

# Offre de stage de Master 2

## Développement d'une nouvelle méthode de clusterisation de familles d'astéroïdes

**Durée :** 4 à 6 mois à compter de février 2023

**Profil recherché :** Master 2 ou équivalent en astronomie ou statistique/machine learning

**Contacts :** Benoit Carry (benoit.carry@oca.eu) et Vincent Vandewalle (vincent.vandewalle@inria.fr)

**Lieu du stage :** Observatoire de la Côte d'Azur ou Centre Inria d'Université Côte d'Azur

**Gratification :** Environ 600 euros par mois.

### Contexte

Les astéroïdes entre Mars et Jupiter sont les vestiges des planétésimaux qui se sont accrétés pour former les planètes. Leurs distributions orbitale, en taille, et compositionnelle portent la trace des différentes étapes de la formation de notre système solaire, de l'accrétion des solides dans le disque aux migrations planétaires.

Néanmoins, 4 milliard d'années d'évolution dynamique et collisionnelle ont imposé leur marque, brouillant ces informations. En particulier, les astéroïdes se détruisent mutuellement lors de collision catastrophique, éjectant de nombreux fragments. On appelle les groupes formés par ces fragments les familles d'astéroïdes. Identifier correctement ces familles est critique pour remonter le temps et déterminer la distribution originelle des astéroïdes dans la ceinture principale.

Deux méthodes sont actuellement utilisées dans la communauté pour classer les astéroïdes au sein des familles. Plus de 200 familles ont été identifiées, par ces méthodes, permettant d'étudier de nombreux aspects de l'évolution des astéroïdes. Toutefois, ces deux méthodes ont plusieurs limites. Le but de ce stage, en co-direction entre l'Observatoire de la Côte d'Azur et l'Inria, est d'explorer des méthodes modernes de clustering pour identifier les familles d'astéroïdes.

### Déroulement

Dans un premier temps, le ou la stagiaire proposera un modèle probabiliste sur la distribution des mesures (excentricité, demi-grand axe, etc.) à l'intérieur d'une famille d'astéroïdes en s'appuyant sur des connaissances physiques liées au phénomène de dispersion des astéroïdes. Ce modèle servira de brique élémentaire pour mettre en place une approche de clustering à base de modèles [1].

Le ou la stagiaire proposera ensuite un algorithme d'estimation des paramètres et réalisera une première implémentation sous R ou Python avec premiers tests sur données réelles. Le modèle développé pourra être étendu pour prendre en compte des contraintes physiques supplémentaires conduisant par exemple à l'observation de distributions tronquées.

Enfin, une fois le modèle stabilisé, les nouvelles familles obtenues seront étudiées en détails et comparées aux familles identifiées par la communauté.

### Références

- [1] Charles Bouveyron, Gilles Celeux, T Brendan Murphy, and Adrian E Raftery. *Model-based clustering and classification for data science : with applications in R*, volume 50. Cambridge University Press, 2019.