using Gaia, 2MASS and ALLWISE was given by O'Callaghan et al. (2024).

Revised extinctions were also derived through improved reductions of the XP spectra (essay 189), viz. the 220 million stars of Zhang et al. (2023), the 175 million sample of Andrae et al. (2023b), and the 48 million sample of Hattori (2025). Using the XP spectra on a slightly different metallicity scale, based on empirically calibrated theoretical models, An et al. (2024) determined both reddening and metallicity for 80 million main-sequence stars, modelling foreground extinction out to approximately 3 kpc.

Amongst the latest, Gontcharov et al. (2023a) used Gaia, Pan-STARRS1, SkyMapper, 2MASS, and WISE photometry for 100 million stars to derive 3D extinction maps in V and G within 2 kpc (resolution 3.6–11.6 pc transverse and 50 pc radial; see their Fig. 10), a 3D differential extinction map, and a 2D map of total Galactic extinction for Galactic latitudes $|b| > 13°$, and with a precision in A_V of 0.06 mag. An analytical model, treating the 3D dust distribution as a superposition of three overlapping layers, is given by Gontcharov et al. (2023b).

THE XP SPECTRA also allow determination of extinction at their spectral resolution, $R \sim 20-100$. Zhang et al. (2023) constructed a 'universal extinction' versus wavelength curve, derived from the data without reference to any previous model, and with each wavelength interval modelled separately. The resulting smooth extinction curve (their Fig. 15) agrees reasonably well with the $A_V = 3.1\,E(B-V)$ model of Cardelli et al. (1989).

LET ME RECALL that this essay has focused on extinction estimates based on the Gaia data, and does not cover other ongoing efforts, such as with JWST (e.g. Wang & Chen, 2024), and others (e.g. Yu et al., 2023).

Let me also mention that specific sky regions are often subject to specialised extinction studies, especially towards the Galactic bulge (e.g. Nataf et al., 2013; Nataf et al., 2016), and towards the SMC and LMC (e.g. Prevot et al., 1984; Fitzpatrick, 1986; Wang & Chen, 2023).

EVEN WITH Gaia DR3, several extinction maps and calibrations are now available. Absent some wider consensus, each user must make their own choice!

192. The structure of molecular clouds

ON THE subject of the structure of Galactic molecular clouds, Dharmawardena et al. (2023) wrote *'Great progress has been made in this field with the arrival of the Gaia mission'*. I will give some background, explain why their 3D structure is relevant, and illustrate how Gaia is contributing. I should also mention that my previous essay, on extinction, is closely tied to this present topic.

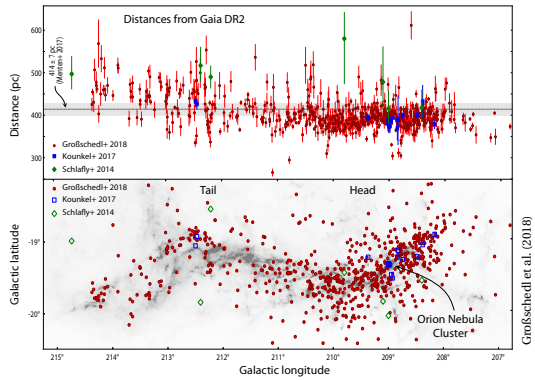
MOLECULAR CLOUDS are regions of the interstellar medium dense enough (10–30 cm^{-3}) to form molecular gas. This results in 'dark nebulae' that obscure light from background stars, and in whose densest regions star formation can occur. *Giant* molecular clouds (GMC) extend to 10–200 pc, and masses $10^4 - 10^7\,M_\odot$.

Maps of the Galaxy in CO, a constituent more easily detectable than the dominant H_2, show that such molecular gas is largely confined to the mid-plane of the Galaxy, with a scale height ~50–60 pc. Since ultraviolet photons from associated star formation transform molecular material back to the atomic state, their close association with spiral arms in turn suggests that molecular clouds form and dissociate on a timescale $\lesssim$10 Myr.

Parallaxes for relatively nearby stars make it possible to derive cloud distances and even their 3D structures, important in understanding how processes such as turbulence and magnetic fields affect star formation within them (e.g. Hartmann et al., 2001; Kennicutt & Evans, 2012; Evans et al., 2014; Heyer & Dame, 2015).

THE PRINCIPLES underlying Gaia's contributions can be seen in the GMC archetype, Orion A, the most active local star-forming region, from which many key observables, including star-formation rates and history, multiplicity, initial mass function, and protoplanetary disk populations, have been derived. Estimates of cloud mass, physical size, and star formation all depend critically on accurate distance measurements.

The Orion Nebula Cluster, the richest cluster at the northern end of Orion A, lies at ~400 pc. Pre-Gaia observations, while suggesting a distance gradient across the cloud (Schlafly et al., 2014; Kounkel et al., 2017), could say little about its shape or 3D structure.



The Gaia DR2 parallaxes for a number of Young Stellar Objects (YSO), with ages $\lesssim$ 3 Myr, can now be used as tracers of the cloud's 3D shape (Großschedl et al., 2018; Kounkel et al., 2018). Großschedl et al. (2018), for example, used a sample of 682 YSOs, to trace distances of young objects within the Orion A cloud, shown here projected on the Planck–Herschel dust column-density.

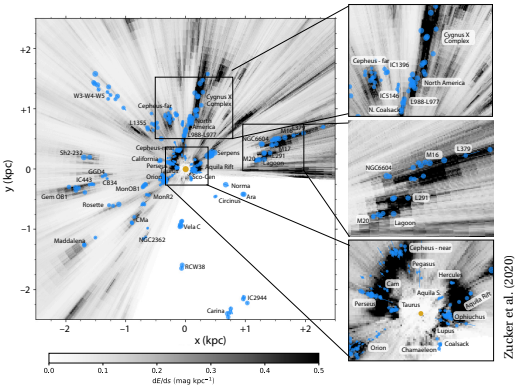
They concluded that Orion A is not the straight filamentary cloud seen in projection, but instead a cometary-like cloud with aspect ratio 30:1, and two distinct components: a denser and enhanced star-forming 'head', and a lower density 75 pc long 'tail'. The true extent, 90 pc rather than the projected 40 pc, makes it the largest molecular cloud in the local neighbourhood.

Zucker et al. (2019) also used DR2 parallaxes to derive distances and extinctions to tens of thousands of stars towards several dozen molecular clouds within 2 kpc. They gave average distances (accurate to 5–6%) to most of the major named clouds in the CO survey of Dame et al. (2001), including Perseus, Orion A, Taurus, Ophiuchus, California, and Cepheus. An interactive version of their Fig. 1 shows their distance–extinction relations towards the Perseus cloud as characterised before and after Gaia, with the most probable distance and extinction to each star indicated with a red cross.

Yan et al. (2019) similarly used DR2 parallaxes and G-band extinctions to derive distances to 52 high Galactic latitude ($|b| \geq 10°$) molecular clouds, 13 for the first time, and again with errors of typically 5%.

THE 2-VOLUME Star Formation Handbook (Reipurth, 2008a; 2008b) contains 60 star forming regions within 2 kpc. Although well studied, their distances have remained poorly constrained. Amongst them, several have pre-Gaia estimates that vary by more than a factor of two (e.g. Circinus, Coalsack, NGC 2362, IC 5146).

Zucker et al. (2020) derived DR2-based star distances and extinctions in sightlines towards these clouds. They used an implementation of the method pioneered by Wolf (1923), viz. by fitting a line-of-sight dust model to identify the distance at which a discontinuity appears in the stellar reddening. Comparison with maser (parallax) distances results in consistent estimates, with 10% scatter, across their entire distance range 150–2500 pc.



AN ALTERNATIVE to identifying molecular clouds from their CO emission is from dust extinction measurements. Already from 2MASS colour excess maps, Dobashi (2011) had identified 7000 new molecular clouds, although without distance information.

With Gaia DR2, focusing on low Galactic latitudes, $|b| < 10°$, and using the Gaia-based colour excesses of 56 million stars previously derived by Chen et al. (2019a), Chen et al. (2020) identified 567 dust/molecular clouds with a hierarchical structure identification method, with distance uncertainties ~5%. The clouds lie along the known Sagittarius, Local and Perseus Arms, with a possible spur connecting the Local and Sagittarius arms.

SEVERAL STUDIES have exploited the Gaia distances and photometry in various combinations to probe the 3D structure of several of these molecular clouds. Leike et al. (2020) used Gaia, 2MASS, PANSTARRS, and ALLWISE data to create resolved dust maps of clouds out to 400 pc with a resolution of 1 pc.

These were then used by Zucker et al. (2021), who first established the skeleton-like 'spines' of each cloud, thereafter determining the radial volume density profile around each of these spines. Well-described by a two-component Gaussian, they inferred that the clouds are characterised by broad lower-density outer envelopes, with narrow higher-density inner layers, consistent with a transition between atomic and diffuse molecular gas.

Other works employing extinctions, dust-mapping, and Gaia DR2 distances to recover the molecular cloud structures include 3D modelling of the thin disk structures (Babusiaux et al., 2020); mapping the disk extinction beyond the range of Gaia parallaxes (Hottier et al., 2020); and using 18 million stars over 450 deg^2 to obtain the extinction and morphology of the Vela complex (Hottier et al., 2021).

Dharmawardena et al. (2023) used their 3D dust-mapping algorithm `dustribution`, which makes use of the stellar distance and extinction estimates from Fouesneau et al. (2022), to recover the 3D structure of 16 Galactic molecular clouds at 1 pc resolution. They then used the `astrodendro` package (Robitaille et al., 2019), which creates dendrograms of hierarchical structures in astronomical data, applied to the dust regions around each complex. From this, Dharmawardena et al. (2023) estimated the volume, mass, and density for each cloud and its associated substructures.

AS summarised by Cahlon et al. (2024), in the absence of the third dimension, pre-Gaia catalogues of molecular clouds were derived using position–position data ('$p-p$' space) on the plane of the sky, and sometimes using position–position–velocity data ('$p-p-v$' space), in which spectral-line observations of the ISM can add a third (non-spatial) dimension. They considered that, previous to their work, only the Chen et al. (2020) and Dharmawardena et al. (2023) catalogues were based on true 3D '$p-p-p$' data from 3D dust mapping.

In their own work, Cahlon et al. (2024), started with the 3D dust map of Leike et al. (2020), and performed a similar dendrogram analysis to produce a catalogue of 65 distinct local molecular clouds, now in true $p-p-p$ space, and at 1 pc resolution. Projecting back into 2D, they could perform a detailed comparison with the previous 2D maps, for example finding a steeper power law of the mass–size relation than in previous 2D mappings.

GAIA IS ALSO contributing to an understanding of the internal dynamics of molecular clouds. Zhou et al. (2022) studied 150 Young Stellar Objects with $d \lesssim 3$ kpc, finding their associated clouds to be elongated, and oriented parallel to the disk mid-plane. As probed by the YSOs, the turbulence is isotropic, and the 2D velocity dispersion is related to size by $\sigma_v \propto r^{0.67}$. The turbulent energy dissipation rate decreases with Galactocentric radius which, they suggest, can be explained if the turbulence is driven by cloud collisions.

Xu et al. (2023a) relate this to the formation of wide binaries, a topic which I will pick up in essay 193.

ALL THESE studies (none so far using DR3) underline the expectation that the new, Gaia-enabled, 3D mapping of local molecular clouds will contribute to many advances in star-formation studies.

193. More on wide binaries

I REVIEWED some of the contributions being made by Gaia on the subject of wide ($\gtrsim 10^3$ au) binaries in two earlier essays. In essay 14, I looked at evidence that their orbit distributions are inconsistent with Newtonian dynamics and, some claim, more consistent with MOND.

In essay 37, I looked at 'ultra-wide' binaries (separations $\gtrsim 30\,000$ au): how Gaia is able to detect them, and their possible formation and disruption mechanisms.

Both date from 2021, and I will look here at more recent Gaia results regarding their formation and orbit distributions. In cases where recent insights have come from Gaia, but without expanding further, I note this with a superscript indicating the relevant data release.

Gaia's contribution to binary star research more widely is reviewed by El-Badry (2024) and Merle (2024).

ON THE EXISTENCE of wide ($\gtrsim 10^3$ au) binaries in the Galactic field and halo, Kouwenhoven et al. (2010) wrote that *'their origin has long been a mystery'*. Still today, their origin *'remains a mystery'* (Xu et al., 2023a), and indeed *'a major puzzle'* (Modak & Hamilton, 2023).

Simulations find that they cannot form by fragmentation of a single molecular cloud core, nor in typical clusters (e.g. Clarke, 2019; Deacon & Kraus, 2020).

Other possible formation processes include the fragmentation of turbulent molecular cores (Goodwin et al., 2007), cluster dissolution (Kouwenhoven et al., 2010; Moeckel & Clarke, 2011; Rozner & Perets, 2023), dynamical 'unfolding' of compact triple systems (Reipurth & Mikkola, 2012), formation in adjacent molecular cloud cores (Tokovinin, 2017), chance trapping in the tidal tails of stellar clusters (Peñarrubia, 2021), or from 3-body encounters (Atallah et al., 2024).

Observations nonetheless suggest that the components are likely to have formed in the same star-forming region. Evidence includes the large fraction of wide pairs of young stars in low-density star-forming regions (Tokovinin, 2017), the chemical homogeneity in the components of wide binaries (Hawkins et al., 2020$^{\rm DR2}$; Andrews et al., 2019$^{\rm DR2}$), the metallicity dependence of the wide-binary fraction (Hwang et al., 2021$^{\rm DR2}$), as well as N-body simulations (Kroupa & Burkert, 2001).

ONE OF THE suggested formation mechanisms is that they arise as a natural consequence of star formation in the turbulent interstellar medium. This has found some interesting support from Gaia.

In this scenario, stars form in molecular clouds and, specifically, within the high-density filaments and cores at the collisional interfaces of converging highly supersonic turbulent flows (Federrath et al., 2009; Mocz & Burkhart, 2018; Inoue et al., 2018; Xu et al., 2019).

Most recently, Xu et al. (2023a) found that binaries with separations 1000–100 000 au form as a natural consequence of such a turbulent environment, and that such systems implicitly exhibit a 'superthermal' eccentricity distribution similar to that observed, as expanded on below. These ideas have been further reinforced by the simulations of Hamilton & Modak (2024).

GAIA OBSERVATIONS have provided support for this specific formation channel through two interesting and independent avenues.

The first is by measuring the eccentricity distribution of wide binaries, which has been possible for the first time with Gaia. At fixed separation a, the eccentricity distribution is expected to be 'thermal', $P(e) = 2e$, if binaries are spread out uniformly in phase space.

In contrast, from the DR2-based binary catalogue of El-Badry et al. (2019), Tokovinin (2020) used the direction and speed of relative motions in wide pairs to infer, statistically, their eccentricity distribution. He found that for separations $\gtrsim$1000 au, there is an excess of very eccentric orbits (termed 'superthermal'), with the distribution well-fit by a power law $P(e) = (1+\alpha)e^{\alpha}$, with $\alpha \sim 1.2$. This was further substantiated, using values of the relative separation and velocity vectors of components from Gaia EDR3, by Hwang et al. (2022).

Given that such a distribution does not arise as a result of Galactic tides (Modak & Hamilton, 2023) nor as a result of stellar encounters (Hamilton & Modak, 2024), the latter authors concluded that the eccentricity distributions measured by Gaia favour a turbulent fragmentation origin. A testable prediction is that α should be a monotonically decreasing function of binary age.

THE SECOND PIECE of evidence favouring a turbulent origin of wide binaries is that, in this model, the turbulent velocities of the gas are expected to be imprinted on those of newly formed stars. And indeed, recent Gaia observations find that the velocity differences and spatial separations of young stars statistically follow the power-law velocity scaling of interstellar turbulence.

Ha et al. (2021) used Gaia DR2 and APOGEE–2 data to derive positions and velocities of 1439 young stars in the Orion Molecular Cloud. The velocities exhibit the expected characteristics of turbulence, with their 'velocity structure function', over scales of 1–100 pc, consistent with the Kolmogorov-like predictions of Larson (1981).

The results have been further substantiated by Ha et al. (2022) based on similar studies using Gaia EDR3 in the star-forming regions of Ophiuchus (at 140 pc using 107 stars), Taurus (140 pc, 139 stars), Perseus (320 pc, 138 stars) and Orion (410 pc, 2468 stars).

Zhou et al. (2022) extended the study, using DR2, to 15 149 YSOs in 150 associations younger than 3 Myr within 3 kpc. They found that their associated clouds are elongated, and oriented parallel to the disk mid-plane. As probed by the YSOs, the turbulence is isotropic, and the 2D velocity dispersion is related to size by $\sigma_v \propto r^{0.67}$. The turbulent energy dissipation rate decreases with Galactocentric radius which, they suggest, is explained if the turbulence is driven by cloud collisions.

RETURNING to the topic of essay 14 (April 2021), on the orbit distributions of wide binaries. At that time, based on Gaia DR2, there was a consensus that there is a non-Newtonian tail of their component velocity ratios beyond $\gtrsim 7000$ au. But it was unclear whether this is evidence for some form of Modified Newtonian Dynamics, MOND (Hernandez et al., 2019), or explicable as stars born in the same cluster and currently undergoing a chance close 'flyby' (Pittordis & Sutherland, 2019), or as a population of hidden triple systems (Clarke, 2020).

In the past year, two papers (Chae, 2023; 2024), both using wide binary stars from DR3, have argued for a gravitational anomaly showing up at low accelerations, $\lesssim 10^{-10}\,\mathrm{m\,s^{-2}}$.

In the first, Chae (2023) selected 26 615 wide binaries within 200 pc with accurate distances, proper motions, and reliably inferred stellar masses, then de-projected the observed velocities and separations to 3D relative velocities and separations by Monte Carlo modelling.

The difference between the observed (kinematic) acceleration, v^2/r, and the Newtonian expectation, GM/r^2 (where M is the total system mass), shows a significant deviation of 0.109 ± 0.013 at accelerations below $10^{-10.15}\,\mathrm{m\,s^{-2}}$. *'What is even more surprising'*, Chae concludes, *'is that the trend and magnitude of the gravitational anomaly agree with what the AQUAL (Bekenstein & Milgrom, 1984) theory predicts'*.

In the second paper, Chae (2024) confirmed the conclusions for separations 2000–5000 au, with a more stringently selected sample from Gaia DR3. From 2463 binaries with a precision on parallaxes and proper motions better than 0.005, and radial velocities better than 0.2, the ratio of observed to predicted acceleration is $1.49^{+0.21}_{-0.19}$ for accelerations $\lesssim 10^{-10}\,\mathrm{m\,s^{-2}}$.

At about the same time, Hernandez (2023) repeated his earlier analysis using Gaia DR3, now using the Gaia-determined stellar masses and estimates of binary probabilities for each star using spectroscopic information, along with a larger sample of radial velocities. He confirmed his previous findings that the resulting relative velocity scalings accurately trace Newtonian expectations for the high-acceleration regime, but remain markedly inconsistent at low accelerations. As he states, *'A non-Newtonian low-acceleration phenomenology is thus confirmed'*.

Meanwhile, Pittordis & Sutherland (2023) provided an update of *their* 2019 study using Gaia EDR3 for 73 159 wide binaries within 300 pc and with $G < 17$ mag. Component masses were estimated from a main-sequence mass–luminosity relation. The frequency distribution of pairwise relative projected velocities as a function of projected separation was compared to simulations.

As in their earlier study, their distributions show a peak at a value close to Newtonian expectations, along with a long 'tail' which extends to much larger velocity ratios. Hypothesising that this is caused by hierarchical triple systems with an unresolved or unseen third star, they fit the observed distributions with a simulated mixture of binary, triple and flyby populations, for GR or MOND orbits, finding that standard gravity is somewhat preferred over one specific implementation of MOND.

Extending these tests, Banik et al. (2024) used DR3 data for 8611 wide binaries within 250 pc, with separations 2000–30 000 au. They integrated the binary orbits in a gravitational field that includes MOND's 'external field effect' but, crucially, admit line-of-sight contamination and undetected close binary companions to the component stars. They found that their best Newtonian model is preferred over MOND at 19σ confidence.

WHERE DOES this leave us with Gaia's present insights into wide binaries? Still with an incomplete understanding of their formation, although with the remarkable findings that, compared to models of a turbulent origin, the expected turbulent gas velocities appear to be matched by those of the newly formed stars.

Whatever the formation mechanism, Gaia DR3 confirms that the orbits of very wide binaries deviate from the Newtonian prediction. Proponents of a MOND-like explanation have confirmed their earlier conclusions, while undetected close binary companions perhaps still offer a more conventional explanation.

194. Cosmological simulations and Gaia

In the early 1980s, when Hipparcos was accepted by ESA, space astrometry was far from the mainstream of astronomy. Its goal was to extend distance measurements to ~100 pc, determine luminosities in the upper parts of the Hertzsprung–Russell diagram, and map distances and motions in the solar neighbourhood.

It was a bold choice by ESA's advisory committees, but somewhat reluctantly accepted more widely. Some doubted its feasibility, and further advances below a milli-arcsec seemed implausible. The rich complexities of dynamical phenomena across the Galaxy's vastness were unknown, beyond reach, and never discussed.

How the landscape has changed in the decades since! In the curious way that science advances across many fronts in parallel, Gaia is, today, central to problems being tackled in stellar evolution, in exoplanet science, in solar system studies, and in cosmology.

In the last of these fields, impressive developments in numerical simulations are now guiding many of Gaia's advances: amongst them, interpreting stellar streams, modelling resonant motions, characterising the dynamics of local group galaxies, and comprehending the multiple manifestations of past satellite interactions.

Here, I will briefly outline these cosmological simulations. I will then show how Gaia is confirming many of their detailed predictions, and helping in their interpretation and presumably their future development.

Over the past 20 years, very large massively parallel N-body simulations, resting on the Big Bang ΛCDM paradigm, have been developed to investigate how dark and baryonic matter structures have evolved over time (Vogelsberger et al., 2020).

My introduction here will refer only to CDM simulations on the largest cosmological scales, but they are also used to study effects on smaller systems. From a substantial literature I could mention, as examples, galaxy and halo formation (e.g. Navarro & White, 1994), black hole accretion (e.g. Di Matteo et al., 2008), and the first protostars (e.g. Yoshida et al., 2008).

The first of the very large-scale projects was the Millennium Simulation (Springel et al., 2005). The simulation followed the growth of dark matter structures from $z = 127$ to the present. It used 2160^3 particles, each representing $10^9\,M_\odot$ of dark matter, within a cube of side 700 Mpc. It occupied the main supercomputer of the Max Planck Society, Garching, for more than a month.

Successive versions incorporated improved parameters and input physics. Millennium II (Boylan-Kolchin et al., 2009) simulated a smaller volume with the same number of particles, each of $7\times10^6\,M_\odot$. Millennium XXL (Angulo et al., 2012) used a cube of side 4 Gpc, with 6720^3 particles each representing $7\times10^9\,M_\odot$.

Other such simulations now include Bolshoi (Klypin et al., 2011; Trujillo-Gomez et al., 2011), Eris (Guedes et al., 2011), the widely-cited EAGLE (Schaye et al., 2015), HESTIA (for the Local Group, Libeskind et al., 2020), and NewHorizon (Dubois et al., 2021), along with 'zoomed-in' developments such as ARTEMIS (Font et al., 2020).

I will say more on just one of these, Illustris (Vogelsberger et al., 2014a; Genel et al., 2014). This simulation started 12 Myr after the Big Bang, evolved over 13 Gyr, and used 12 billion resolution elements in a cube of side 100 Mpc. It generates (for example) elliptical and spiral galaxies, galaxy clusters, the distribution of hydrogen on large scales, and the metal and hydrogen content of galaxies on small scales (Vogelsberger et al., 2014b).

Galaxy formation processes include, amongst others, stellar evolution and feedback, gas recycling, chemical enrichment (following nine elements independently), and black hole growth and mergers (e.g. Vogelsberger et al., 2013; see also Somerville & Davé, 2015).

The Illustris framework has been used for other derivatives, specifically: Auriga (high-resolution simulations of Milky Way-like dark matter halos; Grand et al., 2017); IllustrisTNG ('The Next Generation'; Pillepich et al., 2018); Thesan (for the reionisation epoch; Kannan et al., 2022); MillenniumTNG (for the massive end of the halo mass function; Hernández-Aguayo et al., 2023); and TNG-Cluster (for galaxy clusters; Nelson et al., 2024).

I WILL GIVE a few of the growing number of examples where these simulations are guiding interpretation of the Gaia data. And I refer to essay 118 for some specific words on the ΛCDM 'missing satellites' problem, the 'core–cusp' problem, the 'too-big-to-fail' problem, and the 'plane of satellites' problem.

MAJOR MERGERS: it is now accepted that the Gaia–Sausage–Enceladus (GSE) merger played a key role in the formation of our Galaxy's inner stellar halo (Helmi et al., 2018; Gallart et al., 2019), as well as the disk (Haywood et al., 2018; Xiang & Rix, 2022). Are these discoveries supported by cosmological simulations?

Dillamore et al. (2022b) showed that about one-third of galaxies from the ARTEMIS simulations contain accreted stars on highly radial orbits, similar to the GSE event. The major mergers also result in disk rotation, and changes in shape and orientation of their dark matter halos. Early mergers result in retrograde stars, analogous to the 'splash' or 'plume' feature also discovered with Gaia (Di Matteo et al., 2019; Belokurov et al., 2020b).

Khoperskov et al. (2023a) similarly analysed six M31 and Milky Way analogues from the HESTIA simulations of the Local Group. They found that all experienced between one to four mergers with stellar mass ratios between 0.2–1, with five them occurring 7–11 Gyr ago. The most massive mergers result in a sharp increase in the orbit eccentricity of disk stars of the main progenitor.

THE CENTRAL BAR: Gaia is providing new insights into the morphology and dynamics of our Galaxy's central bar (essays 112 and 196). Arising as a consequence of stellar orbits, are they also influenced by past mergers?

IllustrisTNG, NewHorizon, and EAGLE all confirm the emergence of bars, but with a large variation in bar fractions, ranging from 5–55% (Cavanagh et al., 2022).

Fragkoudi et al. (2020) showed that, in their Auriga simulations, galaxies which best reproduce the chemodynamical properties of the Milky Way bulge have quiescent merger histories since $z \sim 3.5$. Their last major merger was more than 12 Gyr ago, with any subsequent mergers having a stellar mass ratio of 1:20 or lower.

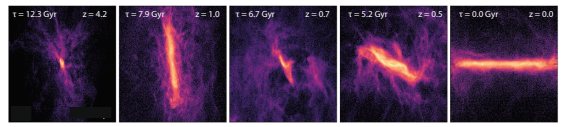
In turn, this suggests an upper limit of a few percent for the mass ratio of the GSE merger event. They inferred that the Milky Way has had an 'uncommonly quiet merger history', and hence an essentially *in situ* bulge.

EVOLUTION OF THE MILKY WAY: as I described in essay 190, Chandra et al. (2024) used [Fe/H] and [α/Fe] estimates for 9.9 million Gaia red giants to characterise their angular momentum as a function of metallicity. Taking this as a proxy for age, they identified three distinct evolutionary phases: a disordered/chaotic protogalaxy, a hot old high-α disk, and a cold young disk with more ordered and circular orbits.

Does this proposed 'three-phase evolution' of the Milky Way find any support from these large-scale cosmological simulations?

Semenov et al. (2024) had already selected a representative sample of 61 Milky Way-like galaxies from the TNG50 simulations (the highest-resolution box from IllustrisTNG). Of these, 11 matched the 'early spin-up' previously inferred by Belokurov & Kravtsov (2022).

Chandra et al. (2024) showed that halo #519311, one of the earliest of their spin-up galaxies, exhibits a 'three-phase' structure in the orbit circularity versus metallicity space 'remarkably similar' to the Milky Way. This halo also experiences a gas-rich major merger 8 Gyr ago, albeit slightly later than the estimated GSE merger. The merger adds a large amount of low-metallicity gas and angular momentum, from which the kinematically cold low-α stellar disk is subsequently born.



Chandra et al. (2023)

MANY MORE Gaia studies are calling on these cosmological simulations to assist their interpretation.

I counted more than 20 other papers between 2018 and mid-2024 using EAGLE or Illustris to further the understanding of the GSE-like accretion event, where I will simply list some of the most recent (viz. Wu et al., 2022; Belokurov et al., 2023; Dillamore et al., 2023a; Carollo et al., 2023; Khoperskov et al., 2023a; 2023b; Rey et al., 2023; Lane & Bovy, 2025; Carrillo et al., 2024).

VARIOUS STUDIES EMPLOY these simulations in interpreting the improved globular cluster orbits (Chen & Gnedin, 2022; Ishchenko et al., 2023a; 2023b; Chen & Gnedin, 2024; Ishchenko et al., 2024), and the improved orbits of the dwarf spheroidals (Pardy et al., 2020; Pawlowski & Kroupa, 2020; Borukhovetskaya et al., 2022; Martínez-García et al., 2023).

THEY ARE ALSO being used in discussions of the mass of M31 (Patel & Mandel, 2023), warping in the orbits of Cepheids (Dehnen et al., 2023), Milky Way analogues (Grand et al., 2018), infall times for Local Group galaxies (Barmentloo & Cautun, 2023), hypervelocity stars and the Galactic escape speed (Deason et al., 2019), the Milky Way mass profile and halo mass (Li et al., 2020), halo anisotropy (Bozorgnia et al., 2020), and the cosmological core–cusp problem (Wang et al., 2022).

AND, OF COURSE, more accurate and complex data sets, more stringent tests, and greater clarity on the accuracy and fidelity of the ΛCDM parameterisation, will come with future Gaia data releases.

195. Imaging of the Galaxy bar

I DISCUSSED SOME aspects of our Galaxy's central bar in essay 112: how bars in general are believed to originate, the evidence for a bar in our own Galaxy, and inferences being made from kinematic studies with Gaia.

These included estimates of the bar's length, its 'pattern speed', why it affects motions in the solar neighbourhood via resonances, and evidence that the bar is slowing down, over cosmological time scales, due to dynamical friction dominated by the dark matter halo.

Here, I will say more about the attempts (and difficulties) of elucidating its morphology, and include a surprising demonstration of Gaia's ability to image it.

OCCUPYING OUR Galaxy's inner few kpc is a dense concentration of mostly old stars in a roughly spheroidal 'bulge'. Departures from circularity were inferred from 21-cm HI observations by de Vaucouleurs (1964), and its triaxial structure first discerned more directly from 2.4 μm observations (Blitz & Spergel, 1991), and near-infrared COBE data (Weiland et al., 1994; Dwek et al., 1995; Freudenreich, 1998).

Many subsequent efforts have been made to discern its detailed structure, as a step to understanding its origin. Work focused on the bar's morphology includes use of the colour–magnitude diagram of red clump stars from OGLE (Stanek et al., 1994; Rattenbury et al., 2007); and using infrared star counts from 2MASS, DENIS, ISO, GLIMPSE, UKIDSS, and VVV (López-Corredoira et al., 2001; Picaud & Robin, 2004; Babusiaux & Gilmore, 2005; Benjamin et al., 2005; Cabrera-Lavers et al., 2008; Wegg et al., 2015; Simion et al., 2017).

The detailed morphology still remains unclear, e.g. whether it is peanut-shaped, or whether there are additional components. Estimates still vary widely concerning its axis ratio, its half-length (1–5 kpc), and its orientation with respect to the Galactic centre ($10-45°$).

Its stellar population content, and its formation and evolutionary history, also remain poorly constrained. Two or more overlapping populations might exist, complicating estimates of its shape and orientation (e.g. Martinez-Valpuesta & Gerhard, 2011).

IN ONE SPECIFIC model, which I will refer to later, Robin et al. (2012b) made a two-component fit to star counts, colour–magnitude diagrams, and metallicities. Their principal component was a triaxial ellipsoid with scale lengths 1.46/0.49/0.39 kpc, at an orientation of 13°.

IN THEIR REVIEW, Bland-Hawthorn & Gerhard (2016) emphasised the difficulties in separating the bar and bulge, the inner disk and inner halo, and the possibility of a 'super-thin' contribution. They gave best estimates as: *stellar* mass ($7\pm1\times10^9\,M_\odot$), half-length (5.0 ± 0.2 kpc), scale height (180 pc), orientation ($28-33°$), and an age of 6–7 Gyr (see also Sormani et al., 2022a).

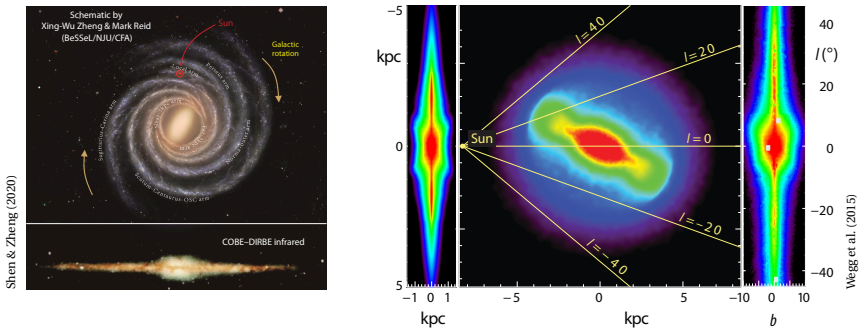
As they stated at the time, its pattern speed ($43\pm9\,\mathrm{km\,s^{-1}\,kpc^{-1}}$) and corresponding corotation radius (4.5–7 kpc), derived mainly from numerical and hydrodynamical simulations, also remain poorly determined.

MUCH NEW INSIGHT has come from recent kinematic measurements and associated modelling using the Gaia data. I will give just a brief summary of some of the findings I mentioned in essay 112.

From Gaia DR2 and VVV–VIRAC, Sanders et al. (2019) and Clarke et al. (2019) showed that differential rotation between the double peaks of the magnitude distribution of red clump giants confirms the X-shaped nature of the bar–bulge identified by Wegg & Gerhard (2013), and consistent with a pattern speed of $37.5\,\mathrm{km\,s^{-1}\,kpc^{-1}}$.

Bovy et al. (2019) used APOGEE DR16 with Gaia DR2, while Leung et al. (2023) updated the analysis using APOGEE DR17 with Gaia EDR3, to identify the minimum in rotational velocity, and the quadrupole signature in radial velocity, expected for stars orbiting in a bar. They estimated a pattern speed of $40.1\pm1.8\,\mathrm{km\,s^{-1}\,kpc^{-1}}$.

Other Gaia-based studies have given other estimates of the bar length and pattern speed (Clarke & Gerhard, 2022; Drimmel et al., 2023; Lucey et al., 2023). A remarkable if preliminary conclusion, based on models of resonance sweeping, is that the bar is decelerating, an effect attributable to dynamical friction of a dark matter halo (Chiba et al., 2021; Chiba & Schönrich, 2021).



THE LEFT FIGURE ABOVE, from Shen & Zheng (2020), is an artist impression of the Milky Way viewed face-on (top), with the infrared COBE–DIRBE image showing the box-like bar (bottom). The right figure shows the bar and bulge projections derived from near-infrared star counts by Wegg et al. (2015).

More details of the bulge–bar structure, interpretation of the X-shaped feature, and of associated dynamical models, are given by Shen & Zheng (2020).

And as summarised in their review of the chemo-dynamical history of the bulge, Barbuy et al. (2018) noted that *'Despite impressive progress, we do not yet have a successful fully self-consistent chemo-dynamical bulge model in the cosmological framework, and we will also need a more extensive chrono-chemical-kinematic 3D map of stars to better constrain such models'*.

THINGS HAVE progressed with Gaia since then, but here I will look only at one study which has essentially been able to image the bulge, at least more directly than reconstructions based on number counts.

Anders et al. (2019) used the `StarHorse` code on a combination of Pan–STARRS1, 2MASS, and AllWISE photometry with Gaia DR2 parallaxes to derive stellar parameters, distances, and extinctions for 137 million stars with $G \leq 18$ mag. Resulting median precisions were 5% in distance, 0.20 mag in V-band extinction, and 245 K in $T_{\rm eff}$ for $G \leq 14$, degrading at fainter magnitudes.

`StarHorse` (Santiago et al., 2016; Queiroz et al., 2018) finds the (Bayesian) posterior probability over a grid of stellar evolutionary models, distances, and extinctions, given a set of observations plus a number of priors: here, the initial mass function, density laws for the thin disk, thick disk, bulge, and halo, as well as broad metallicity and age priors for each component.

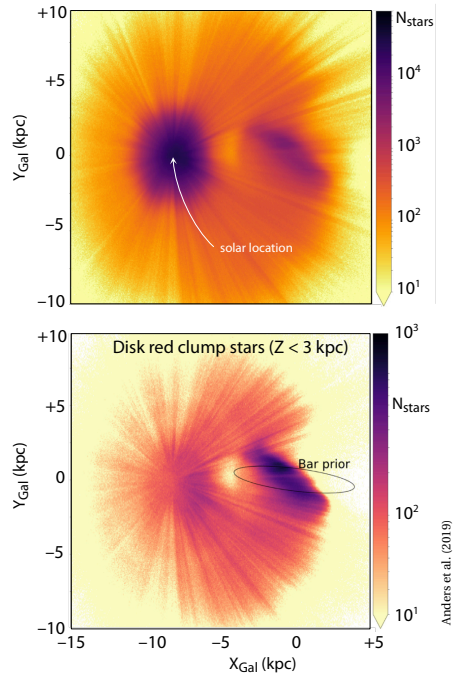
The upper figure opposite shows the face-on projection of the stellar density distribution of their final sample in Galactic XY coordinates. The lack of stars towards the inner Galaxy is attributed to identified effects of the high dust extinction and the adopted quality criteria.

Especially noteworthy is the presence of a stellar overdensity coinciding with the expected position of the Galactic bar, inclined by about 40° with respect to the solar azimuth, and with a semi-major axis of ~4 kpc.

The presence of the bar in the Gaia DR2 data is even more prominent when focusing only on the red clump stars (lower figure). Interestingly, its revealed shape and inclination are significantly different from their adopted prior, which was given by the Galactic bulge–bar model derived by Robin et al. (2012b), and noted above. They also identified a kinematic imprint of the coherent motion of the Galactic bar in the proper motions.

That Gaia is probing stellar populations in the Galactic bulge and beyond is, at first sight, surprising. They reason that their distance estimates are not more *precise* than those prescribed by, e.g., Bailer-Jones et al. (2018), using which the bar is barely evident. However, at large distances, their values (they argue) are more *accurate*, due to the use of more informative Galactic priors.

Further discussion of the velocities and chemo-dynamical implications, based on Gaia EDR3, were reported by Queiroz et al. (2021), who also interpreted their results in the context of numerical simulations of barred galaxies (e.g. Debattista et al., 2017; Bovy et al., 2019; Carrillo et al., 2019; Fragkoudi et al., 2020).



196. The rotation curve of our Galaxy

A GALAXY'S ROTATION CURVE describes the circular rotation speed as a function of radial distance. It provides important constraints on the galaxy's mass distribution. Since the work of Rubin et al. (1978), the flat rotation curves observed in spiral galaxies, in contrast to those expected from the *observed* mass distribution, is generally attributed to the presence of dark matter.

For our own Galaxy, the rotation curve similarly provides strong constraints on the mass distribution of its major structural components (bulge, thin and thick disk, and halo), including the effects of dark matter (e.g. Sofue et al., 2009; Salucci et al., 2010; Weber & de Boer, 2010; Sofue, 2012; Wang et al., 2015; Wang et al., 2020).

DEFINING OUR Galaxy's rotation curve is not straightforward. In the inner regions, it has been mainly determined from radio observations of the gas in the spectral lines of H I and CO (e.g. Burton & Gordon, 1978; Clemens, 1985; Fich et al., 1989; Levine et al., 2008).

Beyond the solar circle, it can be derived from the kinematic motions of suitable tracer populations, including H II regions (Fich et al., 1989; Brand & Blitz, 1993), planetary nebulae (Schneider & Terzian, 1983; Amaral et al., 1996; Maciel & Lago, 2005), classical Cepheids (Pont et al., 1997; Gnaciński, 2019; Mróz et al., 2019a), blue horizontal branch stars (Xue et al., 2008; Deason et al., 2012; Kafle et al., 2012), red clump giants (Bovy et al., 2012; López-Corredoira, 2014; Huang et al., 2016), and masers (Honma et al., 2012; Reid et al., 2014; Gromov & Nikiforov, 2021).

This has been extended into the outermost regions of the Galaxy using the dynamics of globular clusters or dwarf galaxies (Spitler & Forbes, 2009; Watkins et al., 2010; Boylan-Kolchin et al., 2013; Li et al., 2017), most recently using data from Gaia DR2 (Li et al., 2020).

A compilation of 2780 measurements between 3–20 kpc, `galkin`, is given by Pato & Iocco (2017). And although often subject to large random and systematic errors, the general consensus is that the rotation curve is roughly constant at $R \sim R_0$, and fairly flat or with a slow decline at larger radii, at least out to $2R_0$, implying the presence of invisible or dark matter in the outer parts.

NEW DETERMINATIONS are being derived based on Gaia data, exploiting the much larger number of stars, much greater accuracies in distances and proper motions, and star classification.

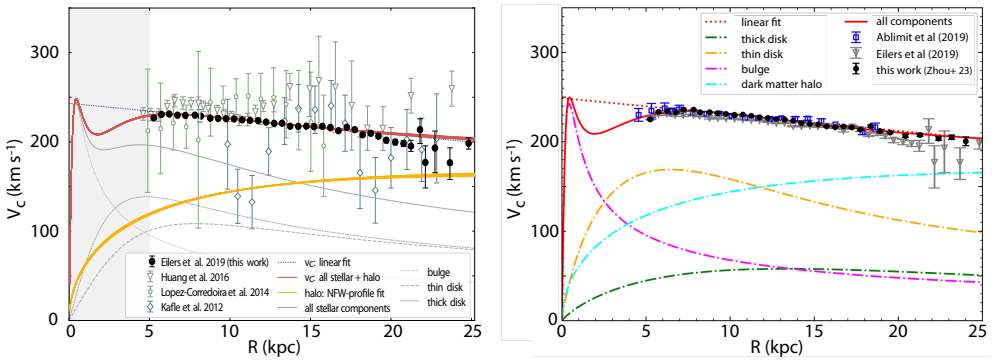
FROM AN APOGEE DR14 sample of red giants from Hogg et al. (2019), and Gaia DR2 data, Eilers et al. (2019) used 23 000 thin disk stars to determine the rotation curve between 5–25 kpc, based on a Jeans model, and assuming an axisymmetric gravitational potential.

The Gaia-based error bars of Eilers et al. (2019) are significantly smaller than previous determinations, but still result in large uncertainties beyond 20 kpc. This is clearly seen as the filled circles in their Fig. 3 (shown over, top left). This figure also shows the contributions due to the thin and thick disks (assuming the model profiles of Miyamoto & Nagai, 1975), the bulge (assume the spherical potential of Plummer, 1911), as well as their fit to a Navarro–Frenk–White (NFW) dark matter halo profile (Navarro et al., 1997).

THEIR *'most precise measurements of the circular velocity to date'* shows a gently declining slope beyond 5 kpc, -1.7 ± 0.1 km s^{-1} kpc^{-1}, in reasonable agreement with the classical Cepheid determination by Mróz et al. (2019a), but less shallow than some earlier studies which have suggested a flat circular velocity curve (e.g. Bovy et al., 2012; Bovy & Rix, 2013; Reid et al., 2014).

The gradient is important because disk galaxies in the local Universe typically show a flat or even *increasing* rotation curve (e.g. Rubin et al., 1980; Sofue et al., 1999), while galaxies with declining curves have only been reported at higher redshift. This effect seen in the early Universe has been attributed to baryons efficiently condensing at the centres of dark matter halos at a time when gas fractions were higher and dark matter less concentrated (Genzel et al., 2017; Lang et al., 2017).

The rotation curve derived by Eilers et al. (2019) was used by de Salas et al. (2019) to evaluate the sensitivity of estimates of the local dark matter density to the assumed details of the dark matter halo, and the distribution of matter in the baryonic disk.

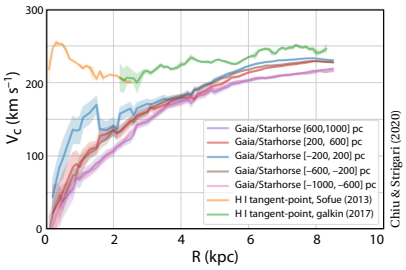


Gaia-based determination of the Galaxy's rotation curve from Eilers et al., 2019 (left) and Zhou et al., 2023b (right)

FOR THE rotation curve within R_0, the assumption of circular orbits has underpinned use of the so-called 'tangent-point method' in interpreting the gas kinematics. This assumes that the observed radial velocity extremum along any line-of-sight is at the tangent point, allowing calculation of both Galactocentric radius and circular velocity (van de Hulst et al., 1954; Sofue, 2013).

Specifically, and pre-Gaia, Chemin et al. (2015) used numerical simulations to predict that, in part due to the effects of the bar, the tangent-point method overestimates the inner rotational velocity, and instead reflects local motions, concluding that *'the quest to determine the innermost rotation curve of the Galaxy remains open'.*

Indeed, a rather radical finding along precisely these lines was reported by Chiu & Strigari (2020). They used 6 million stars from Gaia DR2 to determine the rotation curve within the solar circle, which they compared with that derived from gas kinematics using the tangent-point method.



They found significant differences between the two methods: while the tangent-point curve is mostly in the range 200–250 km s^{-1}, that based on stellar kinematics increases to ~150 km s^{-1} at 2 kpc, rising to a plateau at 6–7 kpc. The peak using the tangent-point method at 1 kpc is absent in the curve based on stellar kinematics.

USING 54 000 thin disk red giants, selected from the APOGEE and LAMOST surveys and using astrometry from Gaia EDR3, Zhou et al. (2023b) constructed the rotation curve from 5–25 kpc. With a 2–3 times larger sample than Eilers et al. (2019), their rotation curve also shows a weak decline beyond 5 kpc, with a similar gradient of $-1.83 \pm 0.02 \pm 0.07$ km s^{-1} kpc^{-1}, a circular velocity at R_0 of $234.04 \pm 0.08 \pm 1.36$ km s^{-1}, and a significantly better accuracy beyond 20 kpc (see figure, top right).

They used this to construct a mass model for the Galaxy, yielding a mass of the dark matter halo $M_{200} = (8.05 \pm 1.15) \times 10^{11} M_\odot$, with $R_{200} = 192.37 \pm 9.24$ kpc, and a local dark matter density 0.39 ± 0.03 GeV cm^{-3}.

FURTHER IMPROVEMENTS came with Gaia DR3. Wang et al. (2023b) extended the mapping to 30 kpc, and confirmed that the rotation curve indeed shows a significant decline of ~50 km s^{-1} between 15–30 kpc. The rotation curve (and azimuthal velocity) also presents a dependence on height above the disk for $R < 15$ kpc.

Ou et al. (2024) used APOGEE DR17 spectra combined with Gaia DR3, 2MASS, and WISE photometry, to construct the rotation curve out to 30 kpc. From more than 30 000 disk red giant branch stars they confirmed the previous findings. But they found an even faster decline beyond ~25 kpc, better matched to a 'cored Einasto profile' (Einasto, 1965) than the generalised Navarro–Frenk–White profile. It yields a significantly lower halo virial mass, in tension with mass measurements from globular clusters, dwarf satellites, and streams.

Jiao et al. (2023) compared the methods used in these Gaia-based determinations (Zhou et al., 2023b; Wang et al., 2023b; Ou et al., 2024). They also concluded that there is a sharp decrease in circular velocity of ~30 km s^{-1} between 19.5–26.5 kpc, matching a *Keplerian* decline, and so implying the absence of significant mass beyond 20 kpc (Zobnina & Zasov, 2020). This would imply a total Galaxy mass $2.0 \times 10^{11} M_\odot$, 4–5 times smaller than currently favoured estimates (essay 93).

IT IS too soon to conclude that the dark matter halo terminates at around 20 kpc, with all the implications that this would involve, and with the other inconsistencies that this might raise (e.g. Jiao et al., 2023, §6).

Meanwhile, work continues in probing other effects that might be complicating the determination and interpretation of the rotation curves, including the assumptions underlying use of the Jeans equation, the effects of coordinate transformations, and alternative models of (MOND-like) gravity (e.g. Chrobáková et al., 2020; Petersen & Frandsen, 2020; Natalia Cisneros et al., 2023).

197. The Gaia Sausage-Enceladus stream

THE TOPIC of stellar streams, key in understanding the Milky Way's formation, has been transformed by Gaia. I have written several essays on them already: the Enceladus stream (essay 15), other stream discoveries (71 and 156), and their use as probes of black holes (176) and ΛCDM-type sub-halos (184). More than 100 are now known, and in this and the following two essays, I will focus on three: Enceladus, Sagittarius, and Cetus.

IN ESSAY 15 (April 2021) I gave a top-level picture of what was called the 'Enceladus' stream in the discovery paper by Helmi et al. (2018), and which was described as 'sausage-shaped' in velocity-space in the independent discovery paper(s) by Belokurov et al. (2018) and Myeong et al. (2018c). It is now known as the Gaia Sausage–Enceladus, or GSE, stream. It is not, let me note, the same as the 'Helmi stream', discovered in the Hipparcos data by Helmi et al. (1999).

Refereed papers mentioning GSE in their *abstract* is rising: 10 in 2019, 25 in 2020, 35 in 2021, 40 in 2022, 40 in 2023, and 30 so far in 2024. This underscores the stream's importance in revealing the assembly history of the Milky Way – structurally, dynamically, and chemically, and its use as a Rosetta Stone for understanding our Galaxy's formation in a cosmological context.

I will give only some key points in our understanding of the GSE stream, and keep the detailed discovery steps, and references, to a minimum. The pre-Gaia understanding of the Galaxy's halo and its formation, and much more on what is known about GSE, can be found in the recent review by Deason & Belokurov (2024).

OVER THE PAST 20–30 years, through a combination of observations, theory, and simulations, aided by N-body simulations of the large-scale structure of the ΛCDM Universe, it became clear that the halo should be the graveyard of many other galaxies, gravitationally captured by the Milky Way galaxy over its lifetime, as in this animation. Dynamical timescales in the outer halo are longer than the age of the Universe, so it should preserve a fossil record of these ancient encounters.

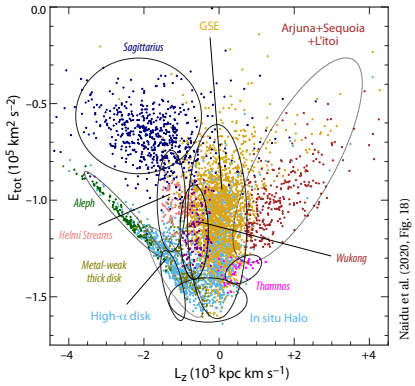
OVER BILLIONS of years, galaxies gravitationally captured by the Milky Way slowly spiral inwards, the more distant (Sagittarius and the Magellanic Clouds) showing only early signs of capture and disruption, the most ancient on their second or subsequent orbital approach. While they steadily dissolve in 6D phase-space, i.e. progressively scattering through the halo in terms of their positions and velocities, some preserve clear signatures of their common origin through a clustering of their abundance patterns *and* their orbital properties.

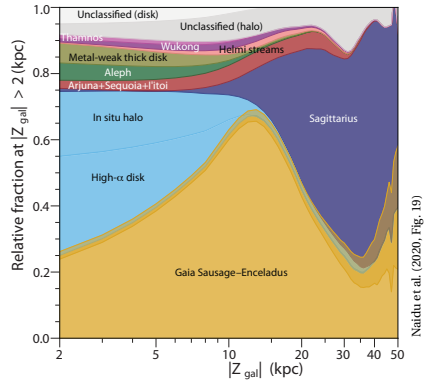
More precisely, while they may be widely scattered on the sky, their common origin is encoded in their similar 'integrals of motion', e.g. in their total energy, $E_{\rm tot}$, and their projected orbital angular momentum (Helmi & de Zeeuw, 2000; Binney & Tremaine, 2008, §3.1.1).

THERE IS NOT SPACE here to elaborate on the many post-discovery studies of the GSE stream which have been piecing together its properties, dynamics and origin, nor on the many other halo streams that have now been identified with Gaia (`Galstreams` lists 126 tracks, in 95 distinct stellar streams, as of March 2022).

But I will pick out the study by Naidu et al. (2020), where some of their conclusions were particularly striking, and which serves as a context. They used 5684 giants with distances $|Z| > 2$ kpc from the Galactic plane, combining spectroscopy and radial velocities from the MMT–H3 survey with Gaia DR2 astrometry, and they identified several clear clusterings in terms of their chemistry ([Fe/H] and [α/Fe]), and their position in energy–angular momentum space, $E_{\rm tot} - L_z$.

They identified various prominent structures and streams already reported in previous studies, viz. the high-α disk, the *in situ* halo, the Sagittarius dwarf galaxy, along with the GSE, Helmi, Sequoia, and Thamnos streams. They also identified four other streams (both prograde and retrograde) with distinct chemistries and orbits: Aleph, Arjuna, I'itoi and Wukong. The top left figure (over) shows all these clusterings, also distinguishing those on radial orbits ($J_z < J_R$, normal font) from those on polar, circular orbits ($J_z > J_R$, italic font).





THE GSE MERGER, some ~9.5 Gyr ago (at $z \lesssim 3$), had a major effect on the halo. Naidu et al. (2020) found that, at $|Z| \gtrsim 15$ kpc, the debris of just two massive accreted dwarfs ($10^8 - 10^9 M_\odot$) comprise 80% of the halo, together explaining the relatively high overall halo metallicity, [Fe/H] ≈ -1.2. They attributed 95% of their sample to one of their dozen identified structures (their inferred composition versus Z is shown, above right), pointing to a halo built *entirely* from accreted dwarfs.

Within 25 kpc, GSE, with [Fe/H] = –1.2, was a radial merger that now dominates the local metal-poor halo. The 'head-on' collision resulted in rapid phase mixing, evident today only in its clustered integrals of motion.

Beyond 25 kpc, the halo is dominated by the Sagittarius (Sgr) dwarf galaxy and stream, [Fe/H] = –1.0, one of the first identified halo streams (Ibata et al., 1994). Here, both the residual galaxy core, as well as two vast preceding and trailing streams, are clearly discernable.

Later simulations by Naidu et al. (2021) suggest a GSE mass at infall of $5 \times 10^8 M_\odot$ in stars and $2 \times 10^{11} M_\odot$ in dark matter, that it arrived on a retrograde orbit, with Arjuna being its retrograde debris, and Sequoia and I'itoi perhaps also being stripped from its outer regions.

MERGERS ALSO HAVE a significant effect on the disk, a prominent example being the Gaia phase-space spiral, or 'snail', attributed to phase-mixing through past pericentric passage of the Sgr dwarf (Antoja et al., 2018).

A major local disk component is the 'high-α disk' and its high-eccentricity tail (or '*in situ* halo'), comprising 15% of the local Galaxy. Naidu et al. (2021) argued that the continuous eccentricity distribution supports a picture in which a primordial high-α disk was disturbed and 'heated', possibly by the GSE merger. A similar origin was attributed to the 'splash' population (Belokurov et al., 2020b; Deason & Belokurov, 2024, §4.2).

But studies are still ongoing, and involve the footprints of past mergers in the form of 'energy wrinkles' and 'phase-space folds' (aka chevrons or caustics). While Belokurov et al. (2023) again attributed these to the GSE merger, other much younger events have also left their own mixed contributions (e.g. Donlon et al., 2022; Donlon & Newberg, 2023; Donlon et al., 2024).

THE DEMONSTRATED correlation between the combined mass of a galaxy's globular clusters and its total mass (e.g. Dornan & Harris, 2023) implies that while many tidal debris remnants are expected to be detected in the stellar halo, only the most massive would be accompanied by globular clusters on similar orbits. This opens up another fascinating aspect of the properties and origin of the many halo streams now known.

In the case of GSE, Myeong et al. (2018c) showed that amongst our Galaxy's *accreted* clusters, a tight grouping have very similar orbital actions (with small pericentres $\lesssim$2 kpc, and large apocentres $\gtrsim$10 kpc), consistent with clusters associated with the GSE merger event.

TO CONCLUDE this outline, I refer again to Deason & Belokurov (2024) for a more detailed discussion of the progenitor mass, the event timing, the effects on our Galaxy disk, and the implications for understanding the ongoing Sagittarius and Magellanic Cloud mergers.

And I should also stress that the large-scale ΛCDM-based simulations of structure formation are also guiding interpretation, and providing support, to the picture of Milky Way-type galaxies which have undergone major GSE-like mergers (e.g. Renaud et al., 2021; Sestito et al., 2021; Dillamore et al., 2022b; Khoperskov et al., 2023a; Orkney et al., 2023; Semenov et al., 2024).

THERE IS much more that we would like to know, and future Gaia releases should provide much greater clarity, and many more answers.

Deason & Belokurov (2024) concluded their review with a number of open questions, amongst them:
* does the luminosity function of destroyed dwarfs down to low masses agree with the ΛCDM predictions?
* how instrumental was GSE in the Galaxy's evolution: when did it merge, how much material did it bring, and how does it link to the disk, bulge, and bar formation?
* why did the disk form so early, and is this reproduceable in cosmological simulations?
* what is the origin of the stellar halo spin, and does this relate to the spinning up of the Milky Way disk?
* what happened to the satellites associated with the pre-infall massive dwarf galaxies like the GSE?

198. The Sagittarius stream

In my previous essay, I summarised our knowledge of the Gaia Sausage–Enceladus halo stream. This was discovered in the Gaia DR2 data, and is now known to dominate the stellar halo within 25 kpc. Beyond that, and out to some 50 kpc, the halo is significantly populated by the Sagittarius stream, the subject of this essay.

Debris of these two massive ($10^8 - 10^9 M_\odot$) accreted dwarfs together appears to comprise 80% of the halo (Naidu et al., 2020). Along with the many other streams discovered over the past few years, they confirm the picture of hierarchical structure formation in the Universe, parameterised by the ΛCDM cosmological model.

The Sagittarius dwarf galaxy is a nearby satellite of our own, discovered by Ibata et al. (1994). It was soon recognised as being tidally distorted, and inferred to have completed several close orbits around the Milky Way (Ibata et al., 1995; Mateo et al., 1996). With a mass of $4 \times 10^8 M_\odot$, and distance ~24 kpc, it is a most striking example of galaxy disruption and ongoing accretion.

Subsequent studies, initially with 2MASS and SDSS, progressively revealed an elongated and weakly rotating central core, with tidally stripped stars forming two long tidal tails, both bifurcated, tracing an orbit almost perpendicular to the Milky Way disk, with much other complexity (Yanny et al., 2000; Newberg et al., 2002; Majewski et al., 2003; Newberg et al., 2003; Martínez-Delgado et al., 2004; Belokurov et al., 2006; Newberg et al., 2007; Łokas et al., 2010.; Koposov et al., 2012; Slater et al., 2013). The tidal debris traces out an orbit with an apogee of 50 kpc for the leading arm, and 100 kpc for the trailing (Belokurov et al., 2014; Hernitschek et al., 2017).

Other work has aimed to further characterise the system's orbit, mass, chemistry, and velocity dispersion (e.g. Frinchaboy et al., 2012; Majewski et al., 2013; Gibbons et al., 2017; Navarrete et al., 2017), as well as its use in determining the Galaxy potential (Law & Majewski, 2010b; Price-Whelan et al., 2016; Fardal et al., 2019). The two most distant known halo stars, at 200 kpc (Bochanski et al., 2014), appear consistent with the simulations of Sgr's maximal extent (Dierickx & Loeb, 2017), and perhaps with the 'edge' of our Galaxy (Deason et al., 2020).

With N-body simulations having to satisfy these many, detailed, observational constraints, a key question is whether the present elongated, prolate, bar-like shape of the core is a result of tidal forces on an initially spherical galaxy (Johnston et al., 1995; Ibata et al., 1997; Helmi & White, 2001; Vasiliev & Belokurov, 2020), or attributable to a more disk-like progenitor (Peñarrubia et al., 2010; Łokas et al., 2010; del Pino et al., 2021).

Łokas (2024) selected an Sgr analog from the IllustrisTNG simulations (essay 194) to demonstrate how such a dwarf, with initial mass $> 10^{11} M_\odot$, evolves around a Milky Way-like host on a tight orbit over seven pericentre passages, with a period of 1 Gyr. At the second pericentre, the disk transforms into a bar, and the bar-like shape is preserved thereafter. Strong mass loss leaves a dwarf with final mass $\lesssim 10^9 M_\odot$. The gas is lost, and star formation ceases, at the third pericentre. The dwarf then retains a bar-like shape, small rotation, and a metallicity gradient, mirroring the observations.

Gaia is providing deep insight into this spectacular system. With DR2, investigations included identifying specific stellar types within the tidal arms, e.g. 400 O-rich AGB stars (Mauron et al., 2019), 164 M giants (Li et al., 2019), 6000 RR Lyrae stars with 5D phase-space coordinates (Ramos et al., 2020), and 3500 RR Lyrae with full 6D phase-space coordinates (Ibata et al., 2020a). The latter work also concluded that the global properties of the Sgr stream were still reasonably well reproduced by the earlier model of Law & Majewski (2010b).

DR2 was also used to: determine the mean proper motion over 2π rad along the stream, using 1500 intensity peaks comprising 100 000 stars (Antoja et al., 2020); to identify a low-metallicity population inferred to have originated in the stellar halo of the Sagittarius progenitor (Johnson et al., 2020a); to identify a faint globular cluster torn from the Ophiuchus stream by a close (5 kpc) passage of the Sgr core 100 Myr ago (Lane et al., 2020); and to define 260 000 members from which the mean proper motions provide the basis for a model in which the Sgr galaxy will be fully disrupted over the coming 1 Gyr (Vasiliev & Belokurov, 2020).

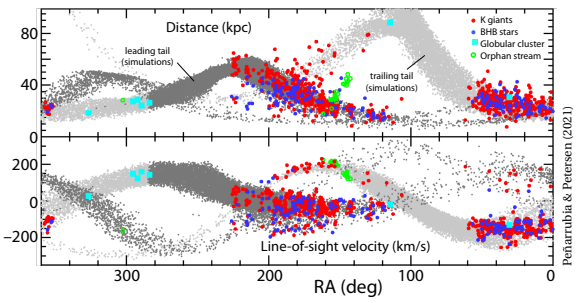
I WILL EXPAND a little on three of the Gaia DR2 results. First, del Pino et al. (2021) used 120 000 stars from DR2 to reveal a bar structure 2.5 kpc long, with the main body of the galaxy, strongly sheared by tidal forces, being a triaxial (almost prolate) ellipsoid. The inner core, of dimension $500 \times 330 \times 300\,\mathrm{pc}^3$ shows no net expansion, but it is rotating, mainly about its intermediate principal axis, with a maximum velocity of $4.13 \pm 0.16\,\mathrm{km\,s}^{-1}$.

In their paper entitled 'Tango for Three', Vasiliev et al. (2021) found a misalignment between the stream track and the proper motion directions in the leading arm, which they interpret as a time-dependent variation of the gravitational potential, in turn attributed to the recent passage of the LMC (of mass $1.3 \pm 0.3 \times 10^{11}\,M_\odot$). They argue that the stream cannot be modelled accurately in a static Galaxy potential, but rather the Milky Way is 'lurching' toward the massive infalling LMC, giving the Sgr stream its peculiar shape and kinematics.

In a very different application, Hayes et al. (2018) used the stream geometry to estimate the solar reflex velocity, and hence the velocity of the Local Standard of Rest, independently of an assumed value of R_0.

WITH THE availability of Gaia EDR3, still outstanding questions included the origin of the bifurcated tails (or parallel streams) in both northern (Belokurov et al., 2006) and southern (Koposov et al., 2012) Galactic hemispheres, although further progress in defining these was made with EDR3 (Ramos et al., 2022).

And while all models imply that Sgr has completed several orbits around the Milky Way, only tidal debris from the last pericentric passage had ever been detected. Peñarrubia & Petersen (2021) introduced a new method to identify clustering in angular momentum space, and identified 925 stars spanning 800° on the sky, thus wrapping the Galaxy *twice*.



In other work, 60 high-velocity stars, including 2 hypervelocity stars, originating from Sgr were identified by Li et al. (2022b). The effect of Sgr on other streams was investigated by Dillamore et al. (2022a), and on the Jhelum stream in particular by Woudenberg et al. (2023). Stream properties as a function of metallicity were investigated by Limberg et al. (2023), and using Gaia DR3 by Cunningham et al. (2024a). Searches for other candidates are also continuing with Gaia DR3 (Li et al., 2024).

THAT THIS SORT OF accretion process should also perturb the disk of the Milky Way had been predicted using test particle simulations (e.g. Ibata & Razoumov, 1998; Quillen et al., 2009; Gómez et al., 2012b). And some evidence of such disk perturbations preceded the Gaia data (e.g. Minchev et al., 2009; Gómez et al., 2012a; Widrow et al., 2012; Carlin et al., 2013; de la Vega et al., 2015; Schönrich & Dehnen, 2018; Laporte et al., 2018).

Gaia DR2 provided unambiguous evidence of such gravitational disturbances of the disk with the discovery of the so-called Gaia 'phase-space spiral', or phase-space 'snail' (essay 117), where ongoing 'phase mixing' from an out-of-equilibrium state is evident in the phase-space projection in the $Z - V_Z$ plane (Antoja et al., 2018).

The Gaia phase-space spiral (and some low latitude overdensities such as the Monoceros Ring) has been widely attributed to pericentric passages of Sgr (e.g. Binney & Schönrich, 2018; Khanna et al., 2019; Laporte et al., 2019a; Laporte et al., 2019b; Laporte et al., 2020; Vasiliev & Belokurov, 2020; Bland-Hawthorn & Tepper-García, 2021; Gandhi et al., 2022; Das et al., 2024).

Others suggest that it may have arisen from several smaller disturbances, rather than a single dominant one (e.g. Hunt et al., 2022; Tremaine et al., 2023).

A remarkable finding along similar lines, using Gaia DR2, was reported by Ruiz-Lara et al. (2020). They modelled the colour–magnitude diagram within 2 kpc of the Sun to identify three conspicuous and narrow episodes of enhanced star formation, estimated as having occurred 5.7, 1.9 and 1.0 Gyr ago. They found that these episodes coincide with modelled Sgr pericentre passages, with the perturbations from Sgr repeatedly triggering major episodes of local star formation.

A SUBJECT WITH ITS own substantial literature is the presence of globular clusters within the Sagittarius system. Ibata et al. (1994) had already noted that M54 (NGC 6715) lies in its densest region. Later work suggested that others probably formed in its gravitational potential well, and have been stripped from it during its extended interactions with the Milky Way.

Pre-Gaia, nine globular clusters associated with Sgr were known (Majewski et al., 2003; Law & Majewski, 2010a), although several of these were only subsequently confirmed by means of the Gaia DR2 data (Bellazzini et al., 2020): M54 in the nucleus; Arp 2, Terzan 7 and 8 in its core; and NGC 2419, NGC 4147, NGC 5634, Pal 12, and Whiting 1 in the tidal streams.

Using VISTA near-infrared data to identify potential RR Lyrae members, and Gaia EDR3 proper motions to confirm membership, 12 new clusters were identified by Minniti et al. (2021b), and a further 8 by Minniti et al. (2021a). This rich system, 29 in total, quite likely completes the census of globular clusters associated with the Sgr system (Arakelyan et al., 2020; Peñarrubia & Petersen, 2021; Kundu et al., 2022).

199. Searching for the Cetus stream

THE TOPIC OF stellar streams, tidal remnants of galaxies captured by our own, has been transformed by Gaia. I have written essays on Enceladus (essay 15), on ongoing discoveries (71 and 156), and on their role in identifying black holes (176) and ΛCDM-type sub-halos (184). In my two previous essays, I focussed on the major Gaia Sausage–Enceladus and Sagittarius streams.

Of more than 120 streams now known (Mateu, 2023), I will look here at the Cetus stream, an example of those displaying multiple 'wrappings' around our Galaxy.

VARIOUS APPROACHES have been used to search for stellar streams, with more powerful algorithms being developed to exploit improvements in the quality of distances and kinematics. Roughly chronologically:

• Pole counts (Johnston et al., 1996): this identifies high-contrast structures on great circle paths, and used in the detection of the Sagittarius stream (Ibata et al., 2002a).

• Co-moving groups: this aims to identify similar star types (e.g. RR Lyrae, BHB) contained in a small phase-space volume. It was used in the discovery of the Arcturus (Arifyanto & Fuchs, 2006), Virgo (Duffau et al., 2006), and Aquarius (Williams et al., 2011) streams.

• Matched filtering (Rockosi et al., 2002, §3; Balbinot et al., 2011): this uses colour–magnitude filtering to find structures that belong to a given stellar population. It was used for Palomar 5 (Odenkirchen et al., 2001), GD–1 (Grillmair & Dionatos, 2006), Orphan (Belokurov et al., 2006), Lethe, Cocytos, and Styx (Grillmair & Carlin, 2016), Eridanus and Palomar 15 (Myeong et al., 2017), and 11 new streams in the DES data (Shipp et al., 2018).

• Given full 6D phase-space coordinates, streams can be isolated by applying cuts in energy and/or angular momentum (Li et al., 2019; Johnson et al., 2020a), or by searching for stars on similar orbits (e.g. Helmi & de Zeeuw, 2000; Yang et al., 2019), using Friends-of-Friends type algorithms (e.g. Behroozi et al., 2013).

• `STREAMFINDER` (Malhan & Ibata, 2018) was developed for Gaia, and is the most successful stream-finding approach to date. It is based on the fact that stars in a (thin and dynamically cold) stream are connected through the progenitor's orbit. Starting with one or more members of a hypothesised stream, others will then be contained in a 'hypertube' whose phase-space dimensions are set by the progenitor's orbit, size and velocity dispersion, moving in the Galaxy's gravitational potential. The trial orbits (and assumed potential) are adjusted to maximise the star counts in the hypertube.

It was used to discover GD–1 from Pan–STARRS1 (Malhan et al., 2018a). In Gaia DR2, Ibata et al. (2018) found Phlegethon, while Ibata et al. (2019b) discovered Slidr, Sylgr, Ylgr, Fimbulthul, Svöl, Fjörm, Gjöll, and Leiptr. In EDR3, Ibata et al. (2021) found 9 new streams.

• STARGO (Yuan et al., 2018; Yuan et al., 2019) works in the 4D space of orbital energy and angular momentum, and hence also requires full 6D phase-space information. It uses unsupervised learning (a 'self-organising map'), which trains a 2D neural network to learn the data set's topological structures from the 4-space of energy and momentum. It has been applied to the Cetus stream (Yuan et al., 2019), to the LMS–1 stream (Yuan et al., 2020), and to the identification of open clusters in the Gaia DR3 data (e.g. Pang et al., 2022; Qin et al., 2023).

• `VIA MACHINAE` (Shih et al., 2022; Shih et al., 2024): is another algorithm developed specifically for the Gaia data. But it departs radically from `STREAMFINDER` in that it is 'model agnostic', i.e. it makes no assumptions about the form the Galactic potential, orbits, or isochrones. It uses a data-driven, unsupervised machine-learning method for anomaly detection (`ANODE`), originally developed for the Large Hadron Collider.

Applied to the positions, proper motions, colour and magnitudes from Gaia, it first identifies stars that are 'anomalous' (overdense) with respect to the background, thereafter restricting selection to overdensities that are broadly consistent with stellar streams. Shih et al. (2024) identified 102 streams in Gaia DR2, of which only 10 had been previously identified.

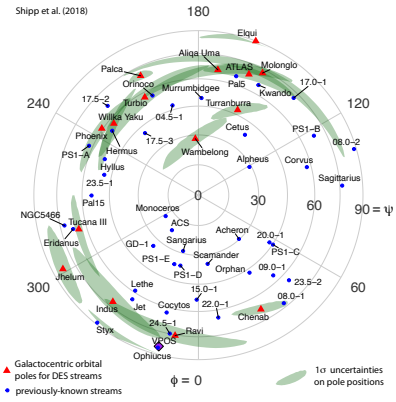
This is not an exhaustive listing, and a number of other stream-searching algorithms can be found in the literature, amongst them another recent model-agnostic algorithm, `SkyCURTAINs` (Sengupta et al., 2025).

THE CETUS STREAM, some 25–40 kpc from the Sun, was discovered from SDSS–SEGUE by Newberg et al. (2009). From the orbit inferred from the stream's radial velocity, they suggested that the globular cluster NGC 5824 is also associated with it. Yam et al. (2013) used N-body simulations to show that the stream could be reproduced by a disrupted dwarf galaxy of $10^8 M_\odot$.

Subsequent studies used the various metallicities to investigate whether NGC 5824 is the disrupted core of the progenitor (Da Costa et al., 2014; Roederer et al., 2016; Mucciarelli et al., 2018), although deep photometry found no evidence of tidal tails around the cluster itself (Walker et al., 2017; Kuzma et al., 2018).

WITH THE ARRIVAL of the Gaia data, Yuan et al. (2019) used data from Gaia DR2, LAMOST, and SDSS, and the `STARGO` stream-searching algorithm, to identify three groups in the metal-poor ([Fe/H] < -1.5) outer halo ($d > 15$ kpc), corresponding to the Sagittarius, Orphan, and Cetus streams. The 150 members of the Cetus stream extended over both sides of the Galactic plane.

While Yuan et al. (2019) confirmed the stream's association with NGC 5824, its metallicity dispersion indicated that the progenitor could not have been a globular cluster. They suggested instead that NGC 5824 was associated with a low-mass dwarf galaxy involved in the merger. Chang et al. (2020) used N-body simulations to argue that NGC 5824 was not the nuclear star cluster of a dwarf progenitor, but rather located off-centre from it.



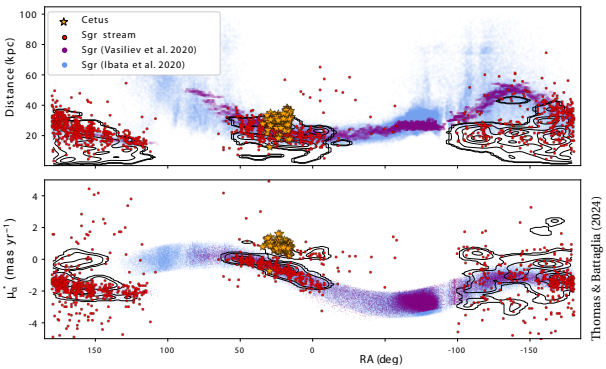
Chang et al. (2020) also predicted that about half of the stream members would be in the southern sky. And their predicted location overlapped the diffuse Palca stream, recently discovered from the Dark Energy Survey, and at a similar distance of ~36 kpc (Shipp et al., 2018). The rich structure of these Dark Energy Survey discoveries gave its name to Palca, Quechua for 'cross of rivers'.

Chang et al. (2020) also suggested that another diffuse substructure, the Eridanus–Phoenix overdensity, was also likely to be related to the Cetus stream.

TO SUMMARISE before proceeding: as of 2020, models of the Cetus stream implied that NGC 5824 is offset from the centre of the dwarf progenitor, and that the Cetus stream extends over the southern (equatorial) hemisphere, overlapping with the known Palca stream.

FURTHER CLARITY came with two papers exploiting the Gaia EDR3 data. With Gaia distances being of limited value beyond 10 kpc, Thomas & Battaglia (2022) used spectro-photometric distances (based on artificial neural networks) for 300 000 SEGUE stars, extended to 6D phase-space using Gaia proper motions, with stream members then identified in their integrals of motion.

They confirmed that the Cetus stream and the Palca overdensity are parts of the same structure, with a combined Cetus–Palca stream mass $1.5 \times 10^6 M_\odot$, and a prominent distance gradient of 15 kpc over the 100° arc on the sky. A second structure, almost parallel to the Cetus stream and extending over 50°, could be a stream resulting from the tidal disruption of a globular cluster that was orbiting *around* the Cetus stream progenitor.



SIMILAR CONCLUSIONS, but based on different data and a different search algorithm, were reached by Yuan et al. (2022). They combined the advantages of both `STREAMFINDER` and `StarGO`, to characterise the Cetus stream as a complex, very metal-poor, nearly polar structure around the Milky Way. They confirmed the southern extensions of the northern Cetus stream as the Palca stream, and identified an additional southern stream, which overlaps on the sky but at a different distance, both extending over more than 100° on the sky.

Their N-body model reproduces both as two wraps (of the progenitor around the Milky Way) in the trailing arm, and yields a progenitor mass of $\gtrsim 4 \times 10^5 M_\odot$, comparable to the Ursa Minor and Draco dwarfs.

In addition, they associated the modelled Cetus–Palca stream with the known Triangulum/Pisces stream (Bonaca et al., 2012; Martin et al., 2013), and with the Willka Yaku stream (Shipp et al., 2018), as had been suggested by Bonaca et al. (2021), and possibly with the C-20 stream discovered by Ibata et al. (2021).

They also concluded that the globular cluster NGC 5824, of similar stellar mass, was not the main progenitor, by possibly accreted in the same group infall. The multi-wrap Cetus stream, they conclude, *'is a perfect example of a dwarf galaxy that has undergone several periods of stripping, leaving behind debris at multiple locations in the halo'*.

200. Chromospheric activity

THE SUN'S chromosphere lies above its photosphere and below the corona. It is seen as a thin (~100 km) annulus above the lunar limb during solar eclipse, and its distinctive red colour is dominated by the Hα 656 nm transition. The term was suggested by Lockyer (1868), although the region's boundaries still remain only imprecisely defined (Linsky, 2017, §2). Spicules and prominences, bright features above the limb and dark features on the disk, provide evidence that the chromosphere is out of radiative equilibrium, with a temperature *higher* than at the top of the photosphere (Beckers, 1972).

There is a substantial literature on the complex physics of the solar chromosphere (Carlsson et al., 2019), and its role in coronal heating (Reale, 2014), and in driving the solar wind (Cranmer & Winebarger, 2019), but I will go no deeper into these aspects here.

THE PHENOMENA collectively known as 'solar activity' originate from the cyclic regeneration of the Sun's large-scale magnetic field (Charbonneau, 2010), and it is in this context that solar variations in Ca II and Hα are interpreted (Durney et al., 1981; Zills et al., 2024).

Similarly, strong and variable magnetic fields in cool stars with convective envelopes are driven by analogous stellar dynamos, themselves evidenced by photospheric starspots, chromospheric plages and coronal flares, and by Ca II and Hα emission (Linsky, 1980; Linsky, 2017, §4; Carlsson et al., 2019, §3.2). For solar-type and cooler stars, UV/XUV radiation from their chromospheres and transition regions also drives photochemical processes in exoplanet atmospheres (Linsky, 2017).

Angular momentum loss via the magnetised wind causes the star to spin down, so that phenomena associated with the dynamo and the surface magnetic field decay over time (e.g. Reiners & Basri, 2008). Chromospheric activity, characterised by Hα and Ca II variability, thus contributes to understanding the complexity of stellar magnetic fields, and the resulting variability as a function of stellar type, rotation and age, as well as probing pre-main sequence mass accretion and mass transfer in binary systems (e.g. Gizis et al., 2002; Pace, 2013).

THE VARIOUS SPECTRAL LINES used as chromospheric activity diagnostics probe different physical conditions (Carlsson et al., 2019, §3.2). The Hα line is the 'classical' diagnostic: the lower level of the ($n = 3 \rightarrow 2$) transition is at an excitation energy of 10.2 eV, meaning that the opacity and line width is temperature dependent, although with large thermal broadening.

The Ca II H and K lines (396.8/393.4 nm) have the ground state as lower level, and trace the dominant ionisation stage for $T \lesssim 13\,000$ K. The parameter $R'_{\rm HK}$, relating line emission to bolometric luminosity (Noyes et al., 1984, § II), facilitates comparison across spectral types.

The Ca infrared triplet (850–860 nm) has metastable lower levels, resulting in temperature-dependent opacities, and with good sensitivity to magnetic field strength.

PAST SPECTROSCOPIC SURVEYS of chromospheric activity include the Palomar/Mount Wilson surveys (Duncan et al., 1991; Gizis et al., 2002), and the HARPS survey of 4454 cool stars (Boro Saikia et al., 2018).

Numbers have been boosted by recent catalogues from the LAMOST survey in these three different lines. A study of 1.1 million solar-like stars using the Ca II HK lines concluded, for example, that the dynamo mechanism of solar-like stars is generally consistent with that of the Sun (Zhang et al., 2024). A survey of 560 000 FGK stars using the Ca triplet at 850–860 nm classified activity as a function of $T_{\rm eff}$, and found that a significant fraction of stars that show a high activity index in both Ca II HK and in the infrared triplet are binaries (Huang et al., 2024c). Finally, of more than 2.3 million G-type stars in LAMOST DR10, 220 000 show excess chromospheric activity in Hα (Su et al., 2024).

Other surveys have focussed on specific spectral types, such as the Cepheids (e.g. Hocdé et al., 2020), and BY Draconis variables (e.g. Chahal et al., 2022).

The increasing quantity and quality of such survey data is leading to a resurgence of interest in chromospheric activity, and the prospects of gaining *'a much improved understanding of chromospheric physics and its wide-ranging impact'* (de Grijs & Kamath, 2021).

WITH THE ABOVE by way of context, Gaia's contribution to the measurement of chromospheric activity can be more easily appreciated. Here, it is important to stress that the wavelength interval of Gaia's Radial Velocity Spectrometer (RVS, 845–872 nm) includes, by design, the Ca II infrared triplet lines (850–860 nm).

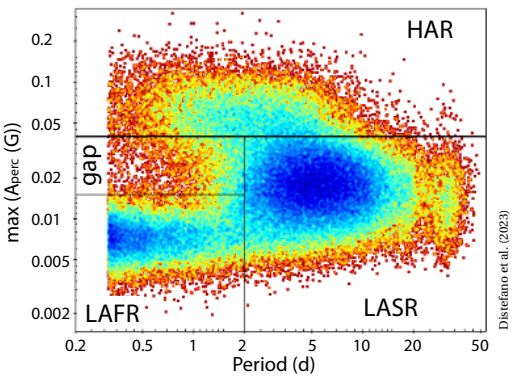
Before detailing the RVS results, some other insights into solar-like variability, including effects of flares and rotating starspots, have also been obtained from Gaia based on the wider photometric variability classification undertaken as part of Data Release 3 (Rimoldini et al., 2023), and specifically focussing on magnetically active stars, initially in the case of DR2 (Lanzafame et al., 2018), and further refined for DR3 (Distefano et al., 2023).

Distefano et al. (2023) analysed a subset of 30 million late-type stars, and derived a catalogue of 474 000 stars with variability induced by magnetic activity, of which 430 000 are newly discovered variables. For each star, their catalogue includes the stellar rotation period P, the photometric amplitude A of the rotational modulation in the Gaia G band (calculated as some maximum over various percentiles), and a correlation coefficient r_0 between magnitude and colour variation.

Their amplitude–period diagram (reproduced below) shows three broad clusterings of rotating stars that they refer to as High-Amplitude-Rotators (HAR), Low-Amplitude-Slow-Rotators (LASR), and a new class of Low-Amplitude-Fast-Rotators (LAFR), the latter not seen, for example, in the Kepler data.

Distefano et al. (2023) conclude that the high-amplitude rotators (HAR) are dominated by long-lived dark spots, while the LAFR stars are dominated by bright faculae and characterised by rapidly variable magnetic fields. The LASR stars are characterised by a moderate correlation between brightness and colour variations, implying that dark spots are still the main cause of their variability, but that faculae spatially or temporally uncorrelated with the spots tend to attenuate the correlation between magnitude and colour variations.

And they interpret the gap between the HAR and LAFR stars as evidence for a rapid transition between their different magnetic configurations.



LET ME TURN NOW to the RVS data, and the insights into chromospheric activity that have been derived from the Ca II infrared triplet lines by Lanzafame et al. (2023).

In Data Release 3, let me recall, the first stage of the RVS processing employs the 'Astrophysical Parameters Inference System' (Apsis, Creevey et al., 2023; Fouesneau et al., 2023), which I summarised in essay 89.

Amongst the 13 Apsis modules, the Extended Stellar Parametrizer for Cool Stars (ESP–CS) computes a chromospheric activity index by comparing the observed RVS spectrum with a purely photospheric (radiative equilibrium) model. For this, $T_{\rm eff}$, $\log g$, and [M/H] are derived from the modules GSP–Spec or GSP–Phot.

Lanzafame et al. (2023) then used the excess equivalent width in the triplet line cores ($\Delta\lambda = \pm 0.15$ nm) as a measure of chromospheric activity (and, in specific cases, of the mass accretion rate in pre-main sequence stars). And, in analogy with the $R'_{\rm HK}$ activity index, they derived a similarly defined $R'_{\rm IRT}$ index, largely independent of the photospheric parameters.

Their derived catalogue contains a stellar activity index, derived from the Ca II infrared triplet, for some 2×10^6 stars. They identified three regimes of chromospheric activity, largely confirming suggestions made in previous work based on much smaller $R'_{\rm HK}$ datasets.

The highest stellar activity regime is populated by pre-main-sequence stars, where the excess flux with respect to radiative equilibrium appears to be dominated by mass accretion, and close binary systems such as RS CVn systems, in which magnetic activity may be significantly enhanced by tidal interaction.

Stars with 3500–5000 K are either very active pre-main-sequence stars, or active main sequence stars with a unimodal activity distribution. A dramatic change in the activity distribution is found for $T_{\rm eff} < 3500$ K, with a dominance of low-activity stars close to the transition between partially- and fully-convective stars, and a rise in activity down into the fully-convective regime. Some evidence of a bimodal distribution in main sequence stars with $T_{\rm eff} > 5000$ K is also found. Interestingly, their $R'_{\rm IRT}$ index is well correlated with the $R'_{\rm HK}$ values from Boro Saikia et al. (2018) noted above.

These are early days in the scientific interpretation. But the Gaia data are clearly allowing activity to be characterised as a function of stellar parameters with unprecedented detail, outlining different regimes of chromospheric heating, and identifying systems for which emission resulting from mass accretion may dominate.

IN RESEARCHING this essay, I was amused by the opening sentence of the review on 'Solar Spicules' by Jacques Beckers (1968), which resonates with my own experience and interest in writing these essays, and which I thought I would share: *'The author of a review article is undoubtedly the one who benefits most from it.'*

201. Young stellar objects

YOUNG STELLAR OBJECTS (YSOs) represent the earliest stages of star formation. The term embraces the *protostar* phase, which starts with the gravitational collapse from the parent molecular cloud (lasting of order 0.5 Myr while the object is still accumulating mass), and the subsequent *pre-main-sequence* phase (PMS), which starts with the exhaustion of the infalling gas.

The PMS phase continues until contraction, and the associated temperature increase, initiates H fusion. Residual gas and dust is blown away, the object becomes visible optically, and the star settles onto the zero-age main-sequence (ZAMS). As the disk material is depleted, its infrared emission decreases, such that YSOs can be classified in evolutionary stages based on their infrared spectral index (classes I, II and III due to Lada, 1987).

Pre-main-sequence stars are classified by mass as T Tauri stars ($M \lesssim 2M_\odot$, with ages $\lesssim 10$ Myr; Joy, 1945), or Herbig Ae/Be stars ($2-8M_\odot$; Herbig, 1960; Thé et al., 1994), with more massive stars contracting too rapidly to be visible as PMS objects.

YSOs are associated with many other early evolutionary phenomena and phases (e.g. Lada, 1985): circumstellar and protoplanetary (proplyd) disks, jets and bipolar outflows, masers, Herbig–Haro objects (associated nebulosity), and dippers, as well as the episodically accreting FU Ori and EX Lup (aka EXor) variables.

So far, I have mentioned some of these phenomena only in passing: YSOs in molecular clouds (essay #192) and wide binaries (193); and T Tauri stars in the context of OB associations (18) and stellar rotation (103).

HIPPARCOS made a contribution to the field by providing direct distance estimates to a number of star-forming regions and individual YSOs, including both T Tauri (e.g. Frink et al., 1998; Hoff et al., 1998; Wichmann et al., 1998), and Herbig Ae/Be stars (e.g. van den Ancker et al., 1998; Vieira et al., 2003), and better identifying their location with respect to the main sequence, and the predictions of evolutionary models.

A catalogue of 1250 proper motions, allowing their origins to be traced back in time to sites of star formation, was given by Ducourant et al. (2005).

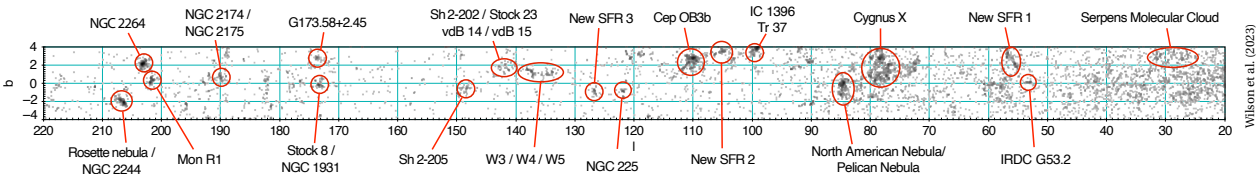
THE FIRST large-scale application of the Gaia data to the identification of YSOs came with a study using DR2 by Marton et al. (2019). They constructed a cross-matched catalogue of 103 million objects from Gaia and WISE (in the infrared), and used machine learning to assign each object to four classes: YSOs, main-sequence, evolved stars, and extragalactic objects. This yielded 1.1 million YSO candidates at 90% probability.

For verification, they showed that the 3D structure of the Orion A star-forming complex agreed with the recent literature. They also assessed the efficiency of the Gaia Science Alerts pipeline (essay 36) in detecting rapid YSO brightness changes caused by episodic accretion. As of 2019 April, the Gaia Alerts database included 7 670 objects, of which 131 were known to be YSOs. Finding 187 YSOs in their DR2–WISE sample suggests that even more of the alerts are attributable to YSO activity. I will return to this in the context of FU Ori and EXor variables below.

WITH GAIA DR3, the Konkoly Catalogue of 12 000 optically selected YSOs within 2 kpc (KYSO, Marton et al., 2023), based the *Handbook of Star Forming Regions* (Reipurth, 2008a; 2008b), was created as a training set for the DR3 variability processing. The classification of 12.4 million DR3 variables into 25 classes, including the class of YSOs, is detailed by Rimoldini et al. (2023).

Their analysis (their §4.24) yielded 79 375 YSOs of various classified sub-types, including dipper stars (DIP), eruptives such as FU Ori-type variables (FUOR), pulsating PMS stars (PULS_PMS), Herbig Ae/Be types (HAEBE), and T Tauri star (TTS), among which are classical (CTTS), weak-lined (WTTS), and late G to early K type pre-main-sequence (GTTS) stars. Some 40 000 of these had not been previously catalogued as YSOs.

Given the relatively low completeness and high contamination expected for optical observations alone, detailed validation by comparison with numerous literature catalogues was carried out by Marton et al. (2023). They found that most candidates are indeed found along lines-of-sight to known star-forming regions and the Galactic mid-plane, with an average of 10% contamination, and completeness at around the percent level.



AN ADS search returns some 130 refereed papers on Gaia–YSO studies to mid-2024, and the following gives a flavour of the many ongoing investigations.

The largest catalogue of mid-infrared (3–9 μm) YSOs in the Galactic plane is from the Spitzer–IRAC wide-area mapping. In their study of 120 000 of these YSO candidates, Kuhn et al. (2021) used DR2 distances to associate these YSOs with the Local Arm, the Sagittarius–Carina Arm, and the Scutum–Centaurus Arm.

In an extensive classification exercise using a 'naive Bayes classifier', and data from WISE, UKIDSS, 2MASS and IGAPS, along with photometric variability from Gaia EDR3, Wilson et al. (2023) identified 6504 candidate Class II YSOs, with good sensitivity to different evolutionary stages (see figure above).

Between the end of the AGB phase and the onset of the planetary nebula phase, the AGB dust shell becomes optically thin, resulting in infrared colours similar to those of YSOs because of their similar cold, dense circumstellar dust envelopes. With YSOs and AGB stars commonly identified from their infrared 2-colour diagrams (e.g., Koenig & Leisawitz, 2014; Suh, 2021), misclassification can arise in the significant overlap regions (e.g. Lee et al., 2021b). Suh (2024) showed that distance and extinction data from Gaia DR3 is highly effective in distinguishing these very different stellar classes.

AMONGST STUDIES of YSOs in specific star-forming regions, Grasser et al. (2021) used the 3D data from Gaia EDR3 of the ρ Oph region to evaluate the existing catalogue of 1114 YSO members from the literature. They found 191 new YSOs (mainly Class III M stars). The proper motions reveal two distinct populations in a similar 3D volume (and 17 non-members): one around ρ Ophiuchi and the main Ophiuchus clouds, and the other ~10 Myr older, more dispersed, and possibly from the Upper Sco group.

Using DR3, Narang et al. (2023) searched for optical counterparts for 62 protostars in Orion, and Gómez de Castro et al. (2024) characterised 63 candidate T Tauri stars in the Taurus–Auriga molecular complex.

Gaia parallaxes and proper motions are also being used to study young stars in the regions of Herbig Haro objects and flows, for example in Canis Major (Pettersson & Reipurth, 2019); around the BBWo 192E nebula (Magakian et al., 2020); in the Dobashi 5006 Dark Cloud (Movsessian et al., 2023); and in the Taurus B18 cloud (Duan et al., 2023).

THE FU Ori variables are PMS stars displaying abrupt changes in magnitude and spectral type. The prototype, FU Ori, increased from 16.5 to 9.6 mag in 1937. The second, V1057 Cyg, brightened by 6 mag and transitioned from dKe to F-type in 1969–70 (Kopatskaya et al., 2013). The flaring appears to be repetitive, cycling between the FUor state and the T Tauri state (Reipurth, 1990). Amongst various models (Audard et al., 2014; Vorobyov et al., 2021) are episodic mass transfer from an accretion disk (Vorobyov & Basu, 2006; Vorobyov & Basu, 2015), or possibly due to a young, inwardly migrating gas giant planet (Lodato & Clarke, 2004; Nayakshin & Lodato, 2012; Nayakshin & Elbakyan, 2024).

Even 50 years ago, the mean time between FU Ori-type outbursts in individual T Tauri stars had been estimated at 10 000 years (Herbig, 1977). Contreras Peña et al. (2019) have further quantified this rate by comparing old photographic surveys with Gaia DR2, giving a temporal baseline of 55 yr for 15 000 Class II objects. They found 139 with $\Delta R \geq 1$ mag (mostly 1–3 mag), most showing irregular variability or long fading events, best explained as hotspots due to accretion or variable extinction. A tail at $\Delta R \geq 3$ mag showed high-amplitude irregular variability over time-scales shorter than 10 yr.

Six objects were consistent with large, long-lasting accretion events, 3 previously unknown. This yielded a recurrence time-scale of 112^{+68}_{-38} kyr, and confirms that YSOs in their planet-forming stage undergo such large accretion events. Outbursts in the Class II stage are a factor ten less frequent than Class I.

Pre-Gaia, around 10 FU Ori stars were known. Gaia is discovering others with ongoing flares from its Science Alerts pipeline, including Gaia 17bpi (Hillenbrand et al., 2018), Gaia 18dvy (Szegedi-Elek et al., 2020), and Gaia 19ajj (Hillenbrand et al., 2019).

THE RELATED class of EXors, named after the prototype EX Lupi, are indistinguishable from T Tauri stars apart from repeated luminosity increases, a year or so in duration, attributed to episodic accretion (Herbig, 2008; Sipos et al., 2009; Lorenzetti et al., 2012; Moody & Stahler, 2017; Magakian et al., 2022).

Again, others are being discovered through the Gaia Science Alerts pipeline, and characterised by Gaia astrometry. They include Gaia19fct (Park et al., 2022; Ghosh et al., 2022a), Gaia20eae (Cruz-Sáenz de Miera et al., 2022), Gaia21elv (Nagy et al., 2023), Gaia21bty (Siwak et al., 2023), and Gaia23bab (Giannini et al., 2024).

202. Update on science alerts

As Gaia scans the sky, it detects and observes all objects brighter than $G \sim 20.7$ mag. Objects bright enough at that specific field-of-view crossing time, including regular or irregular variables, or moving objects within the solar system, are detected and observed.

Using Gaia's own data base of variable stars and solar system objects (such that pulsators, regular variables, and eclipsing binaries are largely excluded as alerts), sudden brightenings or dimmings, or entirely 'new objects', can be flagged. Alerts are then promptly issued to observers worldwide for follow-up photometric observations and spectroscopic classification.

Since the start of the mission, a group at the Institute of Astronomy, Cambridge, has been leading the processing of the satellite photometry, and a sub-group, led by Simon Hodgkin, handles these 'science alerts'.

I gave an introduction to this Gaia Science Alerts system in essay 36 (Sep 2021). Then, by the end of 2020, 15 000 alerts had been issued, 50 with $G < 12$, and 5000 had been classified: 500 cataclysmic variables, 100 AGN, 600 variable quasars, 2000 supernovae (1500 Type Ia, 500 Type II), and 50 microlensing events.

Subsequently, Hodgkin et al. (2021) have given a detailed description of the Gaia Science Alerts system for Gaia EDR3. This includes changes made as more data have become available, and more reliable and efficient detection algorithms have been developed.

The alerts pipeline has to handle many complications. There are modules optimised for detecting *new* sources, for bursts in previously known sources, and for light-curve skewness. 'Bad' transits, including prompt particle events and bright star artefacts, are largely suppressed by requiring detection in two successive fields of view. Solar system objects are identified by forward orbit prediction. A final assessment is done by eye.

Following the convention used for supernova discoveries (the prefix SN, followed by the discovery year, suffixed with a one or two-letter designation), successful 'transients' are designated GaiaYYaaa, GaiaYYaab…

Alerts are published via the IAU Transient Name Server, and as VOEvents using 4 Pi Sky (Staley & Fender, 2016). They also populate their Gaia Science Alert www pages with new alerts, along with the data collected for each source, including detected and historic G-band magnitudes, the light curves, and the low-resolution Gaia (B_P/R_P) spectra (Delgado et al., 2019).

Their Figure 9 (alerts as a function of magnitude between 2014–19) shows that most alerts span the range $G = 13 - 21$, and peak at $G \simeq 18.5$ mag. Their Figure 10 shows the alert rate as a function of time, which has increased from ~1 per day in 2015, to ~12 per day in 2020.

There are, of course, several other transient surveys ongoing, including ASAS–SN (All-Sky Automated Survey for Supernovae; Kochanek et al., 2017), ATLAS (Asteroid Terrestrial-impact Last Alert System; Smith et al., 2020), Pan-STARRS (Chambers et al., 2016), and ZTF (Zwicky Transient Facility; Bellm et al., 2019). As one indicator of fidelity, these multiple surveys often results in the same event being identified by different surveys/names.

Of these, only Gaia and ASAS–SN cover the entire sky, including the Galactic plane. Gaia is also the only *catalogue-driven* transient survey – the others employ difference-imaging techniques. And Gaia is, today, the second-largest contributor to the IAU Transient Name Server (Hodgkin et al., 2021).

Various telescopes around the world are involved in follow-up photometric observations (e.g. Dambljanović et al., 2018; Kvernadze et al., 2023), crucial for characterising the detailed light-curves, and for spectroscopy, crucial for object classification.

Various other studies have fed into the alerts system (e.g. Wevers et al., 2018; Zieliński et al., 2019), including those focussing on microlensing events (Gezer et al., 2022; Wyrzykowski et al., 2024), cataclysmic variables (Breedt, 2019; Mistry et al., 2022), transients in galactic nuclei, including 'tidal disruption events' (Kostrzewa-Rutkowska et al., 2018), gravitational wave event detection (Kostrzewa-Rutkowska et al., 2020), and the contribution of amateur astronomers (Romanov, 2022).

THERE WERE 3484 alerts in 14 classes in Hodgkin et al. (2021). On 2024 Aug 23 the database listed 25 919 alerts, with 7 141 assigned to 23 classes (assigning all SN to one class): AGN (187), BL Lac (154), CCSN (4), CV (779), dK (1) and dM (18), galaxy (4), ILRT (1), LBV (5), nova (46), QSO (1244), RCrB (16), SLSN (59), SN (4139), SN Impostor (1), SSO (2), star (32), symbiotic star (8), TDE (25), ULENS (42), VarStar (72), XRB (13), YSO (283).

[CCSN=core-collapse SN, ILRT=intermediate-luminosity red transients, LBV=luminous blue variable, SLSN=super-luminous SN, SSO=solar system object, TDE=tidal disruption event, ULENS=microlens, YSO=young stellar object]. The SN are further divided as SN (25), SN I (16), SN Ia (2617), SN Ia-CSM (9), SN Ia-pec (21), SN Iax (8), SN Ib (73), SN Ib/c (26), SN Ibm (18), SN Ic (81), SN Ic-BL (30), SN Ic-pec (1), SN II (875), SN IIb (54), SN IIl (3), SN IIn (166), SN IIp (116).

I LIST here a selection of these, with links to the relevant Gaia Science Alert pages. Those marked* (lensing, AM CVn, and TDEs) are detailed in essays 203–206.

Supernovae: among many are Gaia16apd, a UV-bright super-luminous SN (Kangas et al., 2016; Nicholl et al., 2017); Gaia16bvd, the first of a 'pair-instability' supernova (Gomez et al., 2019); Gaia16cfr, an interacting transient (Brennan et al., 2022); Gaia17biu, a nearby super-luminous SN (Xiang et al., 2017; Bose et al., 2018); Gaia17dcj, a SN interacting with dense circumstellar material (Moran et al., 2023); Gaia22cbu, with periodic light-curve modulation (Moore et al., 2023).

Supernova Impostor: Gaia16ada is an example of powerful novae that resemble supernovae but which don't destroy their progenitor (Aghakhanloo et al., 2023).

Changing-Look Quasars: a class of AGN with strongly changing features, possibly attributable to tidal disruption events or episodic accretion (LaMassa et al., 2015; MacLeod et al., 2016, Yang et al., 2018). Gaia19bwn and Gaia19dsk are possible examples (Pursimo et al., 2019).

Microlensing: the 40 events to date include Gaia16aye, the first binary event in the Galactic disk, rather than bulge (Wyrzykowski et al., 2020); Gaia18ajz, a $4.9 M_\odot$ black hole candidate (Howil et al., 2025); Gaia18cbf, a long duration event (Kruszyńska et al., 2022); Gaia19bld*, with gravitationally lensed arcs in rotation from VLTI–PIONIER (Cassan et al., 2022); Gaia19dke*, a very long duration event, with multiple peaks in the light curve (Maskoliūnas et al., 2024); Gaia20bof, a close binary lens, with dense photometry from the OMEGA Key Project (Bachelet et al., 2024).

Gaia22dkv*, is the first Gaia microlensing planet, also towards the disk, and only the third Gaia host listed in the NASA Exoplanet Archive (Wu et al., 2024).

AM CVn: Gaia14aae*, the first fully-eclipsing AM CVn binary (Campbell et al., 2015; Green et al., 2018b: Green et al., 2019b; Sarkar et al., 2023).

Novae: Gaia18blv, a luminous red nova with a supergiant progenitor (Blagorodnova et al., 2021); Gaia22alz, a very slow nova with a 180-d rise (Aydi et al., 2023).

Symbiotic stars: Gaia22eor (V2756 Sgr), a symbiotic detected in a deep eclipse (Merc et al., 2022); Gaia23clr (V2905 Sgr), a symbiotic outburst (Merc et al., 2023; Merc et al., 2024); Gaia23ckh (V390 Sco), a symbiotic outburst of the Mira variable (Merc et al., 2024).

Young Stellar Objects (YSO): I addressed these in my previous essay 201, along with the pre-main-sequence classes displaying abrupt changes in magnitude and spectral type (the FU Ori and EX Lup, aka EXors).

Pre-Gaia, ~10 FU Ori stars were known. Gaia discoveries include Gaia17bpi (Hillenbrand et al., 2018); Gaia18dvy (Szegedi-Elek et al., 2020); Gaia18dvz (Hodapp et al., 2019); Gaia19ajj, (Hillenbrand et al., 2019).

The related class of EXors include Gaia17aeq, (Sicilia-Aguilar et al., 2017; Cieza et al., 2018; Kashi et al., 2019); Gaia19fct (Park et al., 2022; Ghosh et al., 2022a), Gaia20eae (Cruz-Sáenz de Miera et al., 2022); Gaia21elv (Nagy et al., 2023); Gaia21bty (Siwak et al., 2023); Gaia23bab (Giannini et al., 2024); and the possibles Gaia20bwa and Gaia20fgx (Nagy et al., 2022).

Dippers: the science alerts pipeline also flags sources that *fade* significantly. A number of new YSOs (Young Stellar Objects), and other 'dipping' sources, such as VY Scl stars, have been discovered by Gaia, including Gaia17afn (the young star V555 Ori), showing a dipper-like variability (Nagy et al., 2021).

Centres of galaxies: A growing number of transients are found in galaxy centres, where outbursts may be due to a change in the black hole accretion flow, or to a tidal disruption event. Gaia16aax is in a QSO-hosting galaxy, which brightened by about 1 mag over 1 year, before fading to its pre-outburst state (Cannizzaro et al., 2020).

Tidal disruption events (TDE): These results from a star passing close to a supermassive black hole (Hills, 1975; Rees, 1988; Kochanek, 2016). Pulled apart by the black hole's tidal force, the star is disrupted, producing a tidal stream of material that loops around the black hole.

As of May 2024, ~100 TDEs are known, most discovered by the optical transient surveys ZTF, ASAS–SN, and Gaia (Nicholl et al., 2020; Zhang, 2024). Amongst the 25 Gaia discoveries are Gaia20fck* / ATLAS20belb (Charalampopoulos et al., 2023); and Gaia20cjk* /AT2020ksf at $z = 0.092$ (Wevers et al., 2024).

Gravitational waves: Discovery of the electro-magnetic counterpart to the gravitational wave event GW170817 (Abbott et al., 2017) has motivated studies to extend Gaia's alert limit to $G > 19$, while allowing a single transit to trigger an alert (Kostrzewa-Rutkowska et al., 2020). Tests and preparations for the 4th observing run of the LIGO–Virgo–KAGRA collaboration that started on 2023 May 24, GaiaX, are described by Biswas et al. (2023).

203. Gaia's third exoplanet, Gaia22dkvLb

I HAVE ESTIMATED that Gaia will eventually discover tens of thousands of exoplanets based on the astrometric motion of the system's photocentre (Perryman et al., 2014). Admittedly, these estimates are sensitive to the quality of the final astrometric calibrations, and the presence of other massive planets within each system.

Pending the enlarged data sets, improved calibration, and binary system (including sub-stellar mass) processing in the next data releases DR4 (2025) and DR5 (around 2030), only two Gaia exoplanets (from Gaia's transit photometry) so far appear in NASA's Exoplanet Archive: Gaia–1 b and Gaia–2 b (Panahi et al., 2022).

Of nearly 6000 exoplanets discovered globally to date, some 200 have been detected by gravitational microlensing. A new such discovery was added to the NASA archive on 2024–08-07: Gaia22dkvLb. This was discovered as a Gaia 'science alert', and it has a number of interesting features that I will expand on here.

I HAVE COVERED various aspects of gravitational microlensing relevant to Gaia in some earlier essays: focussing on *astrometric* microlensing (11); on microlensing with Gaia more generally (84); candidate black holes from astrometric microlensing (101); and the discovery of new microlensing systems from follow-up of Gaia's 'science alerts' processing system (36 and 202).

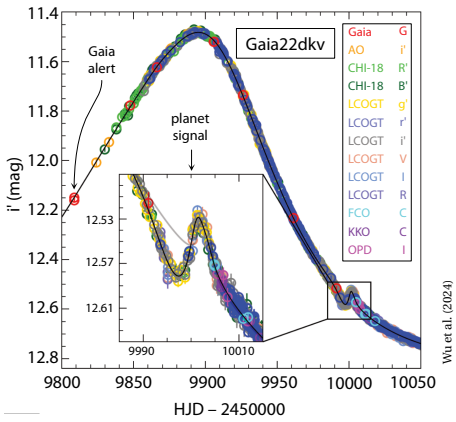
Microlensing relies on the chance (and very rare) alignment of a background source, an intervening lens, and the observer. If the foreground lens is itself of complex structure (whether a cluster of galaxies, or a binary system, or a star orbited by one or more planets), then the background source may show a more complex lensed light curve resulting from the time-varying magnification as the alignment changes.

This has led to discovery of more than 200 microlensing exoplanets. All but 3 have been by the ground-based surveys MOA (Sako et al., 2008), OGLE (Udalski, 2009) and, most recently KMT (Kim et al., 2016). It is the rarity of alignment that has motivated these surveys to focus on the Galactic bulge, where alignment probabilities benefit from much higher stellar surface densities.

GAIA'S latest planet was discovered through the `Gaia Science Alerts`, part of the Gaia Data Processing and Analysis Consortium (DPAC), which processes the real-time photometric data from the scanning satellite, and flags unexpected magnitude increases, allowing rapid follow-up by ground-based photometry (Hodgkin et al., 2021, and essay 202). For microlensing candidates, the goal is to initiate a rapid and dense sampling of the rising light curve (while the lensing event is still ongoing), with the aim of detecting light-curve perturbations signalling the existence of an orbiting planet.

This specific `event announcement` (all include the Gaia light curve and low-resolution B_P/R_P spectra at each epoch) was released on 2022–08–16 at 22:40:56. It drew attention to a bright, Galactic plane source with historic magnitude $G = 13.18 \pm 0.02$, as it brightened by 0.6 mag to an 'alert magnitude' of $G = 12.53$ mag.

THE SUBSEQUENT observations and analysis are reported by Wu et al. (2024). The figure below shows their follow-up photometry, from several telescopes [including Auckland (AO), El Sauce Chile (CHI–18), and Las Cumbres (LCOGT)], and in various filters, here transformed to the photometric band i'. The dense sampling over the subsequent 250 days, the photometric accuracy, and the perturbation on the falling light curve around days 9990–10010 (HJD–2,450,000), are all very striking.



THE OBJECT DESIGNATION, Gaia22dkvLb, denotes the Gaia science alert identifier, Gaia22dkv (itself following the convention GaiaYYaaa, with YY denoting the event year, and a 3-letter encoding of the event sequence in that year), with the suffices Lc adhering to the microlensing exoplanet nomenclature in denoting the planet b associated with the lensing object L.

I SHOULD EMPHASISE that the analysis of most microlensing events is not straightforward. For example, for a single lens event, the Einstein time scale, derived from the event duration, is a degenerate combination of the lens mass, $M_{\rm L}$, and lens distance, $D_{\rm L}$ (or, for a source at finite distance, the 'microlens parallax' defined as the (scaled) relative parallax between lens and source, denoted $\varpi_{\rm E}$ or $\pi_{\rm E}$), and the lens–source relative transverse velocity. Indeed, the clear asymmetry in the main light curve is due to the Earth's orbital motion.

Measurements of $\varpi_{\rm E}$ can be made by exploiting this non-linear (orbital) motion of the Earth around the Sun (an 'orbital microlens parallax'), or in favourable cases by making simultaneous observations of the microlensing light curve over an extended measurement baseline on Earth (a 'terrestrial microlens parallax'), or by observing the same event from Earth and a satellite (a 'satellite–microlens parallax'). The latter has been achieved with both Earth–Spitzer observations, and even Spitzer–Kepler K2 observations (see, e.g., Zang et al., 2020). The Earth–Gaia space parallax was included in the modelling of Gaia16aye by Wyrzykowski et al. (2020).

There are other complications. If the lens is unseen, its mass and distance can be determined only in particularly favourable circumstances. If the lens is visible, solutions depend on the degree of light blending between lens and source, and the possible contribution of another (non-lensing) object along the line of sight. Finite source size effects can also be evident.

ACCOUNTING FOR THESE and other complexities, Wu et al. (2024) concluded that the lens is a $M_{\rm L} = 1.15^{+0.16}_{-0.08} M_\odot$ star at a distance $D_{\rm L} = 1.27^{+0.43}_{-0.25}$ kpc, orbited by a planet of mass $M_{\rm p} = 0.59^{+0.15}_{-0.05} M_{\rm J}$ with a projected orbital separation $a = 1.41^{+0.76}_{-0.36}$ au, and an orbital period $P = 2.96 \pm 0.20$ yr.

The host star and planet masses are typical of exoplanet systems, including microlensing discoveries, but there are other properties that make it more unusual.

AS I NOTED ABOVE, ground-based microlensing surveys have focussed on the direction of the Galactic bulge to provide enhanced lens and source densities, and so improve alignment (and therefore event occurrence) probabilities. Accordingly, all but one of the previous microlensing planets lie in the direction of the Galactic bulge.

The only other non-bulge microlens planet, TCP J0507+244Lb, was found as an alert from ASAS–SN (Nucita et al., 2018; Fukui et al., 2019). It is of order Neptune mass, $M_{\rm p} = 19 \pm 3 M_\oplus$, orbiting a $0.495 \pm 0.063 M_\odot$ star, at a distance of 429 ± 21 pc (Zang et al., 2020). Interestingly, both Gaia22dkv and TCP J0507+244 reached $V < 12$ mag at peak. All-sky rates for such bright events are very low, with only ~0.1 events yr^{-1} (Han, 2008).

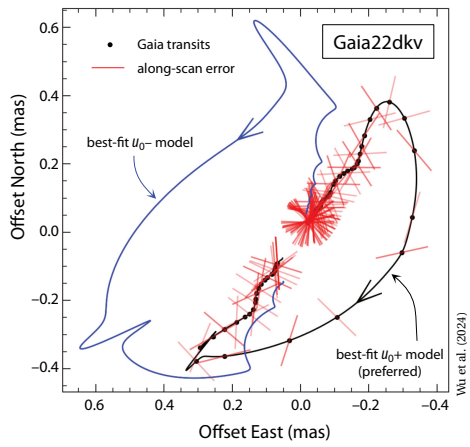
SEVERAL PROPERTIES make Gaia22dkv noteworthy. The first is that, like TCP J0507+244, it lies in a non-bulge direction, although still in the Galactic plane.

Second, the planet's semi-major axis places it within the snow-line, where water may be in liquid form. In common with TCP J0507+244, but nonetheless a rarity amongst microlensed planets, it therefore contributes to an improved understanding of the mass–radius distribution of these systems.

Third is the exceptional quiescent brightness of the lens star, $V \sim 14$ mag. If this represents (as they infer) the magnitude of the star (and not some blended image), it is the brightest microlensing host star to date.

The combination of such a bright host, high planet mass, and relatively small orbit separation, makes it the most promising microlensing planet for (the first!) radial velocity follow-up. Such observations would definitively establish the planet's orbital period and eccentricity, and would allow the search for other co-orbiting planets.

FINALLY, and pending release of Gaia's epoch astrometry, Wu et al. (2024) simulated the host's astrometric motion, based on their system parameters and the known epochs of the Gaia measurements (shown below). They conclude that the *astrometric* effects of the microlens event should be measurable once the Gaia epoch astrometry is made available (black curve). This will constrain the microlens parallax, $\varpi_{\rm E}$, break the current $(u_0 + /-)$ model degeneracy (which I have not detailed), and directly yield the lens mass and distance.



204. The unique AM CVn, Gaia14aae

THE AM CVn-type binary Gaia14aae was discovered as Gaia's fifth 'science alert', soon after the alert system went live in 2014 (Hodgkin et al., 2021; and essay 202).

This specific event (which, as for all Gaia science alerts, includes the Gaia light curve as well as the low-resolution multi-epoch B_P/R_P spectra) was announced on 2014-08-11, as it brightened, at $G = 16.04$ mag (compared with its historic magnitude, $G = 17.56 \pm 0.20$ mag).

Follow-up spectroscopy at WHT (Rixon et al., 2014), identified it as a H-deficient AM CVn-type binary, with double-peaked emission lines, consistent with a compact accreting double-degenerate binary system.

THE AM CVn systems are a rare class of very compact interacting binaries ($P_{orb} \simeq 10-65$ min), comprising a white dwarf accreting He-rich material from a low-mass companion (Nelemans, 2005; Solheim, 2010). They are distinguished from other cataclysmic variables (binaries with significant, irregular brightness increases) by the absence of H in their spectra. Only ~60 have been identified since discovery of the prototype, AM CVn (HZ 29), more than 50 years ago (Smak, 1967).

Ordered by orbital period (e.g. Kotko et al., 2012), AM CVn systems with $P < 12$ min have no accretion disk and show direct impact of the accreting material onto the white dwarf; those with $P = 12-20$ min form a large stable accretion disk, permanently in outburst; those with $P = 20-40$ min have occasional outbursts; and those with $P > 40$ min form small stable accretion disks. The ultra-short periods indicate that both donor and accretor must be degenerate or semi-degenerate objects.

AM CVn systems contribute to the understanding of what is called 'common-envelope evolution' (CEE). In this short-lived evolutionary phase for a wide range of binary stars, the two components orbit inside a single, shared envelope. CEE occupies a possible end point for binary white dwarf evolution, and affects the progenitors of Type Ia supernovae, X-ray binaries and double neutron stars. The systems are also potentially strong sources of gravitational wave emission due to their very compact orbits (Paczyński, 1967 Nelemans, 2003).

IN THEIR EXTENSIVE REVIEW, Ivanova et al. (2013) described common-envelope evolution as *'one of the most important unsolved problems in stellar evolution, and arguably the most significant and least-well-constrained major process in binary evolution.'*

Common-envelope evolution plays a key role in the evolution of many binary systems, and with various outcomes. For example (see Ivanova et al., 2013, Fig. 1, for an informative schematic), SN Ia progenitors can result from double-degenerate mergers, or accretion on a CO white dwarf from a non-degenerate companion (Maoz et al., 2014). And although the fractions of predicted CEE products can be matched to observations using 'binary population synthesis' (e.g., Nelemans et al., 2001; Goliasch & Nelson, 2015), the physics is complex (Ivanova et al., 2013, §2), and particularly challenging both analytically and computationally (e.g., Hatfull et al., 2021).

THREE CHANNELS have been proposed for the formation of AM CVn systems, with one diagnostic being the nature of the donor at the onset of mass transfer. Different channels may dominate for different periods.

In the white dwarf (Paczyński, 1967; Faulkner et al., 1972), and He-star (Savonije et al., 1986; Iben & Tutukov, 1987) channels, the binary passes through two common-envelope stages as each component leaves the main sequence. The surrounding envelope extracts energy from the system, reducing the orbital period to below the minimum for H-dominated, non-magnetic CVs. These two channels differ in the nature of the secondary star following the ejection of the second common envelope; in the former it is left as a low-mass, degenerate or semi-degenerate white dwarf, in the latter it is a non-degenerate He-burning star.

In the evolved cataclysmic variable (or evolved main sequence donor) channel (Tutukov et al., 1985; Podsiadlowski et al., 2003), the donor evolves off the main sequence at around the start of mass transfer. The system then appears as a H-dominated cataclysmic variable in its early evolution, becoming He-dominated as the donor's H envelope is stripped.

OBSERVATIONALLY, eclipsing AM CVn systems, and in particular those where the white dwarf is totally eclipsed, offer excellent prospects for measuring parameters such as component masses and orbit inclination.

However, their extreme mass ratios means that such eclipses are rarely observed. Indeed, only two eclipsing AM CVn systems were known: SDSS J0926+3624 (Anderson et al., 2005; Copperwheat et al., 2011; Szypryt et al., 2014), in which the white dwarf is only partially eclipsed, and PTF1 J191905.19+481506.2 (Levitan et al., 2014), in which only the edge of the accretion disk is eclipsed.

The importance of Gaia14aae, and its role in understanding common-envelope evolution, is as only the third known eclipsing AM CVn system. Its uniqueness is as the first in which the white dwarf is *totally* eclipsed.

GAIA14AAE was seen independently in the alerts reports of ASAS–SN during a separate outburst. A third burst is seen in archival data from Pan–STARRS-1 and ASAS–SN. The three outbursts occurred within a four-month period, while no others are seen in the previous 8 yr of ASAS–SN, Pan–STARRS-1, or CRTS (Catalina Real-time Transient Survey) data (Campbell et al., 2015).

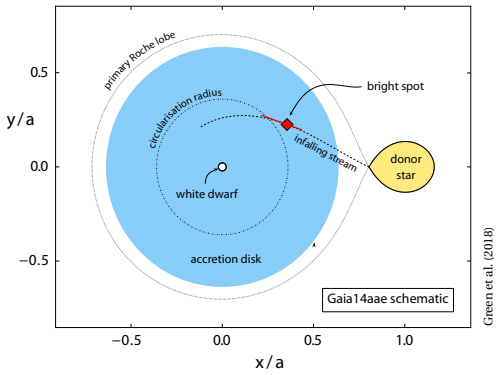
A detailed analysis of the system was reported by Campbell et al. (2015). Their follow-up photometry gave an orbital period of 49.71 min, placing it at the long-period extreme of the AM CVn period distribution. Assuming an orbit inclination $i = 90°$, the contact phases of the white dwarf gave only a lower limit on the mass ratio of $q = M_2/M_1 > 0.019$, and lower mass limits of 0.78 and $0.015 M_\odot$ for the accretor and donor respectively.

They suggested that these masses, and the estimated accretion rate, point to a degenerate donor, and that the system may have resulted from a merging double white dwarf, with a much shorter orbital period in the past as a result of mass transfer (Tsugawa & Osaki, 1997).

They also suggested that the three outbursts within 3–4 months are 'rebrightenings' (aka 'echo outbursts') rather than independent events. And from the observed correlation between period and outburst recurrence time (Levitan et al., 2015), they predict that independent outbursts in Gaia14aae recur only every 10 yr.

FROM high-speed photometry of 53 eclipses over 25 months (see schematic), Green et al. (2018b) derived the mass ratio $q = 0.0287 \pm 0.0020$ (compared with the previous upper limit), with $M_1 = 0.87 \pm 0.02 M_\odot$ and $M_2 = 0.0250 \pm 0.0013 M_\odot$, and the most precise measurement of the donor mass of an AM CVn system to date.

But they were left with a puzzle: their measured donor mass and radius do not fit with models for donors descended from white dwarfs or He stars, being more consistent with systems originating from evolved H-dominated cataclysmic variables. But such systems should show spectroscopic hydrogen… which is not seen in Gaia14aae.



PHASE-RESOLVED spectroscopy by Green et al. (2019b) is consistent with the 'central spike' tracing the motion of the central white dwarf (based on its velocity and phase), but raised other questions about the origin of a second bright spot seen in He I emission. They continued to (marginally) favour Gaia14aae as being an unusual example of an AM CVn system that has formed through the evolved cataclysmic variable channel.

THE LAST PUBLISHED contribution to an understanding of this enigmatic system at the present time is a theoretical study by Sarkar et al. (2023). They considered both Gaia14aae, and another similarly peculiar system ZTF 1637J+49. The latter was discovered in a search for deep eclipses in the ZTF (Zwicky Transient Facility) light curves of white dwarfs selected using Gaia EDR3 parallaxes (van Roestel et al., 2022).

They emphasised that previous work has assumed that the only mechanism which drives the angular momentum loss of AM CVn systems is gravitational wave radiation: for the white dwarf channel (Deloye et al., 2007); for the He-star channel (Yungelson, 2008); and for the evolved CV channel (Podsiadlowski et al., 2003).

As a first step, they showed that white dwarfs with semi-degenerate He-rich donors of mass $0.1 - 0.3 M_\odot$, can explain the observed abundances of H and C in both systems. They then showed that these systems *can* emerge from a common envelope phase if magnetic breaking, as well as gravitational radiation, is included as an energy source in ejecting the common envelope.

DUE TO THE total eclipse of the central white dwarf, the Gaia alert discovery Gaia14aae is the best characterised of all AM CVn-type systems known today.

Campbell et al. (2015) suggest that some 1000 new cataclysmic variables, including a number of evolved systems and AM CVn-type systems, will be found by Gaia over its lifetime. Gaia will also continue to play an important role in their characterisation, in outburst as well as during eclipse. This will further elucidate their formation, and provide further insights into the complex processes of common envelope evolution.

205. Two unusual microlens alerts

MY LAST three essays have been on the Gaia photometric alerts system (Hodgkin et al., 2021). In essay 202, I described the system and some of its discoveries. In essay 203 I looked at Gaia's third exoplanet, the microlens system Gaia22dkvLb. Essay 204 was on the eclipsing AM CVn-type cataclysmic variable Gaia14aae.

Here, I will look at two other interesting microlens systems discovered by means of the science alerts pipeline: Gaia19dke, a long-duration event with multiple peaks, and Gaia19bld, in which arc-like sub-images have been spatially resolved… and observed rotating.

TENS OF THOUSANDS of microlensing events have been discovered in the past 30 years, most by the dedicated search teams OGLE, MOA and KMT, with a growing number found by large-scale photometric transient surveys such as ASAS–SN, Pan–STARRS and ZTF.

I gave an introduction to Gaia's microlensing events in essay 84. Of course, Gaia provides only relatively sparse and irregular sampling: the average number of observations per source over the 34-month interval of DR3 (2014–2017) is about 40, with only 20 or so in the bulge region, although up to some 100 at intermediate ecliptic latitudes. But it has all-sky coverage, high photometric accuracy, and spans a wide range of magnitudes. Accompanying Data Release 3, Wyrzykowski et al. (2023) reported 363 Gaia discoveries, 90% of them unique.

Many complexities enter the characterisation and modelling of microlensing events (e.g. Rektsini & Batista, 2024). The main parameters of relevance for this essay are the angular 'Einstein radius' of the event, $\theta_E \propto M_L^{1/2}$, and the associated linear Einstein radius, $R_E = \theta_E D_L$, where M_L and D_L are the lens mass and distance.

The total magnification of an event varies with time due to the relative transverse motion between source, lens, and observer. For a relative source–lens transverse velocity, $v_\perp$, a typical event time scale is then given by the Einstein radius crossing time, $t_E = R_E / v_\perp$.

For a bulge source (at 8 kpc), and a $1\,M_\odot$ lens half way to the source, this 'Einstein time' is around 35 d. Longer duration events are either more massive, and/or with a small relative velocity between source and lens.

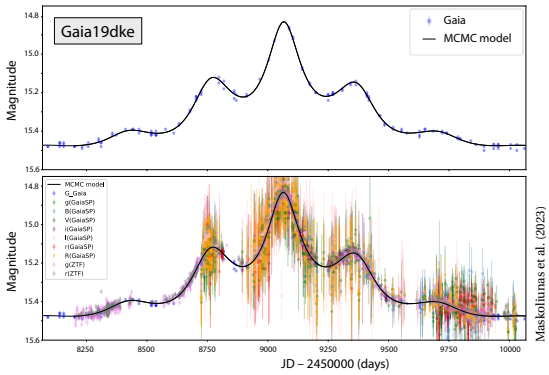
THE LONGEST recorded event, OGLE-1999-BUL-32 (MACHO-99-BLG-22), had t_E = 640 d (Mao et al., 2002). Long-duration events found in the Gaia alerts include `Gaia18cbf`, t_E = 491 d (Kruszyńska et al., 2022), and `Gaia19dke`, t_E = 159 d (Maskoliūnas et al., 2024).

The top panel below shows the G-band photometry for Gaia19dke over five years (other photometry, including ZTF, is shown below). Modelling by Maskoliūnas et al. (2024) shows that the event is due only to a single unresolved lens, with mass $M_L = 0.50\,M_\odot$ at distance D_L = 3.1 kpc (source distance D_S = 4.9 kpc). The prominent *multiple* peaks might therefore seem surprising.

Recall that long events arise from large M_L and/or small $v_\perp$. In the latter case, good source–lens alignment can extend over many months. As first described by Smith et al. (2002), multiple peaks can then result from small variations in this general alignment as the Earth (and Gaia) move in their annual orbit around the Sun.

Incidentally, the time scale of lensing events tends to increase with Galactic longitude (Mróz et al., 2019b), due to the fact that, away from the bulge, both lens and source are generally located in the disk, and are often therefore moving with similar transverse velocities.

Maskoliūnas et al. (2024) used the presence of this annual parallax, due to the Earth's motion, to infer the lens mass. This, combined with their blending analysis, suggests that the dark lens is an isolated white dwarf. Direct verification of these parameters will be possible with the epoch astrometry expected with Data Release 4.

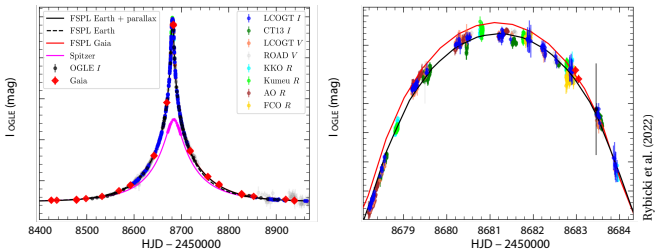


ANOTHER OF Gaia's microlensing alerts discoveries is Gaia19bld. This was announced on 2019-04-18, as it brightened, at $G = 14.44$ mag (compared with the historic $G = 14.82 \pm 0.05$ mag). The Gaia light curve is well defined, with nearly 200 observations between 2014–23, and reaching a peak of $G = 10.5$ mag, three months after the alert was issued. Photometry from the Las Cumbres Observatory (LCO) established it as a microlens event, with LCO–NRES spectroscopy showing features consistent with a K3 supergiant (Rybicki et al., 2019).

The interest of Gaia19bld (with $t_E \sim 107$ d) is that, for the first time, the rotating arc-like sub-images have been resolved using the ESO VLT–PIONIER interferometer.

I SHOULD START with a brief summary of the extensive analysis by Rybicki et al. (2022), which provides the context. They assembled a large collection of photometric measurements of the ongoing event, including by the Spitzer space telescope. Together these defined the entire high-magnification event (left image), as well as resolving the event peak (right image). The best-fit model is found for a point lens, but with a finite source size.

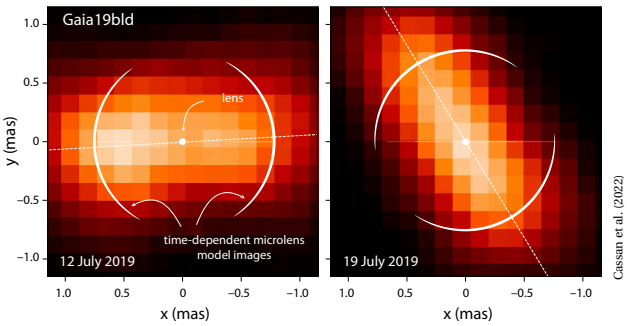
The microlens parallax could be inferred from the difference in light-curve profile seen from Earth and from Spitzer (which, in its Earth-trailing orbit, was at a distance of about 1 au). Their derived parameters gave a (dark/unseen) lens mass, $M_L = 1.13 M_\odot$, at a distance $D_L = 5.52$ kpc. Similar results were found from the spectroscopic follow-up by Bachelet et al. (2022).



THE INDIVIDUAL images of the source in microlensing events are unresolved by monolithic ground-based telescopes, even the largest, and the changing brightness (and position) of an event is simply from the superposed (and changing) flux of the individual images. For example, Gaia19dke is unresolved even at the 20 milliarcsec inner working angle of Gemini North 8-m with the Alopeke speckle imager (Maskoliūnas et al., 2024).

Measuring the image separation would provide another route for determining the lens mass. And note that these (normally unresolved) sub-images rotate on the sky with time as the event alignment changes.

Prospects for resolving the individual images, or at least measuring their angular separation, has improved with the development of long-baseline interferometry (Delplancke et al., 2001; Dalal & Lane, 2003; Rattenbury & Mao, 2006; Cassan & Ranc, 2016; Cassan, 2023).



THE FIRST RESOLVED IMAGING of a microlensing event used the VLTI–GRAVITY interferometer for the ASAS-alerted non-bulge system TCP J0507+2447 (Dong et al., 2019). While not providing a 'picture' as such, the instantaneous (single-epoch) separation of the individual lensed images of the background source was determined from the interferometric 'closure phase'. They derived an image separation 3.78 ± 0.05 mas, and hence the Einstein radius $\theta_E = 1.87 \pm 0.03$ mas.

More strikingly, Cassan et al. (2022) used the VLTI–PIONIER interferometer (Le Bouquin et al., 2011) to observe Gaia19bld. They used the four ESO 1.8-m Auxiliary Telescopes, with baselines up to 128 m, on three dates in July 2019 (just two are shown above). The source was split into two images on either side of the lens, and the arc-like images are seen to be rotating around the lens.

From the image separation, they derived the angular Einstein radius $\theta_E = 0.0818 \pm 0.0020$ mas, in agreement with the Earth–Spitzer (microlens parallax) value $\theta_E = 0.0823 \pm 0.0018$ mas by Rybicki et al. (2022), and subsequent spectroscopic observations by Bachelet et al. (2022). These spectroscopic observations also characterised the source as a red giant of radius $40 \pm 10 R_\odot$.

Combining θ_E and the microlens parallax ϖ_E yields the microlens mass, $M_L = 1.147 \pm 0.029 M_\odot$, which they were able to measure to an unprecedented accuracy compared with other microlensing events. The model fits also provide distance estimates to both the source, $D_S = 8.4 \pm 1.5$ kpc, and the lens, $D_L = 5.5 \pm 0.6$ kpc.

LET ME SUMMARISE. The Gaia alerts pipeline is able to identify microlens systems outside the bulge region, some with relatively nearby (1–2 kpc) sources, and of unusually long duration. In favourable circumstances, the Gaia light curve can contribute to determining the microlens parallax, ϖ_E, while progress in ground-based interferometry is allowing the direct determination of the Einstein radius θ_E, rather than waiting several years to resolve the source–lens separation as the alignment degrades (e.g. Bennett et al., 2006).

In many of these events observed by Gaia, the forthcoming availability of the intermediate astrometry with Data Release 4 (essay 11) will be another important milestone in event classification.

206. Alerts – and tidal disruption events

I WILL CONTINUE with the theme of my last four essays, Gaia's science alerts (Hodgkin et al., 2021), and look here at the remarkable class of 'tidal disruption events'.

In September 2024 the Gaia Science Alerts database listed 25 919 alerts. Of these, 7 141 were assigned to 23 classes, with 25 inferred to be tidal disruption events. Before saying more about two of these events, let me summarise the phenomenon, their discovery history, and the insights that are being gained from them.

TIDAL DISRUPTION EVENTS, or TDEs, result from a star passing sufficiently close to a supermassive black hole that tidal forces overcome its self-gravity. Drawn-out material results in a tidal stream that loops around the black hole, some fraction on unbound orbits, the remainder forming an accretion disk, with bursts of electromagnetic radiation (spanning radio to γ-ray), fading over several months as the material is accreted (e.g., Carter & Luminet, 1983; Evans & Kochanek, 1989).

Their existence was predicted, 50 years ago, when supermassive black holes were still only hypothesised (Hills, 1975; Lacy et al., 1982; Rees, 1988). Interest in them was as a possible fuel source for quasars, with extreme events perhaps explaining some γ-ray bursts.

THE FIRST candidates were found with ROSAT (1990–99). Soft X-ray outbursts from otherwise quiescent galaxies were interpreted as the formation of the implied accretion disk (e.g., Bade et al., 1996; Komossa & Bade, 1999). Transient emission was discovered in the ultraviolet with GALEX (Gezari et al., 2006; Gezari et al., 2008), and at γ-ray wavelengths by the Neil Gehrels Swift Observatory (e.g., Bloom et al., 2011; Komossa, 2015).

Optical emission was first reported from archival SDSS images (van Velzen et al., 2011), and later from the surveys Pan–STARRS (e.g., Gezari et al., 2012; Chornock et al., 2014; Holoien et al., 2019), and PTF (e.g., Bloom et al., 2012; Arcavi et al., 2014). The large transient surveys ASAS–SN (e.g., Holoien et al., 2014; Holoien et al., 2016) and ZTF (e.g., Nicholl et al., 2020; van Velzen et al., 2021) have since found many more.

DISCOVERIES HAVE now been made in the infrared (e.g. van Velzen et al., 2016), and most recently in larger numbers from NEOWISE (Masterson et al., 2024). New discoveries in X-rays have been possible with eROSITA (Sazonov et al., 2021; Khorunzhev et al., 2022).

The first detection at radio wavelengths was made within 24 hr of the (Neil Gehrels Swift Observatory) discovery of Sw J1644+57 (Zauderer et al., 2011). Several dozen have since been detected as follow-up to alerts at optical wavelengths (Alexander et al., 2020), with the first radio discovery by Anderson et al. (2020).

There is an extensive literature on event modelling based on multi-wavelength light curves (e.g., Piran et al., 2015; Shiokawa et al., 2015; Guillochon & Ramirez-Ruiz, 2015; Bonnerot et al., 2016; Hayasaki et al., 2016; Mockler et al., 2019). One conclusion is that spectral properties depend on the viewing angle (Dai et al., 2018).

Meanwhile simulations show, for example, that because the tidal and Schwarzschild radii have different dependences on black hole mass, solar-type stars are accreted whole for black hole masses $\gtrsim 10^8 M_\odot$ (other than by the most rapidly spinning), while white dwarfs can be tidally disrupted for black hole masses $\lesssim 10^5 M_\odot$.

THE GROWING number of discoveries is today allowing increasingly detailed population studies (e.g. Gezari, 2021; Sazonov et al., 2021; van Velzen et al., 2021; Hammerstein et al., 2021; Yao et al., 2023).

In turn, these studies are raising questions about the nature of their host galaxies, and (for example) their suggested prevalence in systems which have undergone recent mergers. And there are outstanding questions about the observed versus predicted occurrence rate, the latter being of order 10^{-4} per galaxy per year (van Velzen & Farrar, 2014; Stone & van Velzen, 2016; Graur et al., 2018; Hinkle et al., 2020).

Today, tidal disruption events are probing the properties of previously dormant supermassive black holes, including their mass and spin (Kesden, 2012; Mockler et al., 2019; Pasham et al., 2019; Wen et al., 2022), and their nuclear environments (Alexander et al., 2020).

As of mid-2024, more than 100 tidal disruption events are known. The 25 discoveries from the Gaia alerts pipeline were typically alerted at $G \sim 17-18$ mag, compared with their historical 18–19 mag. Some were reported independently by the other transient surveys.

Eleven are 'confirmed TDEs' (Gaia19bvo, Gaia19eks, Gaia20ead, Gaia20fqa, Gaia22bdt, Gaia22cgf, Gaia22clk, Gaia22cwy, Gaia23cbw, Gaia24awm, Gaia24beb), and others are simply noted as 'coincident with known galaxies' (Gaia18dpo, Gaia20cxg, Gaia21drx, Gaia22anq, Gaia22dgl). In the following I will look at the two brightest.

Since 2016, the IAU Transient Name Server assigns names to confirmed supernovae, and 'Astronomical Transient' (AT) designators to other transients, including TDEs. And a literature search for a Gaia discovery (GaiaYYnnn) may only find that event under the ASASSN, ATLAS, ZTF, or ATYYnnn identifier.

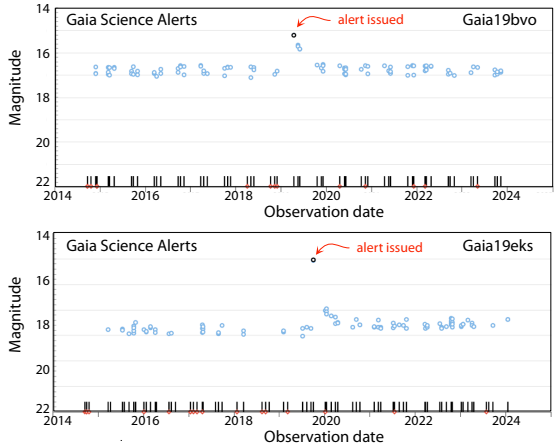
The brightest of Gaia's TDEs, Gaia19bvo, alerted at $G = 15.22$ mag on 2019 April 13 (historic $G = 16.81 \pm 0.14$ mag), was also discovered as ZTF17aaazdba/ASASSN–19dj, and is also referenced as AT2019azh.

The 'Astronomer's Telegram' #nnnnn (ADS reference 2019ATelnnnnn) reported spectroscopy from the dedicated transit instruments NOT–NUTS (#12529) and ESO NTT–ePESSTO (#12530), both showing a featureless blue spectrum superimposed on a $z = 0.022$ (96 Mpc) galaxy.

Swift observations (#12568) found a 15-d optical/UV plateau, at a temperature of 30 000 K. The TDE classification is based on the event location in the galaxy centre, its blue colour, high blackbody temperature, and absence of supernova-like spectral features. The plateau suggests Eddington-limited accretion onto a $4 \times 10^6 M_\odot$ black hole. Radio emission was detected by MERLIN at 5 GHz on 2019 May 21 and June 11 (#12870), and by the (upgraded) GMRT at 1–1.45 GHz on November 8 (#13356). It was detected as a 'late-time brightening' X-ray source by ISS–NICER on October 19 (#13221).

High-cadence optical photometry and spectroscopy from –21 to 392 d relative to the peak emission was reported by Hinkle et al. (2021). For the first 16 d, the rise was consistent with a $L \propto t^2$ power law, peaking at a luminosity $L = 6.2 \times 10^{37}$ J s^{-1}. The X-ray flux increased by an order of magnitude 225 d after the peak, resulting from the expansion of the X-ray emitting region, with the late-time X-ray emission well matched by a blackbody with radius $r \simeq 10^{10}$ m, and temperature $T \simeq 6 \times 10^5$ K.

In the radio, more than two years of monitoring with VLA and MeerKAT, from –10 d to +810 d with respect to the optical peak, was reported by Goodwin et al. (2022). The source brightened slowly over 2 yr, showing fluctuations in the synchrotron emission from 450 d post-disruption. They deduced that the outflow is non-relativistic, and explicable as a spherical outflow from self-stream intersections, or by a mildly collimated outflow from accretion onto the supermassive black hole.



The second brightest of Gaia's TDEs, Gaia19eks, and the closest at the time, alerted at $G = 16.04$ mag on 2019 October 3 (historic $G = 18.79 \pm 0.12$), was a 3 mag brightening of a galaxy, and in the footprint of a gravitational wave event. Referenced as AT2019qiz, it was also discovered as ZTF19abzrhgq/ATLAS19vfr/PS19gdd.

The subsequent 'Astronomer's Telegrams' reported Keck–LRIS spectroscopy, classifying it as a rising TDE at $z = 0.015$ (#13131), along with optical/UV brightening and later decay by Swift (#13146, #13193), X-ray variability with Swift and ISS–NICER (#15217), and variable radio emission with ATCS (#13310, #13334).

Nicholl et al. (2020) found that the velocity dispersion of the host galaxy, combined with fits to the event light curve, indicate a star of mass $\simeq 1 M_\odot$ disrupted by a black hole of mass $\simeq 10^6 M_\odot$. The extensive ultraviolet, optical, and X-ray data indicate an early optical emission dominated by an outflow, with a luminosity evolution $\propto t^2$, consistent with a photosphere expanding at $v \geq 2000$ km s^{-1}. The light-curve rise begins 29 d before maximum light, peaking when the photosphere reaches the radius where optical photons can escape.

Short et al. (2023) found that between the optical flare and subsequent observations, the X-ray spectrum softened dramatically, with the 0.3–1 keV flux increasing by a factor 50, and the hard X-ray flux decreasing by a factor 6. WISE (infrared) fluxes rose over the same period, indicating the presence of an infrared echo. Detailed numerical modelling is described by Kovács-Stermeczky & Vinkó (2023).

With 25 detections to date, Gaia is discovering significant numbers of tidal disruption events, although most of them are also being reported by other transient surveys. Gaia's discoveries are contributing to the event statistics as well as to the light-curve characterisation, and Gaia is providing the accurate reference frame which is essential in associating these events with (the centres of) the previously quiescent galaxies.

207. Gaia science synopsis to 2024

In this ante-penultimate essay of 2024, I will give a concise 'run through' of the scientific areas impacted by Gaia so far. I divide this into five categories: solar system; applications of the photometry and radial velocity spectrometer (RVS) spectra; stellar physics; Galaxy structure and dynamics; and local group and cosmology. In my next essay (208), I will focus on a few highlights.

References for such a concise overview would be challenging – and none are given. To follow-up on these many topics, my 'Gaia Science Tree' links to the various essays giving more details.

Solar system: Data Release 3 (DR3, June 2022) gave an all-sky catalogue of 1.8 billion sources, providing a reference system for the pointing of ground-based telescopes, for interplanetary spacecraft navigation (e.g. the New Horizons flyby of Arrokoth in 2019), and for predicting occultation events in unprecedented numbers (including Europa and Triton), with insights into object sizes, shapes, and atmospheres. The positions also provide an absolute reference frame for the measurement of gravitational light-bending by the Sun and Jupiter.

DR3 gave reconstructed orbits for 155 000 minor solar system bodies, based on 34 months of data. One of DPAC's Focused Product Releases (Oct 2023) gave improved results over the 66-month coverage planned for DR4. These orbits are yielding further insights into their migration and their complex dynamical resonances.

Accurate photometry is yielding the solar-radiation driven Yarkovsky and YORP effects for several hundred near-Earth asteroids, throwing light on their orbital migration, the origin and orbits of Near-Earth Asteroids, and their spin states. Reflectance spectra for 60 000 are advancing studies of their classification, and collisional and orbital evolution, where the effects of space 'weathering' are seen as an increase in spectral slope with age.

The positions and velocities of nearby stars are being used for determining the chronology of stellar flybys and their correlation with past extinction events on Earth, and in efforts to trace back the origins of stellar 'vagabonds', exemplified by Oumuamua and Borisov.

Photometry and spectra: Gaia's multi-epoch photometry is allowing variability classification for tens of millions of stars, of all spectral types, including Cepheids, non-radial pulsators, and eclipsing and ellipsoidal variables, with stellar rotation determined for millions of stars. It provides the basis of an alerts system, which is flagging microlensed events, supernovae, tidal disruption events, and others.

The radial velocity spectrometer is providing tens of millions of radial velocities. The spectra are yielding stellar parameters ($T_{\rm eff}$, $\log g$, and [M/H]), extinction measures, mapping of diffuse interstellar bands, and synthetic photometry in arbitrary passbands.

Constraints on physical effects include the determination of gravitational redshift, asteroseismology models, and limits on the variation of $\dot{G}$, on gravitational waves, and effects of modified (MOND-type) gravity.

Stellar physics: Gaia is providing an unparalleled census of nearby stars (to 25 pc, 50 pc and 100 pc). This is allowing searches for solar analogues and solar siblings, and mapping out the Local Bubble, diffuse interstellar bands, and the effects of extinction.

The substantial improvement in content and accuracy of the Hertzsprung–Russell diagram, over all evolutionary phases, is providing a massive framework for advances in modelling stellar evolution, determining the initial mass function, constraining important effects such as convection (and the mixing length), and identifying subtle features such as the Jao gap in M dwarfs.

Open clusters are being discovered and mapped in unprecedented numbers. This is revealing the structure (and main sequence) of the Hyades, the distance to the Pleiades, and the nature of complex clusters such as Westerlund 1. Gaia is probing the origin of OB associations, moving groups and their traceback ages, and the structure of nearby molecular clouds.

Accurate distances and proper motions are allowing the identification and characterisation of runaway stars and their birth clusters, and the identification of many new hypervelocity stars, and determining their origin.

Gaia's multiple star census has yielded many new results, based on an unparalleled survey out to 1000 pc. This includes samples of spectroscopic binaries, of twin binaries, triple and quadruple star systems, eclipsing binaries, and cataclysmic variables. There are new results on stellar masses from SB2 binaries, and on the nature and origin of wide and ultra-wide binaries. And on variable stars including Cepheids, RR Lyrae, and Mirae.

Studies of white dwarfs have been revolutionised, with deeper and more complete surveys, new results on the mass–radius relation, on their interior physics, and the detailed structure of their HR diagram, including the effects of core crystallisation, and of convective dredge-up as the cause of the prominent bifurcation.

Studies of many other star types include Wolf–Rayet stars, brown dwarfs, planetary nebulae, carbon stars, S stars, and the nature of the tip of the red giant branch; also of young stellar objects and chromospheric activity.

Various studies of the end stages of stellar evolution include neutron stars and pulsars, supernova remnants, nearby black holes in non-active binaries, and black holes in open clusters, stellar streams, and in M4.

For exoplanets, distances are being used to characterise host stars, calibrate exoplanet radii, and pinpoint locations for direct imaging. Gaia's first two transiting exoplanets, and first microlensed exoplanet, have been reported. Constraints are being placed on the nature of Boyajian-type stars, and the existence of Dyson spheres.

GALAXY STRUCTURE AND DYNAMICS: Improved structural insights include the Sun's height above the disk, its motion with respect to the Local Standard of Rest, and the distance to the Galactic centre.

Gaia has provided an improved understanding of the morphology and dynamics of the central bar, the spiral arms and their 'breathing motion', the global warp, and a new structural feature termed the 'Radcliffe Wave'.

Insights have been gained into the resonance origin of velocity structures in the solar vicinity, including the Hercules, Arcturus, and HR 1614 streams, and the complex Gaia 'phase-space spiral' perturbations.

Major advances have been made on the discovery and characterisations of stellar streams in the halo, notably for the Gaia-Sausage-Enceladus, Sagittarius, and Cetus streams. Other clues about our Galaxy's formation include studies of its ancient 'heart', and its infall history inferred from cerium abundances.

Gaia is elucidating the structure and motion of globular clusters (including Omega Cen, and Palomar 5), of their tidal tails, and of our companion dwarf spheroidal galaxies. It is providing an improved rotation curve of our Galaxy, and revealing the tiny effects of aberration arising from our Galaxy's rotation with respect to quasars. It is providing improved knowledge of the local mass density, total mass, escape velocity, and age.

LOCAL GROUP AND COSMOLOGY: With more than 10 million Gaia stars in the Large Magellanic Cloud, nearly two million in the Small Magellanic Cloud, and more than a million in M31/Andromeda, there has been much progress in clarifying the morphology, rotation, interaction, and disruption of our neighbouring galaxies. Improved mass estimates of the Local Group are being derived from dynamical modelling, the 'timing method', and from cosmological N-body simulations.

As noted above, an important contribution from Gaia has been the identification and modelling of many new stellar streams over a range of distances within the halo. Tests of the ΛCDM cosmological paradigm follow from the comparison of these observations with numerical simulations. As part of these, halo streams are also being probed for the presence of black holes, and cosmological 'sub-halos'.

Gaia's global survey embraces not only stars. Gaia's overall survey, all part of DR3, contains more than a million galaxies (whose radial profiles can be determined), and amongst which are found a number of dual active galactic nuclei, signalling past galaxy mergers.

Also identified by Gaia are around one million quasars. These provide the key sample of objects defining Gaia's quasi-inertial reference frame. More than 60 000 show clear evidence for their underlying host galaxy. The quasar sample includes many previously known, as well as newly identified, strongly lensed quasars, including the important class of quadruply-imaged 'Einstein crosses'. Of several cosmological investigations based on Gaia's quasar, the present DR3-based results appear to be compatible with the Planck value of the 'matter overdensity variance', S_8.

Contributions have been made to the problem of the cosmic distance scale (aka the 'Hubble tension') based on the local Universe expansion rate derived from Cepheids. A number of studies have focussed on supernovae, and supernovae remnants and their associated pulsars (including Cygnus, Vela, and Tycho). Searches for the orbital motion of luminous objects around a dark companion should throw more light on the many complexities of the late stages of binary star evolution.

The rich phase-space complexities revealed in the Gaia data, indicators of past galaxy mergers and tidal disruptions, are being tested against the latest suites of large cosmological simulations. Such tests include the occurrence of Gaia-Sausage-Enceladus type mergers in these simulations, as well probing the overall merger and spin-up history of our Galaxy.

Contributions to other cosmological challenges include the 'plane of satellites' problem, and the 'core–cusp' problem, further elucidated by Gaia's discovery of the enormous dwarf galaxy Antlia II in 2019. Another test that may be possible is whether our Galaxy's disk is indeed 'tumbling' with respect to the halo.

208. Gaia science highlights to 2024

In my previous essay, I gave an overview of the many scientific areas being impacted by Gaia, divided into five categories: solar system; uses of the photometry and RVS spectra; stellar physics; Galaxy structure and dynamics; and local group and cosmology.

Here, I will say a little more on some of the highlights, with references to my earlier essays for further details.

Solar system: Gaia has provided state-of-the-art orbits of 157 000 asteroids for the 66-month data interval used for Data Release 4 [159], reflectance spectra from the BP/RP spectrophotometry [180], and rotation periods from Gaia's multi-epoch photometry [181, 182].

This combination is being used to study their dynamical and taxonomic properties; masses from mutual orbit perturbations; asteroid 'families' resulting from collisional fragmentation (which can remain clustered in orbit space, with similar spectra); space 'weathering'; the origin of near-Earth asteroids via solar-radiation driven (Yarkovsky) migration; the excess of fast and slow rotators as a result of the solar-radiation driven (YORP) effect; and sizes and morphologies from the many occultation events that can now be accurately predicted.

Stellar physics: The enormous and strictly defined census of stars within the solar neighbourhood, including the CNS5 within 25 pc [129], and the Gaia Catalogue of Nearby Stars within 100 pc [33], yield the accurate placement of more than 200 000 stars in the colour–magnitude (M_G versus $G_{BP} - G_{RP}$) analogue of the Hertzsprung–Russell diagram [42]. Many advances are being made by connecting these observations with theoretical models of stellar structure and evolution.

Variable stars have long been recognised as offering deep insights into stellar structure and evolution, and more than 10 million variable sources have been classified as part of Data Release 3 [76]. Accurately locating them in the HR diagram reveals the specific occurrences of (for example) pulsating, eruptive, and cataclysmic variables, as well as stars that show apparent variability due to stellar rotation or binary eclipses. A movie of the changing loci of representative variable stars across the colour–magnitude diagram has been given at `cosmos.esa.int/web/gaia/gaiadr2_cu7`.

As one example, the classical ZZ Ceti stars (white dwarfs featuring fast non-radial gravity-mode pulsations) are particularly concentrated in magnitude and colour, with variability seen in about half of them. This concentration is attributed to the partial ionisation of H in their outer envelopes, which is only developed over a narrow range of temperature, and therefore colour.

The fidelity of both the Gaia astrometry and photometry is nicely demonstrated by three striking new features of the Hertzsprung–Russell diagram.

The first of these is a clear discontinuity at the location of the radiative–convective boundary of M dwarfs, today also known as the Jao gap. It has recently been explained in terms of the non-equilibrium burning and mixing of ^{3}He in the stellar core [152].

The second is the prominent bifurcation in the hydrogen- and helium-dominated atmospheres of the white dwarf sequence. Not predicted by previous evolutionary models, this is now attributed to trace amounts of carbon added by convective dredge-up [178].

Another feature addressed in several follow-up papers is a discontinuity in the white dwarf sequence now attributed to ongoing core crystallisation. This is a phase transition that holds up the release of heat, leads to a discontinuity in the rate of cooling, and affects the inferred 'cooling-age' estimates of white dwarfs [108, 139].

More generally, the substantial increase in white dwarf numbers from Gaia [29] is leading to many other insights into their physics, including the mass–radius relation, their asteroseismic properties, and mass-based inferences on stellar mergers [107, 108].

On another topic, many new insights into binary and multiple stars are flowing from Gaia's joint astrometric, photometric, and spectroscopic survey. Amongst these are the occurrence of equal-mass 'twin binaries' [138], and the detailed orbital characterisation of wide binary systems, allowing studies of their origin [193]. Both have important implications for star formation models.

GALAXY STRUCTURE AND DYNAMICS: With Gaia representing a step-change in the observational characterisation of our Galaxy's origin and evolution, the topics of Galactic structure and cosmology now overlap considerably. I will defer the mention of halo streams, although intimately tied to our Galaxy's structure and dynamics, to my next 'cosmology' section.

I will gather the whole subject of star formation under the first highlight of this section. Here, big progress has been made in identifying many new open star clusters [74, 144], and further characterising some of the best known clusters and associations [18]. Amongst these are the superbly-defined main sequence of the Hyades [151]; resolution of the distance controversy of the Pleiades [13]; the identification of various tidal tails [20]; their possible harbouring of black holes [175]; the origin of runaway stars [165], and the somewhat connected hypervelocity stars [22, 166].

The 'peculiar motion' of nearby stars has been known since the time of William Herschel, and the detailed nature of various nearby moving groups and dynamical streams has long been debated. Many studies with Gaia have confirmed that some are 'evaporating' open clusters, and some are the tidal debris of accreted satellite galaxies. Of many new results, the Hercules stream has been associated with a dynamical resonance with the Galaxy's central bar, perhaps even consistent with the bar's deceleration through dynamical friction with the Galaxy's dark halo [115].

Amongst several other profound dynamical insights are the discovery of the Gaia phase-space spiral, a remarkable large-scale feature possibly associated with a passage of the Sagittarius dwarf galaxy [117]; a dynamical estimate of the distance to the Galactic centre [111]; measurement of the micro-acsec level effects of aberration due to Galactic rotation [32]; and an increasingly coherent picture of our Galaxy's three-phase formation, involving spin-up, merger, and cooldown [190].

Data Release 3 in 2022, I should recall, included a catalogue of distances, metallicities, temperatures and gravities for 470 million sources [89], while at least four other community-generated catalogues of [M/H], $T_{\rm eff}$, and $\log g$ have been made available since [189]. DR3 also provides estimates of radius for 470 million sources, mass (140 million), age (120 million), chemical abundances (5 million), diffuse interstellar bands (0.5 million), activity indices (2 million), Hα equivalent widths (200 million), and further classification of spectral types (220 million) and emission-line stars (50 000).

Such huge numbers of uniform metallicities and other stellar properties, which provided accurate chemical abundances across all stellar populations throughout the Galaxy, are now being used as crucial inputs for studies of star formation, detailed nucleosynthesis modelling, and Galactic chemical and dynamical evolution.

LOCAL GROUP AND COSMOLOGY: It is the number and variety of complex phase-space features in our Galaxy that has been one of Gaia's greatest contributions to studies of its structure and evolution to date. And some of these can only easily be interpreted within the context of our Local Group of galaxies, and indeed within the framework of ΛCDM cosmology.

Of those phase-space features with a more cosmological origin, I have already mentioned the Gaia phase-space spiral, a complex phase-space feature possibly associated with a passage of the Sagittarius dwarf galaxy [117], and the possible imprints of a decelerating bar attributable to the Galaxy's dark matter halo [112].

Over the past 20–30 years, through a combination of observations, theory, and simulations, aided by N-body simulations of the large-scale structure of the ΛCDM Universe, it became clear that our Galaxy's halo should be the graveyard of many other galaxies, gravitationally captured by the Milky Way galaxy over its lifetime.

Within this framework is one of Gaia's key advances: the ongoing discovery of large numbers of ancient stellar streams making up our Galaxy's halo. While a few were suspected pre-Gaia, more than 100 are known today, confirming this picture of halo accretion, and their influence on (for example) the dynamics of the thick disk [156]. Many detailed insights are flowing from these discoveries, and the detailed kinematic and chemical data from Gaia associated with them.

The 'Gaia Sausage–Enceladus' stream, lying within 25 kpc, is one of the most prominent and well studied. It was a radial merger that occurred some 9.5 Gyr ago, and which now dominates the local metal-poor halo. The 'head-on' collision resulted in rapid phase mixing, evident today only in its clustered integrals of motion. Beyond 25 kpc, the halo is dominated by the Sagittarius (Sgr) dwarf galaxy and stream, [Fe/H] = −1.0, one of the first identified halo streams. Here, both the residual galaxy core, as well as two vast preceding and trailing streams, are clearly discernable.

The phase-space information provided by Gaia is also allowing the search for black holes within these stellar streams [176], and probing the observational consequences of the dark matter sub-halos (halos within halos) that are predicted, in the standard ΛCDM cosmology, to exist surrounding the Milky Way [184].

More widely, Gaia is confirming many of the detailed predictions of the large-scale cosmological simulations such as Millennium, Illustris, and EAGLE, and helping in their interpretation [194]. Topics range from simulations of the merger epochs, the occurrence of bars, the orbits of globular clusters [30] and dwarf spheroidals [31], the bulk motions of the Magellanic Clouds [38], as well as the ΛCDM 'missing satellites' problem, the 'core–cusp' problem, the 'too-big-to-fail' problem, and the 'plane of satellites' problem.

209. A selection of journal plaudits

I WILL END this fourth year of my weekly essays on Gaia with some of the positive words that have appeared in the scientific literature about the mission.

I hope that all scientists involved, all who worked on Hipparcos and Gaia in ESA and industry, members of the Gaia Data Processing & Analysis Consortium, funding agencies, and members of ESA's advisory committees, will appreciate this wider perspective of the impact that Gaia is having across astronomy.

I will start with just a few comments on Gaia's predecessor, Hipparcos, which – crucially – established the foundations and principles of space astrometry.

HIPPARCOS

Hipparcos is the first time since Sputnik in 1957 that a major new development in space science has come from outside the United States
Freeman Dyson, *'Infinite In All Directions'*, 1988

The scientific impact of the Hipparcos mission is only now beginning to be felt. Perhaps the most inspiring aspect of this effort is the staggering degree of cooperation required
Janet Mattei, AAVSO (Sky & Telescope, July 1997)

Our Galactic star precinct has just been well mapped for the first time, ready for a century of searching stars for the promise of life. A terabit of data from the European Space Agency satellite Hipparcos underlies a magnificent list of over a hundred thousand star distances of unprecedented accuracy out to nearly five hundred light-years
Philip Morrison (Scientific American, February 1998)

Hipparcos was one of our most distinctive all-European missions. No other space agency has attempted anything like it Roger–Maurice Bonnet (Bern SPC, May 1999)

The bedrock of astronomy remains the compilation of what is out there… It is invidious to single out surveys which I find particularly impressive, but I make an exception for the Hipparcos astrometric satellite
Malcolm Longair (Millennium Essay, PASP, 2001)

GAIA IN GENERAL

Galactic astrophysics is currently in a similar phase as geography was in the 15th century: large parts of the Earth were unknown to contemporary scientists, and only crude maps of most of the known parts of the Earth existed… In this context, the astrometric ESA mission Gaia represents a major leap in our understanding of the Milky Way's stellar content Anders et al. (2019)

It's a fundamental rewriting of how we do positional astronomy Marc Buie, Science Magazine (24 Nov 2020)

The stunning revolution being operated by Gaia has often been mentioned during the conference and examples have been shown in a number of talks
Clementini et al. (2020)

Gaia Early Data Release 3 opens a new chapter in the measurement of parallaxes Soltis et al. (2021)

The next generation will never know the struggle to determine a distance or a spectral type thanks to you and the Gaia team Chas Beichman (priv. comm., July 2022)

The ESA Gaia space mission is one of the most successful projects in the history of astronomy
Cantat-Gaudin (2022)

Gaia provides a revolutionary step forward in our understanding of the true locations of stars in the Milky Way
Davenport et al. (2022)

Gaia has revolutionised astrometry Munn et al. (2022)

Among surveys, Gaia holds a unique place because it provides high precision astrometric, photometric, and spectroscopic data for an unprecedented number of sources
Tsantaki et al. (2022)

Gaia has revolutionised astronomy with its high-precision astrometry T.J. Wilson (2023)

One of the most remarkable space science missions that most people have never heard of is Gaia: a mission by the European Space Agency to map out the three-dimensional positions and motions of more stars in the galaxy than ever before
Ethan Siegel, *'Ask Ethan'*, 11 August 2023

Gaia Data Release 3 provides data on positions, parallaxes, and proper motions for a staggering 1.46 billion celestial objects Bisht et al. (2024)

We are now deep into the era of Gaia – the most ambitious Galactic science mission of all time
Deason & Belokurov (2024)

The staggering Gaia DR3 released low-resolution spectra for 220 million stars to G = 17.6 mag, and medium-resolution spectra for one million stars to G = 13 mag
Viswanathan et al. (2024)

Stellar structure and evolution

In the Gaia era, stellar evolution modellers have the daunting task of explaining an increasing number of precisely determined features in the colour–magnitude diagram Davenport & Covey (2018)

Gaia is opening an important new window in the study of the interstellar medium Großschedl et al. (2018)

Another ongoing, and perhaps more important revolution in the white dwarf field is the Gaia mission, which will probably discover about 400 000 white dwarfs when the mission is completed Bergeron et al. (2019)

We address two specific fields where Gaia is really astonishing: the detailed monitoring of stellar populations in different evolutionary phases, and the distance scale
Clementini et al. (2020)

The exquisite quality of the data produced by Gaia is revolutionising our knowledge of the Milky Way, allowing us to resolve and characterise its stellar populations to an unprecedented level of detail Marchetti et al. (2022)

Great progress has been made in [the structure of giant molecular clouds] with the arrival of the Gaia mission
Dharmawardena et al. (2023)

Gaia DR3 contains a stellar activity index derived from the Ca II infra-red triplet for some 2 million stars in the Galaxy. This represents a 'gold mine' for studies on stellar magnetic activity and mass accretion in the solar vicinity
Lanzafame et al. (2023)

Since 2018, data from the Gaia mission are revealing previously unseen and often unexpected 3D distributions of gas, dust, and young stars in the solar neighbourhood
Zucker et al. (2023)

Gaia Data Release 3 represents an unparalleled revolution in Galactic archaeology, providing us with radial velocities and chemical abundances for million of stars, overcoming the spatial biases suffered by spectroscopic surveys from the ground Spitoni (2023)

Gaia has revolutionised the calibration of Leavitt's Law [for Cepheids] in the Milky Way based on trigonometric parallaxes Anderson (2024a)

The advent of Gaia has made it possible to construct ever-more accurate 3D dust-based distances to molecular clouds Cahlon et al. (2024)

This review highlights the role of the Gaia space mission in transforming white dwarf research
Tremblay et al. (2024)

Star clusters

This study shows the incredible wealth of data provided by Gaia for the study of young stellar clusters
Beccari et al. (2018)

Gaia has opened a new window into the internal kinematics of young star clusters at the sub-km s^{-1} level
Kuhn et al. (2019)

The census of open clusters in the Milky Way is in a never-before seen state of flux. Recent works have reported hundreds of new open clusters thanks to the incredible astrometric quality of the Gaia satellite
Kuhn et al. (2019)

The Gaia data have unlocked a deluge of new results related to many astronomical topics, and transformed our ability to study star clusters and stellar structures in the Milky Way Cantat-Gaudin (2022)

Our knowledge of the memberships of open star clusters provides the foundation for a wide range of topics in stellar astrophysics – and Gaia has revolutionised the study of clusters Kastner (2023)

The Gaia revolution has also allowed for membership studies of open clusters on the Galactic scale including thousands of clusters Fritzewski et al. (2023)

It is probably safe to assume that no single instrument will have, in the near future, a transformative impact comparable to Gaia
Cantat-Gaudin & Casamiquela (2024)

The excellent photometry and distances from Gaia have provided systematic isochronal ages for hundreds of previously known clusters and associations and thousands of Galactic stellar ensembles that have been recently identified Miret-Roig et al. (2024)

Star multiplicity

Gaia has revolutionised our understanding of stellar multiples, stellar associations, and the 3D structure of the solar neighbourhood Cunningham et al. (2020)

The space astrometry mission Gaia has revolutionised wide binary research Hwang et al. (2022)

Gaia has revolutionised the census of solar neighbourhood by revealing wide pairs of stars with an unprecedented completeness Tokovinin (2023)

Much of the progress in the last decade has been enabled by the Gaia mission. The data probe a wider range of binary separations and mass ratios than most previous surveys, enabling both an improved binary population census and discovery of rare objects El-Badry (2024)

The discovery of co-moving wide binary candidates in the Milky Way has significantly increased based on high-precision parallax and proper motion solutions for a vast number of stars provided by Gaia Lim et al. (2024)

Photometry and Variability

Gaia is a unique transient survey, detecting supernovae, tidal disruption events, cataclysmic variables, and microlensing events Wyrzykowski et al. (2018)

In principle and in perspective, Gaia synthetic photometry may constitute a true revolution in optical photometry Montegriffo et al. (2023)

Gaia is now one of the most successful and leading transient space missions Kvernadze et al. (2023)

Variable star research is about to undergo a revolution driven by large time-resolved imaging surveys such as ESA's Gaia mission Anderson (2024b)

Galaxy structure and dynamics

The field of Milky Way dynamics is reaching an incredibly exciting time, as the successful launch and operation of ESA's Gaia satellite means that we have access to proper motions and parallaxes for a billion stars. This represents an increase of four orders of magnitude over the number of stars with known parallaxes from Gaia's predecessor, Hipparcos. This information will revolutionise how we understand our own Galaxy, and by extension, the Universe as a whole McMillan (2017)

Gaia stellar measurements are currently revolutionising our knowledge of the evolutionary history of the Milky Way Lallement et al. (2019)

The Gaia Early Data Release 3 direct measurement of the acceleration of the solar system within the Milky Way galaxy from the apparent proper motion pattern of distant quasars that it induces is a revolutionary moment for Galactic astrophysics Bovy (2020a)

Since its first data release in 2016 and especially since its recent second data release in 2018, Gaia has impacted nearly every field in astrophysics. Nowhere is this more the case than in the study of Milky Way dynamics Bovy (2020b)

The history and present evolutionary state of the Milky Way are currently being deciphered at all spatial scales, boosted by the remarkable Gaia measurements Ivanova et al. (2021)

The recent Gaia EDR3 opens a new chapter in the measurement of parallaxes, placing the precise and accurate determination of the distances to nearby Galactic globular clusters within reach Soltis et al. (2021)

Gaia has revolutionised Galactic dynamics by providing positions, parallaxes and proper motions with unparalleled precision for a large number of Milky Way stars Angus et al. (2022)

Gaia has revolutionised the study of Galactic archaeology by measuring the stellar kinematics of stars in the solar neighbourhood with an unprecedented precision Ghosh et al. (2022b)

Gaia's unique capability in stellar measuring proper motion and vast dataset is revolutionising our understanding of Milky Way structure Buckley et al. (2023)

Since the first data release, Gaia has played a major part in revealing the kinematics of our Galaxy Wang et al. (2023b)

Gaia has provided revolutionary observational data that have uncovered detailed kinematical features of stars in the Milky Way Funakoshi et al. (2024)

The Gaia mission has revolutionised our view of the Milky Way and its satellite citizens… It has changed our view on how the Milky Way assembled, it has changed how we model the Galaxy, and it has opened our eyes to new research directions and explorations Deason & Belokurov (2024)

The Gaia Data Releases, DR1 and DR2, changed our entire perspective of visualising the Galaxy and its neighbourhood Dhanush et al. (2024)

In the age of Gaia, the local stellar velocity distribution is now known with exquisite precision
Hopkins et al. (2025)

Stellar streams

Gaia has opened up a spectacular vista onto the Milky Way's halo, revealing a lacework of criss-crossing ancient and ongoing accretions that testify to the violent formation history of our Galactic home Ibata et al. (2021)

The second data release of Gaia expanded our ability to make [6-dimensional phase-space] maps by several orders of magnitude, by measuring proper motions for more than a billion Milky Way stars. This phenomenal wealth of data has already facilitated the discovery of many new streams Reino et al. (2021)

Our view of the variety of stellar structures pervading the local Milky Way has been transformed by the application of clustering algorithms to the Gaia catalogue
Andrews et al. (2022)

The high-precision astrometric survey Gaia has revolutionised [the field of stellar streams] and allowed for measurements of the proper motion of faint stream stars for the first time Ferguson et al. (2022)

The revolution brought about by the deep astrometric (and photometric) survey of the Gaia satellite is truly staggering. With each new data release, the Milky Way's phase space is revealed in a wealth of ever more complex and intricate detail Martin et al. (2022)

This analysis [of the proto-Milky Way] reflects the astounding information content of the Gaia Data Release 3, particularly the XP spectra Rix et al. (2022)

The field of stellar streams is currently in a golden era. It has increasingly grown and all but exploded in the last decade, thanks to deep wide-area photometric surveys and, more recently, to the amazing possibilities opened by the all-sky astrometric information provided by the Gaia mission since its Second Data Release Mateu (2023)

The Gaia mission has brought fundamental changes to our view of stellar streams in the Milky Way
Bonaca & Price-Whelan (2025)

The identification by Helmi et al. (1999) of a handful of stars belonging to a stellar stream passing through the solar neighbourhood is a good example of the type of archaeological experiment that had appeared trailblazing and challenging before Gaia, but has been made effortless and almost mundane thanks to Gaia today
Deason & Belokurov (2024)

Solar system

The results from ESA's Gaia mission, an unparalleled catalogue of stellar position, distances, and kinematic states that have, in turn, unlocked the door to widespread studies of the outer solar system M. Buie (2022)

The precision of the Gaia catalogues has marked the beginning of a new era in predicting stellar occultation
Souami et al. (2022)

There are so many revolutionary advances that it is difficult to pinpoint a single most significant advance
Karri Muinonen, Space.com, 28 July 2022

The release of Gaia catalogue is revolutionary to the astronomy of solar system objects Guo et al. (2023)

Exoplanets

Thanks to Gaia, uncertainties in stellar radii are no longer the limiting uncertainties in determining planetary radii for the majority of Kepler planets
Petigura (2020)

Gaia is revolutionary for almost every field in astrophysics, and exoplanets is no exception
Sanderson et al. (2022)

The Gaia space mission is impacting astronomy in many significant ways… Exoplanet science has greatly benefited from the unprecedented accuracy of the stellar parameters obtained from Gaia Swastik et al. (2023)

Telescope operations

Many of our new observatory operational and scientific developments have directly benefited from Gaia data releases in ways that were simply not possible before. Its high-precision astrometric and spectroscopic data has been transformative for both our operations and our science.
Stephen Potter, SAAO (priv. comm., December 2024)

Gaia has been transformative in various ways. It has provided, for the first time, a reliable and dense set of guide stars for wide-field spectroscopy across the entire sky, especially for the southern hemisphere. Before Gaia, many many hours were spent cobbling together various sources of guide stars, ensuring they were on the same system, estimating magnitudes, etc. This instantly became a completely solved problem when Gaia data was available
Mike Blanton, SDSS (priv. comm., December 2024)