2025 Sagan Summer Workshop

Exoplanet Demographics

Hands-on Session II: Distant Giant Planets with Astrometry

Group Projects

Kyle Franson (UT Austin), Rocio Kiman (Caltech), Tim Brandt (STScI)

Group Project 1

For this group project, your goal is to identify the best targets for imaging exoplanets or brown dwarf companions by extending the framework outlined in the tutorial notebook. Start by deciding whether you want to image young exoplanets from the ground or older exoplanets from space with JWST. For either direction, tinker with the parameters in the notebook until you have a substantial list of promising accelerating stars. Then, whittle down the list target-by-target. Some things to consider for each target as you work down the list:

(1) Have the targets been previously observed with exoplanet imaging instruments, like GPI, SPHERE, or NIRC2? You can check this by querying the Gemini (for GPI), ESO (for SPHERE), and Keck (for NIRC2) archives. If a target was just observed with a long sequence, it may not be the best target to try to image a new exoplanet from the ground. Note, though, that planets can move around in their orbits. If it's been a few years since a target was observed, or if the sequence is short, it may be worth reobserving! For imaging older exoplanets with JWST, you can check whether stars have been observed or are on the schedule for the upcoming year with the JWST Duplicate Observations tool.

(2) In addition to checking against Gaia parameters in the notebook, we want to be sure that there isn't evidence of visual binaries that could be causing these accelerations. For each target, you can quickly check whether there are any nearby comoving sources in Gaia by querying this table on Vizier. The default search radius of 2 arcmin is pretty large and Gaia goes down to G ~ 21 mag, so you're gonna get a ton of background stars. You can check if there are any nearby comoving stars by seeing if there are two stars with similar proper motions in RA and Dec (pmRA, pmDE) and parallaxes (Plx). The table for a star with a resolved binary in Gaia will look something the following:

I/355/gaiadr3 — *Post annotation*

Gaia DR3 Part 1. Main source (Gaia Collaboration, 2022) — 2022yCat.135

Gaia data release 3 (Gaia DR3). (original column names in green) *(1811709771 rows)*

start AladinLite   plot the output   query using TAP/SQL

Full	RA_ICRS deg	DE_ICRS deg	Source	e_RA_ICRS mas	e_DE_ICRS mas	Plx mas	e_ mas	PM mas/yr	pmRA mas/yr	e_ (...)	pmDE mas/yr	e_ (...)	RUWE
1	079.69522832905	-21.42155340014	2962192459620864128	1.4625	1.6718								
2	079.72496838240	-21.40061818707	2962192528344018944	0.0167	0.0192	0.4571	0.0296	5.442	2.076	0.021	5.031	0.025	0.970
3	079.70608268654	-21.39655370849	2962192558404939904	0.0896	0.1103	0.2776	0.1644	7.988	7.549	0.114	2.612	0.147	0.982
4	079.70826559010	-21.40364267321	2962192562703759616	0.0063	0.0079	1.8541	0.0114	6.937	5.333	0.008	-4.437	0.010	0.997
5	079.71270996082	-21.39670422768	2962192597063496576	0.0488	0.0585	0.4797	0.0891	1.089	-0.062	0.062	-1.087	0.075	1.018
6	079.72108345466	-21.38538863649	2962192970721794176	0.0419	0.0497	0.5495	0.0755	5.506	4.878	0.053	2.553	0.062	0.997
7	079.67465205382	-21.39807777995	2962193279959430272	0.0433	0.0509	0.6208	0.0785	3.824	3.406	0.056	1.738	0.063	0.964
8	079.67091968590	-21.38518657775	2962193284256468352	1.0136	1.1442	-0.9779	1.7974	7.023	6.204	1.046	-3.291	1.432	1.154
9	079.69228700748	-21.39500185286	2962193314319773056	0.1254	0.1499	2.3203	0.2180	5.371	-4.304	0.154	-3.212	0.180	1.009
10	079.69521121956	-21.39463819065	2962193318614307200	0.1276	0.1207	1.5864	0.1950	4.101	4.029	0.175	-0.766	0.171	1.417
11	079.69935356632	-21.38788783735	2962193318616049920	0.4613	0.5331	0.6864	0.8869	4.367	4.205	0.522	1.178	0.672	1.036
12	079.69741566124	-21.39366198032	2962193318618003584	0.0188	0.0255	49.0080	0.0365	145.508	-138.798	0.025	-43.676	0.031	1.566
13	079.69597282443	-21.39392110555	2962193318618221056	0.0074	0.0097	48.9655	0.0142	141.930	-137.035	0.010	-36.954	0.012	1.054 5
14	079.68110052014	-21.38356314214	2962193383038656256	0.2466	0.3108	0.6678	0.4694	3.177	2.744	0.312	-1.601	0.386	1.016
15	079.71821396321	-21.36895890588	2962193730935069312	0.0323	0.0355	0.5140	0.0565	9.653	4.455	0.040	8.563	0.046	1.078
16	079.70642105103	-21.36505228636	2962193795355529856	0.1535	0.1866	0.4057	0.2762	2.089	-0.716	0.198	1.962	0.248	0.977
17	079.70294209348	-21.36279027303	2962193795355531264	0.4876	0.5562	1.5366	0.8582	23.613	-0.797	0.650	-23.599	0.741	4.717
18	079.70287166277	-21.36263308612	2962193799650617600	1.0524	0.5899	3.0006	0.6143	26.540	-0.679	0.873	-26.532	0.713	1.054

Target retrieved from simbad (79.696627, -21.393757)

plot the output   query using TAP/SQL

Here, rows 12 and 13 are the binary components.

(3) Another catalog to check against for visual binaries is the Washington Double Star catalog. This is a compendium of binaries and optical pairs maintained by the Naval Observatory. Many of the pairs in the catalog are not physical binaries; rather, they're nearby sources. Comoving pairs or pairs with orbital motion have the flags "T," "V," or "O." If you find a target has an entry within WDS with one of those flags and an angular separation <10 arcsec, you likely want to avoid that target. For binaries with wider separations, you can get a sense for whether it could be consistent with causing the acceleration using the mass_causing_acceleration function in the notebook. Note that the function returns mass in M_{Jup}. You can convert this to solar masses by dividing it by 1047.57.

(4) Finally, you should check whether these targets are known spectroscopic binaries. Spectroscopic binaries can masquerade as astrometric exoplanets by causing photocenter shifts that can mimic the astrometric orbits of exoplanets. This might be a significant source of false-positives among the Gaia DR4 exoplanets (see Stefánsson et al. 2025). It's best to avoid known spectroscopic binaries here as well. To check this, you can query the SB9 catalog.

(5) A good final check is to search the internet for any additional information on each target, such as a known planet or binary. You can search for previous papers that reference a star on NASA/ADS using the object field (e.g., object: "HIP 27321").

Any stars that pass these cuts will be some of the most promising astrometry-informed targets for imaging exoplanets. How would you convince the community (or a Time Allocation Committee) that these are *the stars* around which we should look for planets ? Is there additional information or other science cases that could further bolster the case for observing your targets? Pick your 3–5 favorite targets and write up a short case for observing them.

Group Project 2

This activity has the goal of studying the sensitivity of different Gaia parameters to identify binaries. For this we are going to use the Washington Double Star (WDS) catalog (https://www.astro.gsu.edu/wds/), which has identified binaries with their separation. We will cross-match this catalog with Gaia DR3 to identify the unresolved binaries in Gaia, and study different parameters which are used to identify binaries as a function of the separation.

You can choose one of two options:

> A) Follow the instructions below in Part 1 to put together the sample of binaries, and then go to Part 2 to make plots with the sample. Note that you will need to create an account in the Gaia ESA Archive as a step in Part 1 (instructions are provided in step 3a).
>
> B) Use the file with the sample we prepared (WDS_unresolved_original.csv; it can be found in the SSW2025/ImagingGaiaPlanets directory in your Google Drive), and go directly to Part 2 to make the plots.

Part 1:

For this part you can use the Colab notebook we put together which has guidelines for each step (see SSW2025_ImagingGaiaPlanets_group_project_02.ipynb; you can find the link in the Google Colab instructions document), or you can start your own.

1) Using the object `Vizier` from the package `astroquery.vizier`, download the Washington Double Star Catalog catalog (WDS, https://www.astro.gsu.edu/wds/), with the name *'B/wds/wds'*. For this step you can use the same technique we used for the main tutorial, in the sample notebook (SSW2025_ImagingGaiaPlanets_sample.ipynb). In order to download the full catalog, add explicitly all the columns we need:

```
columns=['Comp','RAJ2000','DEJ2000','pa2','sep2','mag1','mag2','pmRA1','pmDE1','pmRA2','pmDE2']
```

Note: you might need to replace the elements which are '<NA>' with `np.nan`.

2) Separate the coordinates of each component of the binary using astropy (https://docs.astropy.org/en/stable/coordinates/index.html). For this step you will need to use `SkyCoord` and `directional_offset_by` to obtain the position of the secondary from the primary using the position angle and the separation. Then create a table including the primary and the secondary in two consecutive rows. Given that we are going to crossmatch this sample with Gaia, select the stars which have total proper motion < 62.5 mas/yr, meaning that in 16 years (difference in the epoch for the positions of WDS and Gaia DR3) the stars moved less than 1 arcsecond. This will reduce the amount of mismatches. The complete way of solving this problem would be to use the proper motion to move the positions of the stars forward in time, so the epochs agree, but we are not going to concentrate on that in this exercise. Also make sure to select objects which do not have a note in the

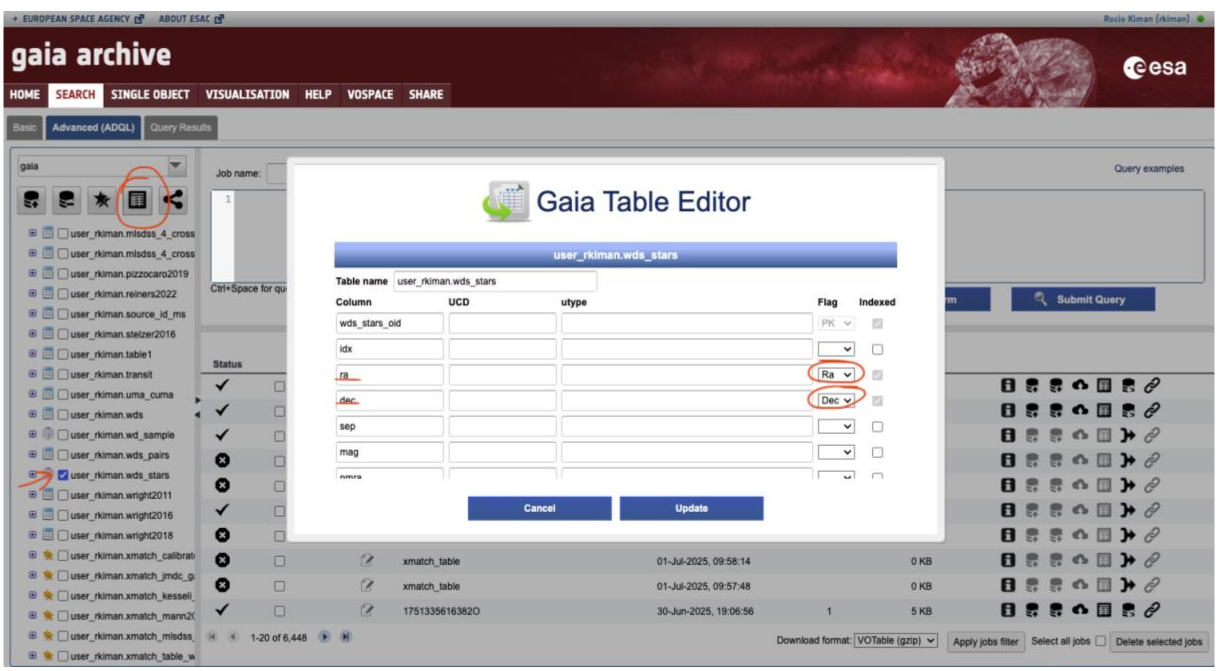
column 'Comp'. Note: You might want to add an index to mark primary and secondary with the same number so you can identify the pairs later.

3) Cross-match with Gaia each component of the binaries with a 3 arcsec radius. We will split this task given that it requires several steps:

a) Create an account in the Gaia ESA Archive if you don't have one already. Go to https://www.cosmos.esa.int/web/gaia-users/register to register for an account. Registration is free and quick, and it offers features such as space to store query results, the option to create and share your own tables, and longer timeout for large or complex query jobs.

b) In your notebook add `Gaia.login()`, where you can import Gaia using `from astroquery.gaia import Gaia`. When you run the cell you will have to use your user name and password to login.

c) Now use the `Gaia` object to make the crossmatch with Gaia DR3 for your table with the pairs, using a 3 arcsec radius. You might find this website useful https://astroquery.readthedocs.io/en/latest/gaia/gaia.html. First use the function `Gaia.upload_table()`, to upload the pairs table to the Gaia archive. After you run this code, go to the Gaia archive https://gea.esac.esa.int/archive/, login with your account, and go to Search -> Advance (ADQL) -> Scroll down to User tables on the menu to the left and look for your table which is called user_"your user"_"name of the table". Select your table and click on the bottom for the Table editor (red circle in the image below). Select the ra and dec columns, indicate they are ra and dec in the Flag column, and update the table. This will make the table ready for crossmatching.



Then follow the website https://astroquery.readthedocs.io/en/latest/gaia/gaia.html and use the functions Gaia.cross_match() and Gaia.launch_job_async(query=query) to perform the crossmatch. An example of a query can be found in the Colab notebook we shared at the beginning if you need help with it. Make sure to extract all the information from Gaia DR3.

4) Select the binaries for which both components have the same match. In your results you will notice that some of the pairs (stars with the same index) have the same source_id, and some have a different ID. We are going to identify pairs that have the same source id as unresolved binaries. Select only those stars, and note that you should keep only one of the stars.

Part 2:

For both cases, whether you just went through the process of making the sample or whether you are using the file with the sample we prepared, now you have a list of stars that we know have a binary companion but are unresolved in Gaia, we have the separation thanks to the WDS, and we have the information from the Gaia DR3 catalog.

1) Define binaries according to the Gaia cuts using ruwe and ipd_frac_multi_peak (and any other you want).

2) Study the fraction of binaries that are detected with the cut as a function of separation on the known binary from the WDS catalog.

3) Analyze your results. Answer questions such as:

- How sensitive is each cut?
- At which separation do the binary flags stop being sensitive?
- Which cases would require each of the types of cuts?
- Is there a criterion that is better than the others to identify binaries?

Group Project 3 (note: this is a more challenging group project)

For this group project, you will derive the mass of a companion from an astrometric or radial velocity trend.

Follow the instructions on the following two pages.

A Mass from an Astrometric or Radial Velocity Trend

Long-period companions to main sequence stars can induce accelerations. These accelerations can be astrometric, in the plane of the sky, and RV, along the line-of-sight. RV accelerations are also known as trends. Denoting the mass of the companion tugging a star as M_B, the RV and astrometric accelerations may be derived from Newton's Law of Gravity as

$$a_{\mathrm{RV}} = \left|\frac{d\mathrm{RV}}{dt}\right| = \frac{GM_B}{r^2}\cos\varphi \tag{1}$$

and

$$a_{\mathrm{astrometric}} = \frac{GM_B}{r^2}\sin\varphi. \tag{2}$$

where φ is the angle between the line-of-sight and the separation vector $\mathbf{r}$ between the star and its companion. Equation (2) assumes that only the acceleration of the main star is measured, i.e., that there is no light from the companion contaminating the measurement.

We might want to relate the astrometric or RV trend to the projected separation if we are looking to image a companion. The projected separation, the product of the distance D from Earth to the star and the angular separation ρ, is

$$D\rho = r\sin\varphi. \tag{3}$$

Very roughly speaking, you could adopt some value for φ and get an approximate relationship between M_B and $D\rho$. We did something like this in the earlier part of the hands-on activities when we multiplied the astrometric acceleration by 1.5 to get an estimate for the total acceleration. In this activity you will derive the actual relevant distribution for the relationship between mass and separation, both for RV and for astrometric acceleration. The astrometric case is a bit easier to work through, so we will do that one first. The RV case is a bit harder since you need to invert a multi-valued function, but it's still tractable.

1. Combine Equations (2) and (3) to get M_B as a function of astrometric acceleration, projected separation $D\rho$, and φ. Write your expression in the form

$$M_B = A a_{\mathrm{astrometric}}(D\rho)^2\Phi(\varphi) \tag{4}$$

where A is a constant and $\Phi(\varphi)$ is some function of φ.

Solution:

$$M_B = \frac{1}{G}(D\rho)^2 a_{\mathrm{astrometric}}\Phi(\varphi) \tag{5}$$

with

$$\Phi = \frac{1}{\sin^3\varphi}. \tag{6}$$

If you assume that the orientations in space are random, the probability distribution of φ is given by

$$\frac{dp}{d\varphi} = \sin\varphi \tag{7}$$

for $\varphi \in [0, \pi/2)$ (it's the same between $\pi/2$ and π). If you prefer, you can use the equavalent

$$\frac{dp}{d\cos\varphi} = 1. \tag{8}$$

Use Monte Carlo to generate a large number of φ values and compute a histogram of Φ. For the histogram, use `np.hist` with `log=True` and a large number of bins.

Solution: see notebook

Next, use the equations to derive the probability distribution $dp/d\Phi$ of $\Phi(\varphi)$ and plot it over your empirical histogram. You will want to use the chain rule, e.g.,

$$\frac{dp}{d\Phi} = \frac{dp}{d\cos\varphi}\left(\frac{d\Phi}{d\cos\varphi}\right)^{-1}. \tag{9}$$

Solution:

$$\frac{d\Phi}{d\varphi} = \frac{-3\cos\varphi}{\sin^4\varphi} \tag{10}$$

so

$$\begin{aligned}
\frac{dp}{d\varphi}\left(\frac{d\Phi}{d\varphi}\right)^{-1} &= \frac{-\sin^5\varphi}{3\cos\varphi} \\
&= -\frac{\Phi^{-5/3}}{3\cos\phi} \\
&= -\frac{\Phi^{-5/3}}{3\sqrt{1-\Phi^{-2/3}}} \\
&= \frac{-1}{3\sqrt{\Phi^{10/3}-\Phi^{8/3}}}.
\end{aligned} \tag{11}$$

We would take an absolute value for the probability density.

Finally, compute the median of your distribution, either using the Monte Carlo or the analytic result. This median could be used in place of the 1.5 adopted in the earlier tutorial notebooks.

Solution: the median is about 1.54 (it's actually $(4/3)^{3/2}$).

2. Combine Equations (1) and (3) to get M_B as a function of RV trend, projected separation $D\rho$, and φ. As for the astrometric case, write your expression in the form

$$M_B = A a_{\rm RV} (D\rho)^2 \Phi(\varphi) \tag{12}$$

where A is a constant and $\Phi(\varphi)$ is some function of φ. As before, generate a large number of φ values and make a histogram of Φ using Monte Carlo.

Solution:

$$M_B = \frac{1}{G}(D\rho)^2 a_{\rm RV} \Phi(\varphi) \tag{13}$$

with

$$\begin{aligned}\Phi &= \frac{1}{\cos\varphi \sin^2\varphi} \\ &= \frac{1}{\cos\varphi - \cos^3\varphi}\end{aligned} \tag{14}$$

See the notebook for the histogram.

Then derive the probability distribution $dp/d\Phi$ of $\Phi(\varphi)$ and plot it. This time, you may find that getting $d\Phi/d\cos\varphi$ in terms of Φ is hard, because the same value of Φ can arise from multiple values of $\cos\varphi$. You can treat the different solutions separately and combine the results. You might want to use Mathematica or Wolfram Alpha or similar to get these inverse functions.

Solution:

$$\begin{aligned}\frac{d\Phi}{d\cos\varphi} &= \frac{3\cos^2\varphi - 1}{(\cos\varphi - \cos^3\varphi)^2} \\ &= \Phi^2 \left(3\cos^2\varphi - 1\right).\end{aligned} \tag{15}$$

This time, I need to get $\cos\varphi$ from Φ. Unfortunately, the function is multiply valued. The two (real) roots that apply between $\varphi = 0$ and $\varphi = \pi/2$ are

$$\cos\varphi_1 = \frac{2}{\sqrt{3}} \cos\psi \tag{16}$$

and

$$\cos\varphi_2 = \sin\psi - \frac{1}{\sqrt{3}} \cos\psi \tag{17}$$

with

$$\psi = \frac{1}{3} \cos^{-1}\left(-\frac{\sqrt{27}}{2\Phi_{\rm RV}}\right). \tag{18}$$

Each of the equations for φ_1 and φ_2 applies over a different range of φ values. When I compute a probability density via the chain rule, I end up dividing each probability density by the size of the domain of φ values to normalize, but then multiply by the same domain size when adding the two probability densities together. I finally have

$$\begin{aligned}\left|\frac{dp}{d\cos\varphi}\left(\frac{d\Phi}{d\cos\varphi}\right)^{-1}\right| &= \left|\left(\Phi^2\left(3\cos^2\varphi - 1\right)\right)^{-1}\right| \\ &= \left|\left(\Phi^2\left(3\cos^2\varphi_1 - 1\right)\right)^{-1}\right| + \left|\left(\Phi^2\left(3\cos^2\varphi_2 - 1\right)\right)^{-1}\right|.\end{aligned} \tag{19}$$

This is plotted in the notebook.