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 was used in the detection of the Sagittarius stream (Ibata et al., 2002).

• 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., 2020), 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 data (Malhan et al., 2018). In Gaia DR2, Ibata et al. (2018) found Phlegethon, while Ibata et al. (2019) 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.

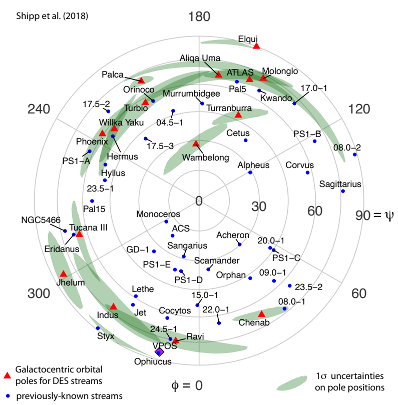
This is not an exhaustive listing, and a number of other stream-searching algorithms can be found in the literature, amongst them another recent model-agnostic algorithm, `SkyCURTAINs` (Sengupta et al., 2024).

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.



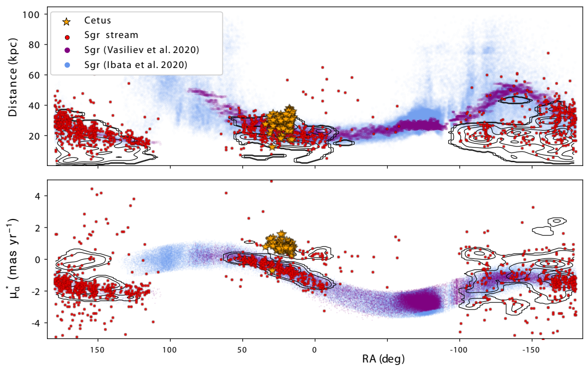
Chang et al. (2020) also predicted that about half of the stream members would be in the southern sky. And their predicted location overlapped the diffuse Palca stream, recently discovered from the Dark Energy Survey, and at a similar distance of ~36 kpc (Shipp et al., 2018). The rich structure of these 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.



Thomas & Battaglia (2024)

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'*.