CONSTRAINING PLANET LOCATION THROUGH GRAVITATIONAL MODELING

By William J. Oldroyd

A Dissertation
Submitted in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy
in Astronomy and Planetary Science

Northern Arizona University

August 2022

Approved:
Chadwick A. Trujillo, Ph.D., Chair
Tyler D. Robinson, Ph.D.
Cristina A. Thomas, Ph.D.
William M. Grundy, Ph.D.

ABSTRACT

CONSTRAINING PLANET LOCATION THROUGH GRAVITATIONAL MODELING

WILLIAM J. OLDROYD

Throughout history, tracking the location of planets has been at the forefront of astronomy. Modern computational resources allow for extreme precision in tracking planets and other objects through their orbits with the use of gravitational modeling. The classical planets in the solar system have well-known orbits, but there are several types of planets that require further orbital analysis in order to develop a solid understanding of these worlds and their impact on the solar system and beyond.

In this dissertation we discuss our efforts to gravitationally model three separate systems. The first of these is the relationship between a hypothetical distant giant planet in the outer reaches of our solar system and the distant objects that provide evidence for its existence. This population of distant dwarf planets has orbits that are clustered together in a way that is suggestive of another large planet, which we refer to in this work as Planet X, orbiting beyond Neptune. Here we explore another orbital feature related to these objects, a gap between two sub-populations of distant dwarf planets that cannot form under the gravitational influence of the known planets alone. By including Planet X in the solar system, we show that both the gap and this population of distant objects are formed naturally over the age of the solar system.

The second group of systems analyzed in this work centers on exoplanets, planets orbiting other stars. One major challenge for direct imaging studies of exoplanets is the large amount of telescope observation time that must be allocated to orbit determination. We have developed a method for reducing the telescope observing time required to determine the orbits of these exoplanets by reducing the number of revisit observations required to constrain directly imaged exoplanet orbits. This will allow for a sizeable fraction of observing time to be repurposed for further study of the surfaces and atmospheres of these worlds.

Our final study is focused on a minor planet in the solar system, 282P. This object shows signs of comet-like activity and it clearly has close encounters with Jupiter both in the last few hundred years, and in the next few hundred as well. These close approaches cause the orbit of 282P to be chaotic beyond the time of these encounters, so we employ statistical techniques to determine likely outcomes and histories of 282P. We find that 282P is in the Quasi-Hilda region, which likely serves as an intermediate zone between comets and active asteroids.

These related projects all focus on constraining the orbital parameters of the planet in question in order to better understand the system as a whole. By improving our understanding of the gravitational influence exerted by and on these bodies, we can develop a more complete picture of the formation, composition, and evolution of the solar system and other planetary systems as well.

Copyright

Chapter 2 contains a published manuscript.

> *This is the Accepted Manuscript version of an article accepted for publication in the Astronomical Journal. IOP Publishing Ltd is not responsible for any errors or omissions in this version of the manuscript or any version derived from it. The Version of Record is available online at* https://iopscience.iop.org/article/10.3847/1538-3881/abfb6f.

Chapter 3 contains work that will shortly be submitted for publication, thus, the content is presented in manuscript format. Chapter 4 is derived from work that has been submitted to Astrophysical Journal Letters for which I am second author.

Acknowledgements

I am grateful for all of the help and support I have received throughout my education from family, friends, and teachers. First, I would like to thank my wonderful wife Jamie for loving me through all of our phases and for providing strength and support when I need it most. Thank you to my children Garrett and Claire for always brightening my day (and for moving my desk up and down as often as you can get away with it). Thank you to the rest of my family as well for encouraging me and listening to hours of facts and details about the universe. I am particularly grateful to Dad, Pops, and Grandpa Paulson for instilling in me a love of space and discovery from a young age, for Mom, Grandma Paulson, and Grandma Gina for helping me figure out what I'm supposed to do, and to Tom and Shelly Claridge for their love and trust.

I would like to thank my committee, Chad Trujillo, Ty Robinson, Cristina Thomas, and Will Grundy, for their excellent guidance throughout my graduate studies as well as for their confidence in me. I am also thankful for the support of the many wonderful teachers and friends I have worked with and interacted with during my time at NAU. A big thank you to Colin Chandler with whom I have spent many hours observing, writing, and repairing telescopes, and to Erin Aadland and Stefan Nelson for helping Jamie and I save the world. I would also like to thank my students, especially Lucas McClure who taught me more than I taught him.

Through the course of this journey there have been hundreds of exceptional individuals who have mentored me and who have helped me to learn and grow. Thank you to Darin Ragozzine, Jani Radebaugh, Nathan Powers, Eric Hintz, J Moody, Denise Stephens, Clark Snelgrove, Valerie Hegstrom, Dale Pratt, Doug Corey, Steve Call, JB Haws, and many others who mentored me during my time at BYU. Thank you to David Fulmer, Alan Myrup, Dave Shelton, Alan Allred, Paul Larson, Dave Creer, Alison Van Orden, and many others who helped teach me to love science, study, and practice. I would also like to thank Matt Clement and Juliette Becker for helping me find my footing as an astronomer and for creating opportunities for me to contribute to the astronomy community. I am grateful to Nozomu Okuda for his continued friendship and example of following your dreams. Additionally, I am thankful to Timothy and Judith Jones and Raymond and Laurie Cutler and the many others I served with for their mentorship during my time as a missionary.

Most of all, I am grateful to the creator of the planets, and everything else, who has given me so many amazing opportunities to learn and grow.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Piazzi, Gauss, and Ceres

On 1 January 1801, Italian astronomer Giuseppe Piazzi was hard at work making a catalog of stars when he discovered a "new star" in the constellation of Taurus (Piazzi, 1801). As he mapped out the sky, he would painstakingly remeasure each star for several nights in a row in order to make sure the coordinates of each star were correct. To his surprise, he found that his new "star" would move each night and he quickly announced his discovery of a new star that might even be a comet (Foderà Serio et al., 2002).

After two weeks of observations, it occurred to Piazzi that his discovery might be more significant than a new star or comet. The presence of a planet orbiting between Mars and Jupiter had long been predicted (Kepler et al., 1596) and was still being actively searched for at the time (Bode, 1772; von Zach, 1801). Piazzi, believing he may have found this planet, called it "Ceres Ferdinandea" after both the patron goddess and the king of Sicily (Piazzi, 1801) and wrote to a few of his colleagues on the matter.

While waiting for a reply, Piazzi fell ill. Between his recovery, other responsibilities as an astronomer at Palermo Observatory, and not wanting to let someone else be the first to calculate an orbit for the new planet, Piazzi did not publish his observations until September 1801 (Piazzi & von Zach, 1801). These observations seemed to confirm the discovery of a new planet, however, because of the delay in their publication, it was no longer the right season to observe the new planet and would not be for several months[1]. Additionally, the calculated orbit for the new planet was inaccurate because of the small number of observations Piazzi had made while the planet was still observable. This inaccuracy would make observing the planet again almost

[1] As the Earth orbits the Sun, different parts of the night sky are visible in different seasons. This is why we have Summer and Winter constellations.

impossible, and, when no one was able to locate the new planet when it should have been observable again, several prominent astronomers began to doubt the existence of the new planet.

During this time of uncertainty, a 24 year old mathematician named Carl Friedrich Gauss saw this as a singularly important mathematical challenge with enormous ramifications. Building off of the work of Kepler (1619) and Newton (1687), Gauss ingeniously developed, in less than three months, a new method for computing orbits that laid the foundation for modern predictive astronomy. Gauss used this method to calculate the location of the planet and it was definitively recovered on 31 Dec 1801, almost exactly one year after its initial discovery, less than half a degree from the position he predicted (see Gauss, 1809).

Upon its recovery, the name of the new planet was shortened to Ceres and was lauded as the most important astronomical finding since the discovery of Uranus in 1781. Within the next six years, however, three additional "planets" were discovered in the region between Mars and Jupiter (the same region as Ceres). During this time William Herschel, the discoverer of Uranus, noted that these bodies were much smaller than the planets and referred to them as "asteroids," (Herschel, 1802) a designation that would start to stick roughly 100 years later (see Section 1.3.2.3).

The story of Piazzi, Gauss, and Ceres illustrates the importance of planetary orbit determination based on only a small amount of observational data (see Foderà Serio et al. (2002) for a detailed account of the discovery of Ceres). Calculating accurate orbits and positions of astronomical objects remains at the forefront of astronomy for its applications to exoplanets (Blunt et al., 2020), binary (Grundy et al., 2019) and satellite-primary interactions (Oldroyd et al., 2018), Near-Earth Asteroids (de la Fuente Marcos & de la Fuente Marcos, 2019), spacecraft targets (Thomas et al., 2018), newly discovered outer solar system (Sheppard et al., 2019) and interstellar objects (Bailer-Jones et al., 2020), large observational surveys (Juric et al., 2020), and a host of other sub-fields in astronomy. In fact, methods used today for calculating orbits are based on the techniques pioneered by Gauss in his search for Ceres.

Now, more than 200 years after its discovery, Ceres remains one of the most interesting objects in our solar system. Currently classified as both an asteroid and a dwarf Planet (see Section 1.3.2.6), Ceres was orbited by the NASA Dawn spacecraft which found it to be a fascinating icy world with cryovolcanoes (that erupt ice instead of lava) dotting its surface (Sori et al., 2018). Recent studies also suggest that Ceres may have originated in the outer solar system and migrated to its current location, where it comprises roughly 1/3 of the mass in the asteroid belt (Ribeiro de Sousa et al., 2022). Regardless of its origin, Ceres remains at the forefront of planetary science and solar system astronomy[2] and its history and future will continue to inspire countless people to explore and discover the universe around them.

[2] It's also well loved in science fiction.

1.2 Orbital Dynamics

1.2.1 A Brief History

Orbital dynamics is the sub-field of astronomy focused on the gravitationally-driven motions of objects in space. This includes stars, planets, galaxy clusters, dust, and everything in-between. Although a full treatment of the topic is beyond the scope of this work, we will touch on a few key advancements throughout the history of orbital dynamics (sometimes called celestial mechanics).

In many ancient cultures the positions of the Sun, Moon, and planets served as astrological indicators for important life events. The planets were also often associated with creation stories in many ancient cultures. For example, *Popol Vuh, the Sacred Book of the Maya* personifies the planets as various deities from Mayan mythology and gives them roles in the creation of the world (Christenson, 2003). Similarly, the ancient Greeks and Romans associated the planets with the gods of their pantheon and, as our modern understanding of astronomy has evolved out of Greco-Roman philosophy, we continue to use the Roman names for the planets today.

Because of the importance placed on the location of the planets by the ancients, many civilizations developed methods for predicting when each planet would be visible, when they would come to a conjunction and align in the sky, and when eclipses would occur. They often built buildings and even entire cities aligned with solstices and equinoxes to help them precisely determine the seasons and measure planetary motion (e.g., Jenkins, 2002). These predictive techniques laid the foundation for the bridge between ancient astronomy and the early modern astronomy of the 17th century.

The oldest surviving treatise on astronomy is the *Almagest* by Ptolemy written sometime in the 2nd century AD. In this definitive work, Ptolemy provides predictions for planet positions and eclipses, based mainly on the methods of the Egyptians and Babylonians. He also explains the prevailing theory of the heavens as outlined by Aristotle followed by the magnitude system for quantifying the brightness of stars developed by Hipparchus, which we still use today[3]. The defining trait of the *Almagest*, however, is that it outlines a geocentric model for the universe (where the Earth is the center of everything, see Figure 1.1) and invokes complicated geometric orbit shapes called epicycles in order to force the Sun, Moon, and planets onto orbits that match both observations and the geocentric theory (Ptolemy & Toomer (1984), see Figure 1.2).

[3]In the stellar magnitude system, the brightest stars are assigned a magnitude of 1, the next brightest 2, and so on (although modern astronomers have recalibrated the scale so that the star Vega has a magnitude of 0). The faintest stars visible to the human eye in an exceptionally dark location are about 6th magnitude. This scale is quite useful for qualitative observations, however, it is often inconvenient when trying to make quantitative measurements since it is easy to forget, particularly for beginning astronomers, that smaller magnitudes correspond to brighter objects.

Figure 1.1 A representation of the Ptolmatic geocentric model of the universe. Reproduced from *Cosmographia* (Apian, 1524), in the public domain.

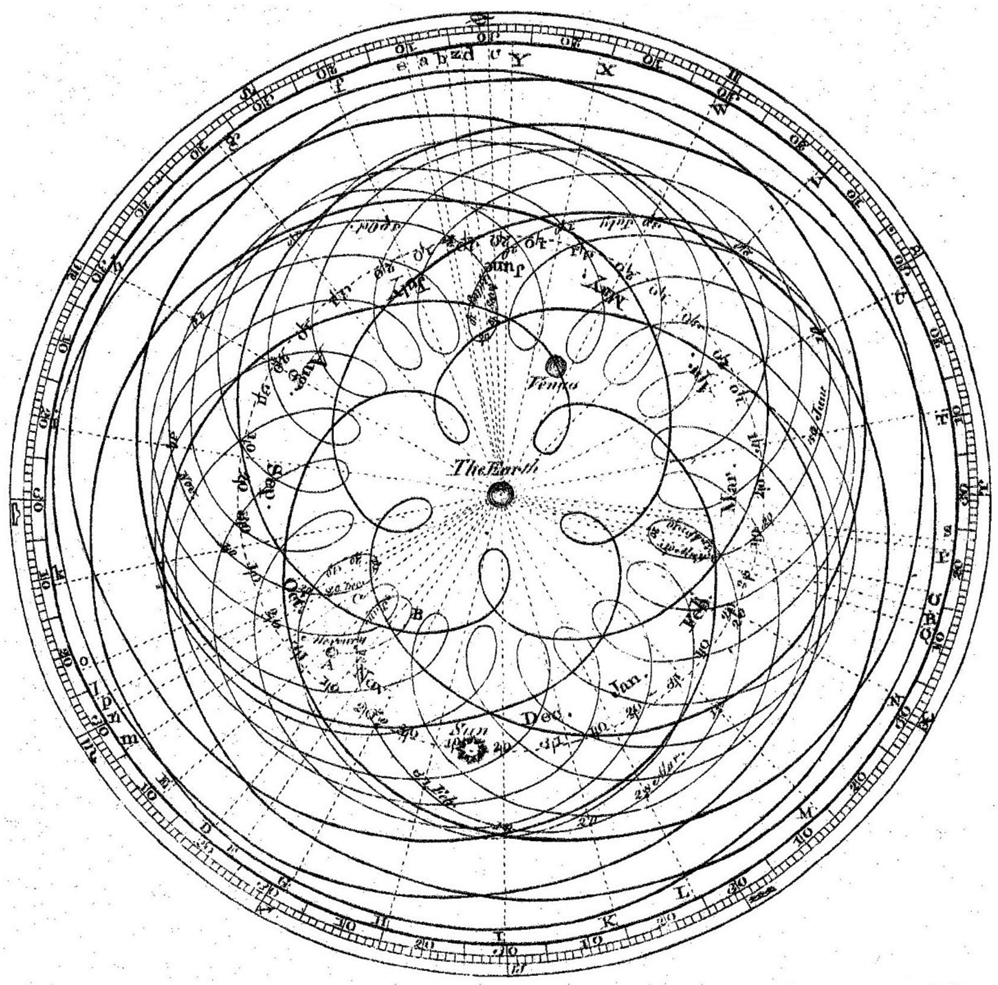


Figure 1.2 Epicycles for Mercury and Venus in the geocentric model. Reproduced from *Astronomy* in *Encyclopaedia Britannica 1st Edition* (Ferguson & Bell, 1771), in the public domain.

Although we now know the geocentric model to be inaccurate, it was an improvement on more ancient methods, since it provided a geometric explanation for the motion of the planets. This strength combined with the wide disbursement of the *Almagest* cemented Ptolemy's geocentric model as the basis of astronomy for more than 1,000 years.

It wasn't until the 16th century that another theory began to take hold. In 1543, Polish astronomer Nicolaus Copernicus published his work *De Revolutionibus Orbium Coelestium*[4] (Copernicus, 1543) while on his deathbed. He had spent nearly 30 years developing a heliocentric model of the universe, where the Earth, the planets, and the stars orbit the Sun and the Moon orbits the Earth (see Figure 1.3). Although few accepted this radical new theory at first (although he was not the first to propose that the Earth was not the center of the universe; e.g., Aristarchus c. 300 B.C.), by the time Newton had finished his work on the theory of gravity (Newton, 1687), the Copernican heliocentric theory was almost universally accepted, in part due to the support of Brahe and Kepler.

Relatively soon after Copernicus published his heliocentric model, Italian astronomer Galileo Galilei built some of the first telescopes. Although he did not invent the telescope, he was the first to make and use an effective telescope for astronomy[5]. In 1610, Galileo discovered that the planet Jupiter has moons, similar to Earth's Moon (see Section 1.3.2.1). The four largest moons of Jupiter, which he discovered, are named the Galilean satellites in his honor (although Galileo initially named them after members of the Medici family, his sponsor). In addition to his observations of Jupiter, Galileo also observed Saturn, the Moon, the Sun, and most importantly Venus. He found that Venus has phases like the Moon, an impossibility under a geocentric model, but necessary in a heliocentric system (Galilei, 1610). This finding supported the theory proposed by Copernicus, and some of the other astronomers of the day began to adopt the new model.

The development of the telescope along with Galileo's findings acted as a catalyst, increasing the rate of astronomical discovery and beginning to solidify astronomy among the physical sciences. It was during this transitional period that German astronomer Johannes Kepler, building on the work of his mentor Tycho Brahe, developed his laws of planetary motion (Kepler, 1619). These laws provided a modified version of the Copernican heliocentric theory where, instead of all planets orbiting the Sun in circular paths, their orbits were described as ellipses. Kepler's laws revolutionized the mathematical side of astronomy now known as orbital dynamics. They still form a critical foundation to understanding orbits and will be discussed in detail in Section 1.2.2.

[4] *On the Revolutions of the Heavenly Spheres*

[5] Galileo's telescopes were similar to and precursors to the spyglasses that became widely used for seafaring.

Figure 1.3 Heliocentric model of the universe put forth by Copernicus with Earth, the planets, and the stars orbiting the Sun. Page 9 in *De Revolutionibus Orbium Coelestium* (Copernicus, 1543), in the public domain.

Possibly the most important advancement in the history of orbital dynamics was the formulation of Isaac Newton's theory of gravity (although Einstein's theory of relativity may be more famous[6]). William Stukeley, an archaeologist colleague of Newton at Cambridge, wrote about the experience that inspired Newton in his *Memoirs of Sir Isaac Newton's Life*:

> after dinner, the weather being warm, we went into the garden, & drank thea under the shade of some appletrees, only he, & myself. amidst other discourse, he told me, he was just in the same situation, as when formerly, the notion of gravitation came into his mind. "why should that apple always descend perpendicularly to the ground," thought he to him self: occasion'd by the fall of an apple, as he sat in a comtemplative mood: "why should it not go sideways, or upwards? but constantly to the earths centre? assuredly, the reason is, that the earth draws it. there must be a drawing power in matter. & the sum of the drawing power in the matter of the earth must be in the earths center, not in any side of the earth. therefore dos this apple fall perpendicularly, or toward the center. if matter thus draws matter; it must be in proportion of its quantity. therefore the apple draws the earth, as well as the earth draws the apple."

(Stukeley, 1752), see also Voltaire (1727).

Following this insight, Newton (who had recently invented calculus for this purpose) began piecing together the mathematics of this phenomenon. In his hugely influential work, *Philosophiae Naturalis Principia Mathematica*[7] (Newton, 1687), Newton outlines three laws of motion: (1) objects in motion stay in motion and objects at rest stay at rest, unless acted on by an outside force; (2) when a force does act on an object, the force equals the rate of change of its momentum ($F = d\rho/dt$)[8]; and (3) when two objects exert a force on each other, these forces are equal and in opposite directions[9]. He completes his theory of gravity by defining the law of universal gravitation: everything attracts everything else with a force that is proportional to the product of their masses and inversely proportional to the square of the distance between them ($F = Gm_1m_2/r^2$)[10]. These laws are fundamental, not only to our understanding of orbital dynamics and

[6] Albert Einstein's theory of relativity, first proposed in 1905, adds to Newton's theory of gravity. It applies corrections to Newton's equations that are most evidently needed when objects are traveling at high speeds (near the speed of light). This was crucial for some aspects of orbital dynamics and provided a way to accurately predict the orbit of Mercury, which had been a puzzle to orbital dynamicists for centuries (some had even suggested another new planet interior to Mercury). To read more about relativity, see Einstein's quite accessible book on the subject *Relativity: the Special and General Theory* (Einstein, 1920)

[7] *Mathematical Principles of Natural Philosophy*

[8] Momentum is mass times velocity, so, in order to change the motion of something that is massive and/or moving quickly (like a train, for example) it requires a large force.

[9] If the Earth pulls on an apple, the apple pulls on the Earth just as much. It's just a lot easier to move an apple than the Earth, so the apple falls and the Earth hardly moves at all. For example, if a 1 kg apple falls a distance of 1 meter, the Earth would move approximately 10^{-25} meters toward the apple. That's about 10 million times smaller than the width of a proton.

[10] This is similar to Newton's 3rd law of motion, but it also depends on how far things are away from each other. If an apple is a billion miles away from the Earth, the force between the apple and the Earth is minuscule and the apple would fall to the Earth very slowly and that's only if there was nothing else with a stronger gravitational force on the apple (either more massive or closer or both) to change its course (if there were no other forces, i.e., the Sun and the other planets, and both the apple and the Earth were stationary it would take about 3,600 years for the apple to collide with the Earth, but it would burn up in the atmosphere first).

astronomy, but also to all of physics. They provide a framework for how we quantify motion in the universe as well as a mathematical method for predicting the outcome of interactions between objects, including the Sun, Moon, and planets.

In the intervening time between Newton and Gauss, several prominent astronomers — including Edmond Halley, Leonhard Euler, Johann Lambert, Joseph-Louis Lagrange, and many others — made contributions to orbital dynamics by way of mathematical predictions and observational measurements. One of the most influential discoveries during this time period was made by William Herschel and his sister Caroline. One night, using a telescope they had built, William searched the sky for double stars while Caroline recorded his observations as part of a detailed survey of the sky. During this search, they found what they initially thought to be a comet, however, after consulting with colleagues it was determined that the new discovery was a planet orbiting the Sun beyond Saturn (see Schaffer (1981) for more details on the discovery of Uranus). This discovery of the planet Uranus[11] in 1781 opened new doors of opportunity for the future of planet discovery. It paved the way for Piazzi to discover Ceres (by bringing the possibility of new planets to the forefront of astronomy), and, consequently, for Gauss to improve methods of orbital calculation from observations (Gauss, 1809). Additionally, the discovery of Uranus would provide both the inspiration and the evidence for the next planet to be discovered, Neptune.

In 1821, detailed tables of the predicted orbit of Uranus were published (Bouvard, 1821). However, when new observations were taken, Uranus was in a slightly different position from where it was predicted to be. French astronomer Urbain Le Verrier was interested in this discrepancy and in 1846 he showed that not only was the predicted location of Uranus incorrect given the theoretical understanding of orbits, but that if there was an additional planet orbiting exterior to Uranus, its gravitational effect would be enough to explain the difference between the predicted and actual orbits (Le Verrier, 1846). Excited by the prospect of a new planet, Le Verrier — who focused mainly on celestial mechanics rather than making telescopic observations — predicted the location of this mysterious planet and sent his calculations to his German colleague, Johann Galle. Galle received these calculations on 23 September 1846 and discovered the planet Neptune on the very same night, less than a degree away from where Le Verrier had predicted it would be. This was a tremendous success for orbital dynamics as it showed the powerful predictive possibilities of detailed orbit computation. Around the same time that Le Verrier made his calculations, British astronomer John Couch Adams performed similar calculations predicting the existence and location of this new planet, however his calculations were not as accurate and he gave full credit to Le Verrier and Galle for the discovery (Adams, 1846).

[11] Uranus was originally named "the Georgian star" by Herschel after King George III of England, but the French astronomers didn't like that and it was eventually renamed Uranus.

Enthusiastic because of the previous success of predicting a new planet from known planet locations and orbit calculations, American astronomer Percival Lowell began a search for a ninth planet, beyond the orbit of Neptune. Lowell called this planet "Planet X" and he predicted its mass and location based off of discrepancies in the calculated and observed orbits of Uranus and Neptune, much like Le Verrier had done for the discovery of Neptune. Unlike the story of Neptune, however, Lowell searched for Planet X for years, unsuccessful up to his death in 1916. After his death, the search continued sporadically among the staff at Lowell Observatory in Flagstaff, Arizona until 1929 when a 24 year old research assistant named Clyde Tombaugh was tasked with revitalizing the search for a new planet. Tombaugh discovered Pluto in 1930 after just under a year of his diligent search for Planet X. The world fell in love with the new planet and many adamantly show their support for Pluto by still calling it a planet despite its "demotion" to dwarf planet status in 2006 (see Section 1.3.2.6). The discovery of Pluto appeared to be another great success for predictive orbital dynamics, however, it was later determined that Pluto was not the planet Lowell had anticipated. Although Pluto was discovered near the location predicted for Planet X by Percival Lowell (about 6° away), it became apparent with the discovery of its large moon Charon and subsequent mass determination of the system[12] that Pluto was much smaller than any of the other planets and even smaller than Earth's Moon (Christy & Harrington, 1978). This meant that Pluto could not possibly account for the discrepancies in the orbits of Uranus and Neptune that Lowell had calculated (improved measurement of the positions of Uranus and Neptune later eliminated these discrepancies). Ultimately, Pluto's small size contributed to its "demotion," nevertheless, this distant icy world remains a crowd favorite and will always be remembered as the original ninth planet[13].

1.2.2 Essential Elements of an Orbit

In order to develop a solid understanding of orbital dynamics, one must first have a firm grasp of orbits and the parameters that describe them. Some of the most basic aspects of orbits were described by Kepler in his three laws of planetary motion. Kepler's first law states that the orbits of the planets and other bodies ("body" and "object" are generic terms for something in space) are ellipses[14] with the Sun at one of the foci. This is illustrated in Figure 1.4, where the blue orbit on the left is the orbit of an object (the blue dot) around a star (the black dot). The right side of Figure 1.4 shows the special case of a circular orbit. One of the characteristics of an ellipse is how elliptical it is; this ellipticity is called eccentricity. A circle can be defined as an ellipse with zero eccentricity. Most of the planets in our solar system have nearly circular

[12] Having a moon makes precise measurement of the mass of a planet quite reasonable, but without a moon, it is very difficult and involves many assumptions.

[13] "Cloyingly saccharine." - Chad Trujillo

[14] Ellipses are similar to ovals, however, ovals don't have to be symmetrical or evenly tapered like an ellipse. For example, a 2D projection of an egg would be an oval, but not an ellipse.

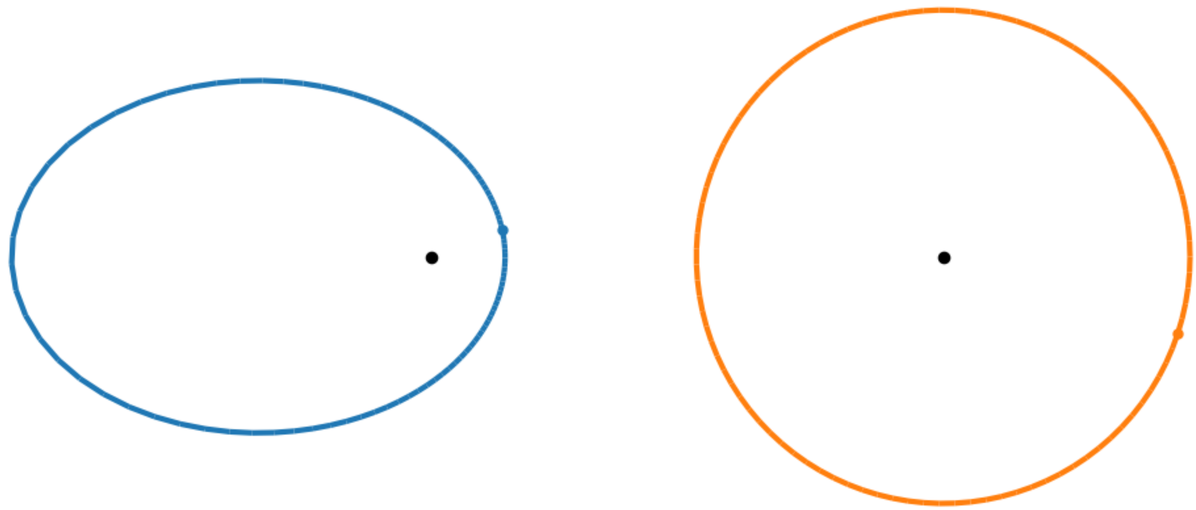


Figure 1.4 An elliptical orbit (left; $e = 0.7$) and a circular orbit (right; $e = 0$) as described by Kepler's first law. Note that a circle is a special case of an ellipse with no eccentricity. The black dots indicate the location of the Sun.

orbits, however, some planets orbiting other stars have even more eccentric orbits than shown on the left of Figure 1.4 (see Section 1.3.4). The orbits of Mercury and Mars are more eccentric than most of the other planets in our solar system and this fact was helpful for Kepler as he was developing his laws of planetary motion.

Next, Kepler stated that planets move faster when they are closer to the Sun than when they are farther away from it. He was able to quantify this geometrically in his second law, which is that a planet "sweeps" out equals areas of its orbit in equal times. This applies to other objects orbiting the Sun or other bodies, such as stars or planets. Figure 1.5 is a visual representation of this where each blue dot is the location of an object at a different time along its orbit. Times t_1 and t_2 are just over a month apart. Then after waiting several months, times t_3 and t_4 are the same amount of time apart as t_1 and t_2, just over a month; the object is moving slower when it is farther away from the Sun. Also, the orange triangular regions have the exact same area, which leads to the formal expression of Kepler's second law, equal areas in equal times.

The third law of planetary motion relates the time it takes an object to orbit the Sun to the average distance of that body from the Sun. Formally, this is written as $P^2 \propto a^3$, or the square of the orbital period of a planet is proportional to the cube of its semi-major axis. The orbital period is the time it takes for a planet to complete one orbit. For Earth this is straightforward since it takes exactly 1 year (by definition) for Earth to orbit the Sun once, $P_{\mathrm{Earth}} = 1$ year. Semi-major axis a is the average distance of an object from the Sun. Geometrically, it is also the distance from the center of an ellipse through a focus to the edge

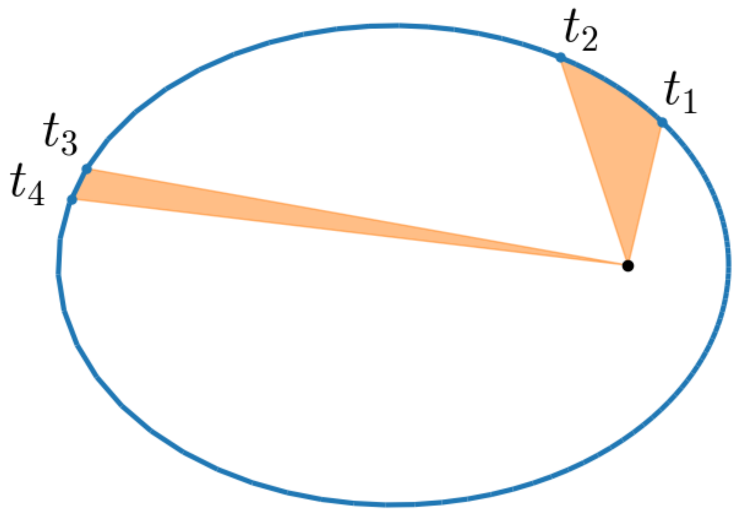


Figure 1.5 The orbit of an object at four separate times. The points t_1 and t_2 are the same amount of time apart from each other as the points t_3 and t_4. The orange triangular areas, which have the same area, illustrate Kepler's second law, which is that an object "sweeps" out equal areas in equal amounts of time.

as is shown in Figure 1.6. For a circular orbit, the semi-major axis is equivalent to the radius of the circle. The semi-major axis of the Earth (which is almost the same as the distance between the Earth and the Sun at any given time, since Earth has a very circular orbit) is defined to be one Astronomical Unit (au)[15]. The au is a common measurement of distance in astronomy and is equal to exactly 149,598,000 kilometers[16]. Because of this definition, the time it takes for a body to orbit the Sun can be calculated from just the average distance of that object from the Sun and vice versa.

Building on Kepler's laws, we can begin to describe an orbit mathematically in order to calculate the motion of planets and other bodies. One of the simplest ways to define the position of any object is to use the Cartesian coordinates X, Y, and Z. Given an origin or zero point, the position of an object can be described by its distance from that point horizontally (X), vertically (Y), and in elevation (Z). Once the location is known, the motion of an object can be broken up into V_X, V_Y, and V_Z velocity components describing the speed and direction of motion along the X, Y, and Z axes. Thus, at any given time, the motion of an object can be described using six variables. Using X, Y, and Z and their velocities to compute

[15]While both AU and au are used as abbreviations for astronomical units, we adopt the lowercase standard here.

[16]This isn't the exact average distance of the Earth from the Sun, but since that has very small gradual changes (over hundreds of millions of years) and to be more convenient, astronomers adopted 149,598,000 km as the value of 1 au. The current real semi-major axis of the Earth with this value for an au is $a_{\mathrm{Earth}} = 1.0000001$ au.

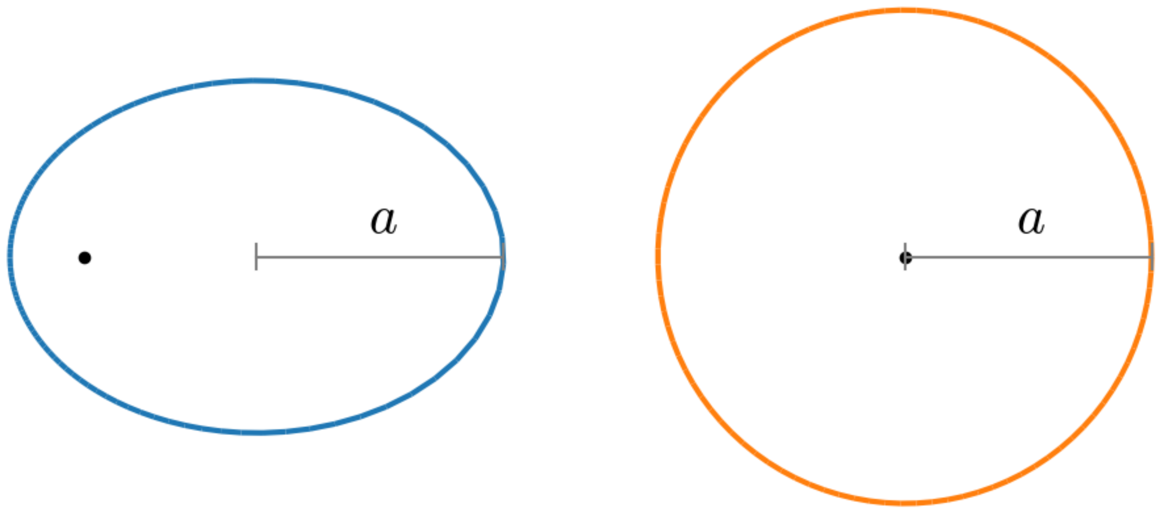


Figure 1.6 The semi-major axis a of two orbits. The semi-major axis is half way across the long (major) axis of an orbit and is also the average distance of an object from the Sun. For a circular orbit (right), the semi-major axis is equivalent to the radius of the circle.

orbits, though simple to conceptualize, does have a major drawback; as an object moves in its orbit, all six of these values constantly change. To make orbit computations simpler, the Cartesian coordinates can be transformed into a set of six orbital elements (see Murray & Dermott (1999) Section 2.8 for a detailed description of this transformation). The six orbital elements are semi-major axis a, eccentricity e, inclination i, longitude of the ascending node Ω, argument of perihelion ω, and true anomaly f. These elements have the major advantage that five of them remain roughly constant and only one (f) changes with time rather than all six[17], allowing an orbit in space to be described with the first five elements and the location of the body along the orbit to be tracked by the sixth.

We have already covered the first of these orbital elements, the semi-major axis a or average distance of an object from the Sun (see Figure 1.6). We also discussed the second element, eccentricity e, which is a measure of how elliptical an orbit is. An eccentricity of 0 means the orbit is circular and the higher the value, the more elliptical the orbit gets until an eccentricity of $e = 1$ is reached[18]. At that point, the orbit is a parabola ($e = 1$) or hyperbola ($e > 1$) and is no longer a bound orbit, but rather describes the path an object not orbiting the Sun would take as it passes the Sun. Here we will focus on eccentricities less than one, since we want to know more about objects that actually orbit the Sun or a star. Most planets in our solar

[17] Assuming there is nothing changing the orbits. In practice this is a good assumption over short time periods, since it usually takes a long time (tens of thousands of years give or take a couple of orders of magnitude depending on the system) for an orbit to change.

[18] In practice, an orbit cannot have an eccentricity of exactly 0 or exactly 1. For example, the orbit of the Earth is typically considered to be very circular and it has an eccentricity of $e = 0.0167$.

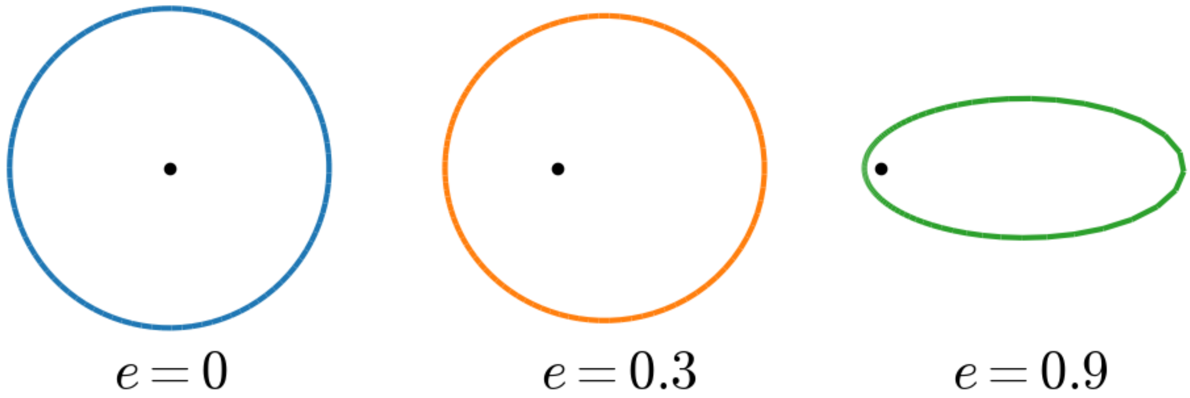


Figure 1.7 Three orbits with different eccentricities e. An orbit with an eccentricity of $e = 0$ is a circular orbit (left). As eccentricity approaches $e = 1$, the orbit becomes more eccentric (elliptical). Orbits with an eccentricity of $e \geq 1$ are no longer elliptical, but are rather parabolic ($e = 1$) or hyperbolic ($e > 1$) and they are not bound to the star, meaning they do not stay in orbit around the star.

system have eccentricities close to $e = 0$, (Mercury has the largest eccentricity at $e = 0.2$), but some solar system objects, such as long period comets, have eccentricities closer to $e = 1$ and planets in other systems have been found with eccentricities ranging from almost 0 to roughly 1. Figure 1.7 shows some examples of orbits with the same semi-major axis, but different eccentricities. The orbit on the left in Figure 1.6 has an eccentricity of $e = 0.7$. One important thing to note is that eccentricity has no units and is simply given as a number.

The third orbital element[19] is the inclination i of the orbit. Inclination is a measure of how much the orbit is tilted or inclined with respect to the reference plane of the system. In the solar system, this reference plane is the ecliptic[20], the disk in which the planets orbit[21]. Figure 1.8 highlights inclination and, as noted in the top right panel, the inclination i is the angle between the plane of the orbit and the reference plane, which, in the case of Figure 1.8, is $i = 30°$.

Fourth in the list of orbital elements is the longitude of the ascending node Ω. The ascending node is the place where an orbit crosses the reference plane, specifically where a body would cross from being below the reference plane to above the reference plane. One analogy for the reference plane is the surface of a swimming pool. A planet on an inclined orbit would be underwater when it is beneath the plane and above water when it is above the plane. The ascending node is where the planet would come up out of the water. On the other side of the orbit there is a descending node where the planet would go back underwater as it crosses the

[19]There is no real order to the orbital elements, but I typically given them in this order because of a mnemonic device I made to help remember them: The orbit vowels (even though they aren't all vowels) are a, e, i, o, w, and f or m. Here o is capital omega Ω, w is lowercase omega ω, and m is an alternative for the true anomaly f called the mean anomaly M.

[20]Not to be confused with elliptic (eccentric).

[21]It is also the path along the sky that passes through the zodiac constellations.

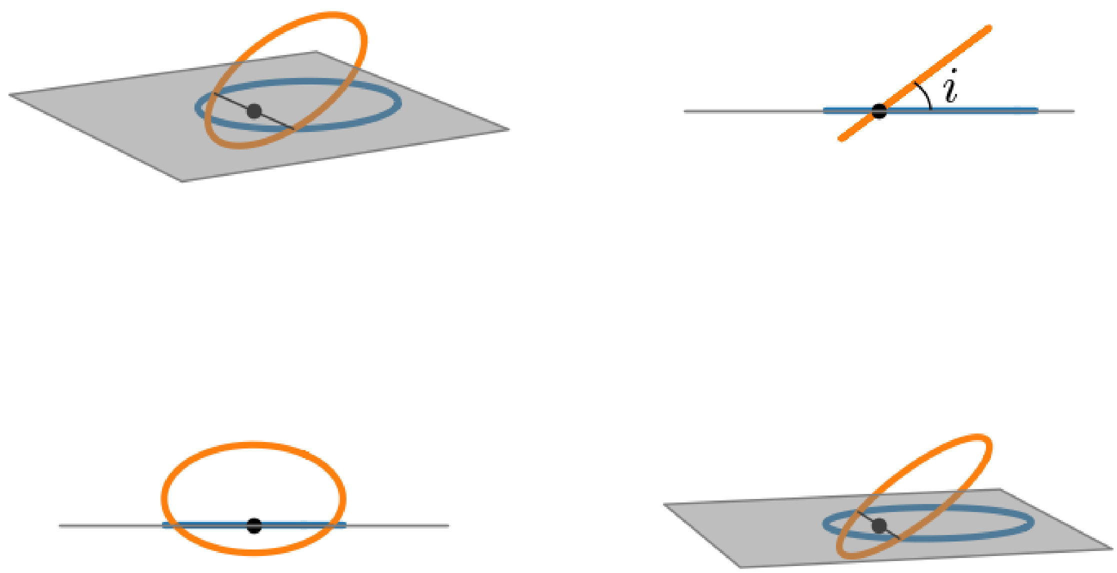


Figure 1.8 An example of orbital inclination i from four perspectives. The gray plane is the reference plane from which inclination is measured. The blue orbit is in the same plane as the reference plane and has an inclination of 0°. The inclination of the orange orbit is highlighted in the top right panel and is the angle between the plane of the orbit and the reference plane. The line passing through the star and connecting to the orange orbit indicates where the orange orbit crosses the reference plane (the line is called the line of nodes).

reference plane. In the top left panel of Figure 1.9, the ascending node is where the black line, indicating where the orange orbit crosses the reference plane (this is referred to as the line of nodes), is coming out of the page for a planet orbiting counterclockwise (the standard prograde (positive) direction for orbits when viewed from the north ecliptic pole, i.e., “above” the solar system). The longitude of the ascending node, is the angle between the direction of the ascending node (from the star) and some reference line. In the solar system, the reference line is called the Point of Aries[22]. It is related to the vernal equinox (associated with the first day of Spring) and serves for the zero point for many astronomical coordinate systems.

The fifth orbital element is the argument of perihelion ω. Perihelion q is the point in an orbit closest to the Sun. This can be seen in Figure 1.10, which also shows aphelion Q, the point furthest away from the Sun in an orbit. Perihelion and aphelion are always on opposite sides of an orbit and are always aligned along the major axis as shown in 1.10. The argument of perihelion, then, is an angle that describes the direction from the Sun to perihelion (with respect to Ω, where the orbit crosses the reference plane). Figure 1.11 gives an example of rotation in the argument[23] of perihelion ω.

With these five orbital elements, a, e, i, Ω, and ω, any regular orbit around a central body (usually the Sun, but it could be a moon around a planet, a satellite around the Earth, or even a galaxy orbiting another galaxy) can be described. The last orbital element, the true anomaly[24] f, describes the location of the orbiting body along the orbit. The true anomaly is defined as the angle between the argument of perihelion ω, the Sun, and the location of the orbiting body at a given time (see Figure 1.12). These orbital elements are summarized in Table 1.1.

An analogy for an orbit, though admittedly somewhat abstract, is someone with a hula hoop waist deep in a swimming pool. The hula hoop represents the orbit and is held around the person, initially level with the surface of the water. The semi-major axis a describes how big the hula hoop is. Next, the person bends the hula hoop into an ellipse, and prepares to start hula hooping (although this is particularly difficult to do in the water). The eccentricity e of the hula hoop describes how elliptical the hula hoop is now that it is bent. Now the person tips the hula hoop up so that the part farthest away from them comes up above the water and the part closest to them is behind them and underneath the water. The inclination i is the angle between the surface of the water and the hula hoop, or how tilted the hula hoop is out of the water. While holding the hula hoop steady so it stays in the same position around the person — with the end that’s away from them in front of their face sticking out of the water — they start to spin in a circle by just moving

[22]The Point of Aries was named in ancient times when the point was within the zodiac constellation of Aries. However, because the position of this point slowly drifts over time, the Point of Aries is currently located in the constellation of Pisces.

[23]One way to think about the argument of perihelion is to imagine the view of a royal ball from a balcony. The princess is wearing a huge dress which represents the orbit, with the Sun being her head. She is angry and is going to get into an argument, so she points in the direction of the argument of perihelion and her dress stretches out behind her in the direction of aphelion.

[24]Here anomaly just means angle.

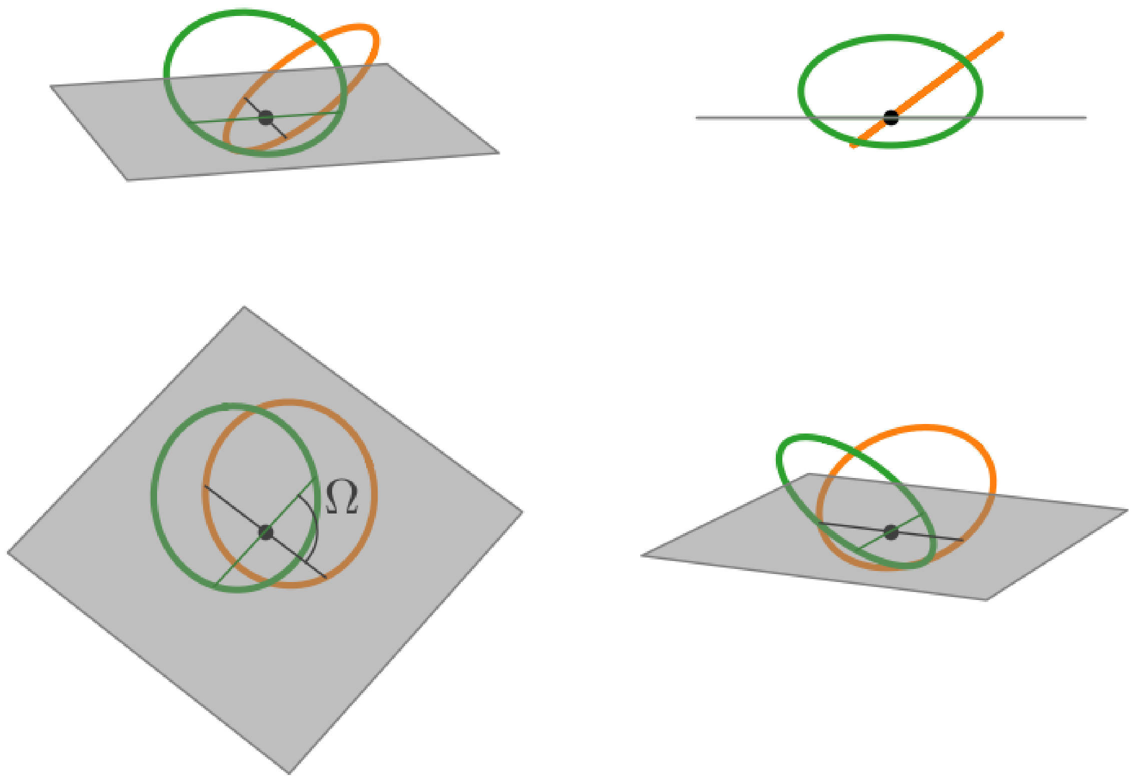


Figure 1.9 An example of the longitude of the ascending node Ω from four perspectives. The gray plane is the reference plane and the lines passing through the star and connecting to the orbit on either side are the lines of nodes (green for the green orbit and black for the orange orbit) and show where each orbit crosses the reference plane. The green orbit is identical to the orange orbit, but is rotated 90° counterclockwise about the Z axis (which is perpendicular to the reference plane). In the bottom left panel the longitude of the ascending node Ω is highlighted as this rotation, assuming that the orange orbit is initially placed at $\Omega = 0°$.

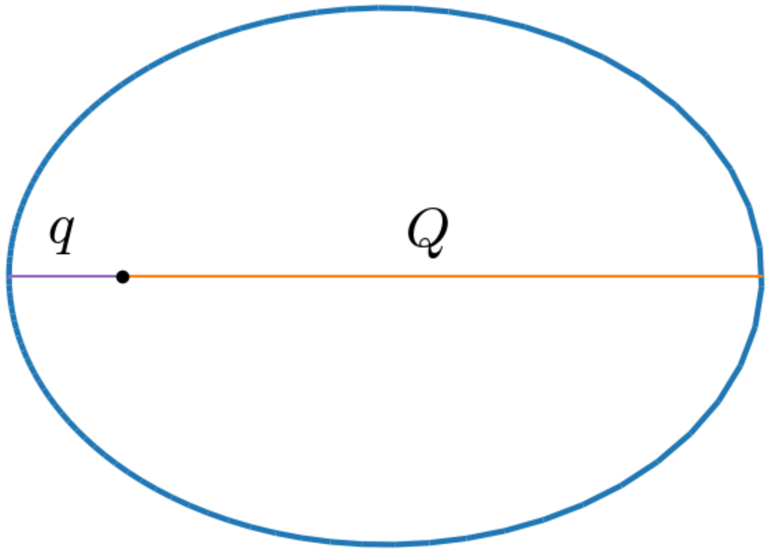


Figure 1.10 An orbit with perihelion q and aphelion Q distances highlighted. Perihelion is also the closest point in a body's orbit to the Sun (where the purple line intersects the blue orbit) and aphelion is the farthest point away from the Sun in its orbit (where the orange line intersects the blue orbit). They are always opposite each other and are along the major axis of the orbit (see Figure 1.6).

their feet while standing in the same spot. This spinning changes the longitude of the ascending node Ω, but their face is still pointed at the same spot on the hula hoop (aphelion Q, perihelion q is behind them). Now they stop spinning so they don't get dizzy. To change the argument of perihelion the person waves their arms from side to side so that their feet do not move, but the part of the hoop in front of their face dips down towards the water with the hula hoop still pointed at the same angle out of the water. Then the true anomaly f is where the planet is along the hula hoop.

In addition to the six standard orbital elements, there are many other parameters that dynamicists use to describe an orbit. The orbital period P, perihelion q, and aphelion Q, which have already been mentioned, are some of the most useful quantities to know about an orbit. They can be directly calculated from the semi-major axis a and, for q and Q, also the eccentricity e. Another useful quantity is the heliocentric distance R of an object at a given time. Here, heliocentric distance just means how far away from the Sun the object is; it can be calculated from a, e, and f. Many dynamicists choose to use an alternate orbital

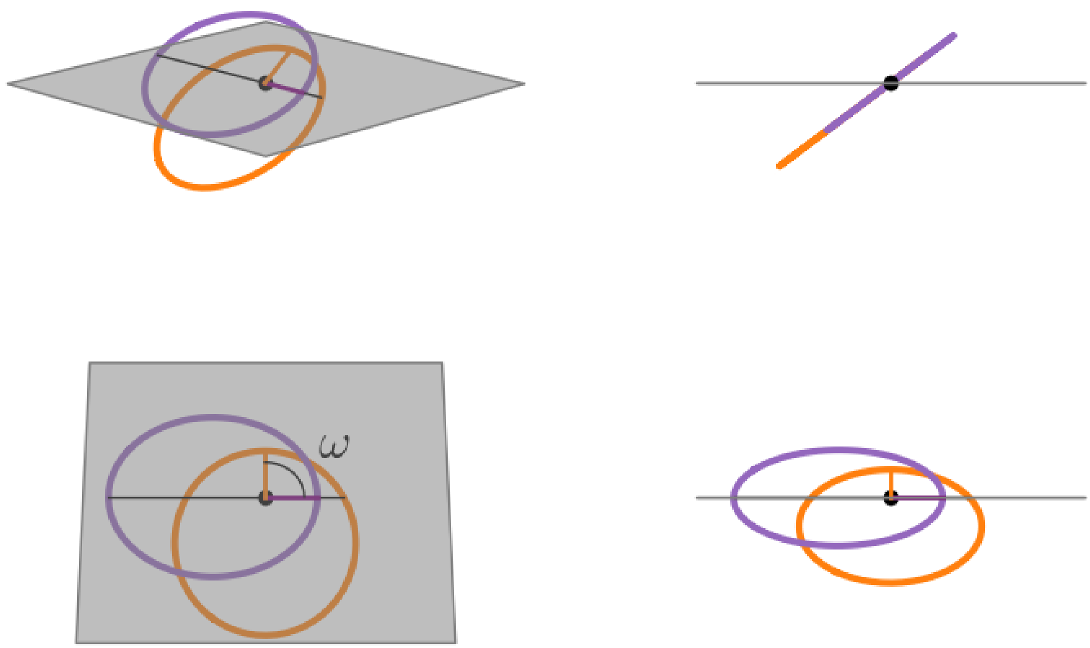


Figure 1.11 An example of the argument of perihelion ω from four perspectives. The gray plane is the reference plane (ecliptic) and the black line is the line of nodes which shows where the orbits cross the reference plane (since both orbits are in the same plane as each other, they both cross the reference plane on the same line). Lines of perihelion for both orbits are shown (see Figure 1.10). The orange and purple orbits are identical except the purple orbit has been rotated 90° in argument of perihelion ω as is highlighted in the bottom left panel.

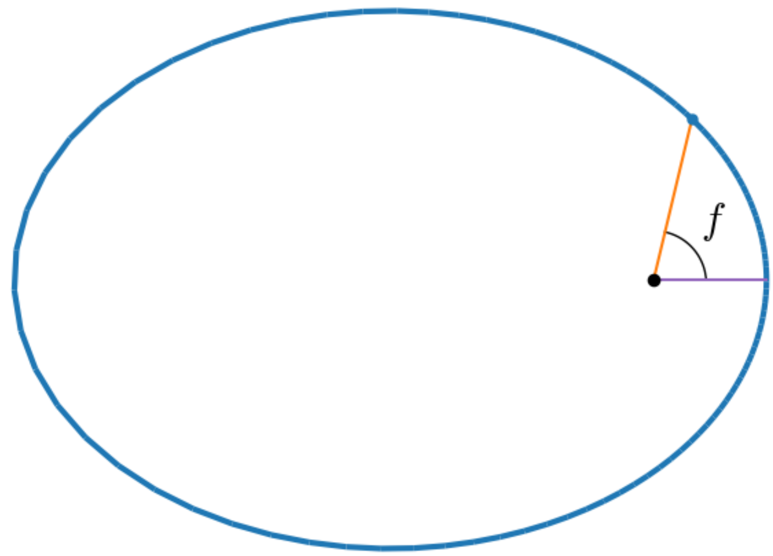


Figure 1.12 The true anomaly f of a body (blue dot) in orbit around the Sun (black dot). Perihelion is indicated by a purple line (see Figure 1.10) and the true anomaly is the angle between perihelion and the line between the planet and the Sun (orange).

element called the mean anomaly M[25] instead of the true anomaly f. The mean anomaly is defined as the fraction of the orbital period that has been completed times 360°. Mean anomaly M has the advantage of having a constant rate of change since it is, in essence, just a fraction of a duration. The true anomaly f, however, changes faster when the planet is close to perihelion and more slowly when the planet is close to aphelion[26]. There is a third anomaly called the eccentric anomaly E, and it is primarily used as a way to convert between M and f. One additional orbital parameter that is used frequently is the longitude of perihelion ϖ[27]. This parameter is calculated by adding the longitude of the ascending node and the argument of perihelion together ($\varpi = \Omega + \omega$), hence it is not a true angle since Ω and ω are in different planes. This hybrid angle is mostly used to look for patterns in the orbits of a group of objects (see Section 1.3.3 for an example). These additional orbital parameters and elements are summarized in the bottom half of Table 1.1.

[25] Capital M is also frequently used to denote the mass of the larger body (e.g., the Sun) in orbit calculations. It is a common issue in astronomy and physics to use the same letter or symbol to represent several different variables.

[26] If a planet is on a circular orbit, M and f are the same.

[27] Called pomega (like combining pi and omega together) or varpi (pronounced var-pie).

Table 1.1. Orbital Elements and Parameters

Element/Parameter	Symbol	Units	Description
semi-major axis	a	au	average orbital distance
eccentricity	e	none	orbit ellipticity
inclination	i	degrees	orbit tilt out of plane
longitude of ascending node	Ω	degrees	Z axis orbit rotation
argument of perihelion	ω	degrees	orbit rotation in the plane of the orbit
true anomaly	f	degrees	angular location of the body from perihelion
orbital period	P	years	time to complete one orbit
perihelion	q	au	closest a body gets to the Sun in its orbit
aphelion	Q	au	farthest a body gets from the Sun in its orbit
heliocentric distance	R	au	distance of a body from the Sun
mean anomaly	M	degrees	fraction of orbital period completed times 360°
eccentric anomaly	E	degrees	used to convert between M and f
longitude of perihelion	ϖ	degrees	$\Omega + \omega$

Note. — The first six entries are the standard orbital elements. Entries following the separating line are other useful derived orbital parameters and alternate orbital elements.

1.2.3 N-Body Integration

When Newton was developing his theory of gravity, the main question he sought to solve was how the planets orbit the Sun. A major implication of his findings is that the orbits of the planets are not only affected by the gravitational interaction between each planet and the Sun, but also by the gravitational interactions between the planets. In general, the Sun is by far the dominant factor in determining the overall gravitational influence on a planet, but the small gravitational pull of the other planets can make a noticeable difference, like the differences in the orbit of Uranus that led to the discovery of Neptune. These small changes to an orbit are called perturbations.

Despite the mathematical simplicity of Newton's law of universal gravitation, an equation that accurately describes the motion of just three bodies interacting gravitationally does not exist, despite diligent study by some of the greatest mathematical minds of the last four centuries (except for special cases with many simplifying assumptions, e.g., requiring $e = 0$, $i = 0°$, and the third body to have negligible mass). This makes an analytical approach to studying systems with many objects (N-bodies) difficult and cumbersome[28]. Computational methods developed in the last few decades greatly reduce the time required to calculate the gravitational interactions between bodies. Hence, the limiting size of gravitationally-interacting systems that can be explored using orbital dynamics has dramatically increased as has the precision of these calculations.

[28] Analytical dynamics, or studying the mathematics of orbit interactions, provides a wealth of information and understanding about orbital dynamics, however, the main focus of this work is on computational orbital dynamics.

N-body simulations are computer models of the gravitational interactions between N objects, where N is a standard shorthand for any number. Typically, N-body usually means simulating 4 or more bodies, such as the Sun and the planets for example. When solving for orbits analytically, the typical approach involves making some simplifying assumptions about the system, solving that simplified system, and then making some minor adjustments to try to account for some of the complexity of the real system. With numerical computations, however, performing billions of calculations in a short amount of time becomes feasible. In practice, there are dozens of different gravitational modeling codes to choose from, many of which are publicly available, and all of them approach the problem slightly differently. In general, in order to integrate (calculate motion from gravitational forces) a group of particles (which could represent planets, the Sun, asteroids, etc.), a simulation is broken up into a series of timesteps[29]. In each timestep, the positions, velocities, and forces acting on each particle are calculated and the motions of the particles are computed. Many integrators break timesteps up into small sub-steps in order to improve accuracy and several also employ coordinate transformations to increase computation efficiency. Typically, each timestep is treated independently and the integrator repeats the necessary calculations and movement of the particles until the desired end time of the simulation.

Despite the high speeds at which computers perform dynamical calculations, simulations involving large[30] numbers of mutually interacting particles are computationally expensive (require large amounts of compute time). One method for reducing the time required to integrate a system is to use test particles, which are bodies with no mass. Although this is impossible in real life, it serves as a good approximation for small bodies that are mostly acted on by the giant planets, but do not appreciably change the orbits of the giant planets. This is useful because it dramatically cuts down the number of calculations that need to be done at each timestep since the test particles only interact with the giant planets and not with each other. Full simulations with massive bodies at large scales are not practical using standard integration techniques. Other techniques can manage to handle large numbers of interacting particles, but require some simplifications to the system.

In this dissertation, we use two modern integrators to compute orbits for a variety of different objects, both in our solar system and orbiting other stars. These integrators are part of the `REBOUND` software package written in `C` and `python`. The first of these is the `REBOUND IAS15`[31] integrator; a high precision code that

[29] Choice of timestep is very important when setting up an N-body simulation. A good rule of thumb is to use a timestep no larger than $\sim 1/16$th of the effective orbital period at perihelion of the particle with the shortest orbital period in the simulation (see Hernandez et al., 2022; Wisdom, 2015).

[30] Here large is arbitrary and depends on the available processing power. For example, a simulation that is "large" when running on a laptop may be quite small in comparison to a "large" simulation on a supercomputer.

[31] 15th order Integrator with Adaptive time Stepping.

handles close encounters[32] well and is accurate to machine precision[33] (Rein & Spiegel, 2015). We use this integrator for maximum accuracy and for systems with only a few bodies (e.g., the Sun, the planets, and an asteroid). The second integrator we use is the `REBOUND MERCURIUS`[34] integrator. This gravity code uses the fast (even in comparison with other integrators) symplectic (transforms coordinates into position/momentum space for increased speed and accuracy, e.g., Holman, 1994) integrator called `WHFAST` (Rein & Tamayo, 2015) for most of the calculations so that it can compute orbits quickly, and then switches to `IAS15` whenever particles get to close to each other (how close depends on the mass of the particles), so it can achieve maximum accuracy. Once close encounters are resolved, `MERCURIUS` switches back to `WHFAST`, effectively optimizing both speed and accuracy (Rein et al., 2019). We use this integrator primarily for simulations spanning the age of the solar system with thousands of particles (see Chapter 2).

For the majority of our dynamical simulations, we use the Northern Arizona High Performance Computing cluster, *Monsoon*. *Monsoon* has over 4,000 computing cores[35], 2.3 PB[36] of storage, and can perform over 200 trillion calculations per second. Moreover, this amazing resource is free for student, faculty, and staff at NAU for research purposes. The use of *Monsoon* has been essential for the completion of this dissertation as over 2 million hours of computations have been carried out on the cluster for this research. Northern Arizona University's *Monsoon* computing cluster is funded by Arizona's Technology and Research Initiative Fund.

1.2.4 Orbital Resonances

Music is among the most universal of languages. We interpret music primarily through sounds, brought together in enjoyable combinations to our brains through our ears. These sounds travel as waves through the air caused by difference in air pressure. The changes in pressure we associate with musical notes are caused by vibrations at specific frequencies (moving a fixed number of times per second). For example, a stringed instrument such as a cello makes sound as the strings vibrate. By placing your fingers on the string, you change the length of the string which changes how fast the string vibrates (the frequency) and this changes the note. Similarly, for a wind instrument like a trombone, you can play different notes by changing the length of the pipe the air flows through and that changes the frequency of the vibrations which changes the note. These vibrations at specific frequencies are amplified by the design of the instrument. The wide opening (called a bell) at the end of a trombone vibrates at the same frequency as the air to amplify the

[32] When planets pass close by one another in their orbits, gravitational interactions between them become much stronger since gravitational force depends on the square of the distance between the objects. This causes perturbations to their orbits to be proportionally stronger and can even change their orbits following the encounter.

[33] The maximum accuracy of standard computers.

[34] It's called `MERCURIUS` because of its similarity to the popular `Mercury` integrator, but it has some improvements.

[35] This means it is has the computing power of 4,000 interconnected computers.

[36] 1 PB = 1,000 TB

sound and a cello has a large hollow cavity inside which similarly amplifies the sound of the vibrating strings. This amplification is called resonance.

Much like the way an instrument can vibrate at a specific frequency to produce a musical note, planets orbit at specific frequencies that are directly related to their orbital periods. These orbital frequencies are often appreciated as being beautiful and music-like and this was particularly the case for the ancient Greeks and also Kepler in the 17th century, who called this *Musica Universalis* (Universe Music) or the "music of the spheres" (Kepler, 1619).

As planets orbit the Sun they do so at different speeds, which causes them to occasionally pass each other in their orbits. From Newton's law of universal gravitation (see Section 1.2.1) we know that the force of gravity gets stronger when two bodies are closer together ($F = Gm_1m_2/r^2$)[37]. Because the distance term r is squared, when objects are closer together the force gets much stronger. Because of this, almost all of the gravitational pull two planets exert on each other happens right at the time when they are passing each other in their orbits. This gives both planets a little gravitational "kick," called a perturbation, that slightly changes their orbits. However, planets typically pass each other at different locations around their orbit. This translates to the "kicks" all being in different directions and, over a long time, they all average or cancel out.

There are some pairs or even groups of bodies that have orbital periods (frequencies) in harmony relative to each other such that rather than passing the other bodies at random locations along their orbits, they always pass each other in the same place. An analogy for this is two runners on a track. A fast runner is on the short inside lane and a slow runner is on the longer outside lane. Rather than the standard staggered starting line used in track, these runners are going to start at the same starting line and that line will also be the finish line. In this race, the fast inside runner has to run around the track twice and the slow outside runner just has to run around once. The runners are exactly fast enough and slow enough that they get to the finish line at the exact same time, every time. This is a resonance and when it happens in an astrophysical system (e.g., the solar system, or an exoplanetary system), rather than all of the gravitational kicks canceling out, they add up and amplify the effect the bodies have on each other, similar to how musical instruments amplify the note forming vibrations to make them louder. This effect can play a major role in stabilizing or destabilizing[38] these systems. Here we will discuss three types of orbital resonance that are focused on in later chapters. Understanding resonances is an extremely important aspect of orbital dynamics and one of the most active areas of study in the field. This is largely because resonances can be one of the

[37]Here, F is the force of gravity, G is the gravitation constant, m_1 and m_2 are the masses of the two objects, and r is the distance between the objects.

[38]Destabilizing a system usually results in the destruction or ejection of some of the bodies from the system.

dominant factors dictating the structure of a given planetary or other system and seemingly subtle resonant interactions can cause dramatic changes to that structure over time.

Perhaps the most common and most basic orbital resonance is the Mean Motion Resonance (MMR). In a Mean Motion Resonance, the orbital periods of the bodies are integer multiples of each other, e.g., if planet A goes around a star once, planet B goes around twice or three times just like in the analogy of the runners on a track. One excellent example of an MMR in our solar system is the resonance between three of the Galilean moons of Jupiter: Io[39], Europa, and Ganymede. These three moons are in a 4:2:1[40] resonance. This means that Io orbits Jupiter 4 times in the same time it takes Europa to orbit Jupiter 2 times, which is also the time it takes Ganymede to orbit Jupiter once. However, having all three lined up at the same time would destabilize the system, so they are staggered in a way that whenever two of the moons line up, the third one is in a different part of its orbit. This staggered separation helps to keep the system stable. Because these moons orbit Jupiter quite quickly (Io's orbital period is less than two Earth days), this resonance is effectively transferring energy within the system all the time. Io ends up with a lot of this excess energy through a process called tidal heating (which is enhanced by this resonance) and, consequently, is the most volcanically active body in the solar system[41].

As seen with the resonance among the Galilean moons of Jupiter, one of the most stabilizing characteristics of MMRs are when they ensure the bodies do not come close together. This is the case for the 2:3 MMR between Pluto and Neptune. Pluto orbits the Sun twice in the same time that Neptune orbits the Sun three times. Additionally, Pluto's orbit ($e = 0.25$) crosses interior to Neptune's orbit. However, because of the resonance between them, they never come within 2 au of each other (see Figure 1.13) and, thus, Pluto can stay on its Neptune-crossing orbit without being ejected from the solar system.

Another important type of resonance, secular resonances, rely on a process known as apsidal precession. Apsidal refers to the argument of perihelion ω, which is also sometimes called the argument of periapsis[42]. Precession is a slow rotation of the axis of a body spinning about that axis. One way to visualize this is to imagine a helicopter; the body spinning on the axis is comprised of the blades of the helicopter. As the blades rotate rapidly, the rotor (the shaft that spins the blades) can tilt in different directions and represents the axis of rotation. If the rotor tilts forward, the helicopter moves forward, if the rotor tilts sideways, the helicopter moves sideways. Precession can then be described as a helicopter facing a given direction,

[39] Pronounced EYE-oh, but the ancient Greeks would have said EE-oh.

[40] Pronounced 4 to 2 to 1.

[41] Io is roughly the same size as Earth's Moon, and there are images taken by spacecraft (e.g., New Horizons) showing gigantic volcanoes erupting and ejecting material from Io into space.

[42] The prefix *peri* is Latin for near. The Apsides (plural of apsis) are, by definition the closest and farthest points from a focus on an ellipse, so the same as perihelion and aphelion, if referring to objects orbiting the Sun. *Helion* means Sun, so perihelion is near the Sun. Periapsis is the general term that means near the central body an object is orbiting. Stars in Greek are *astron*, so if a planet is orbiting a star other than the Sun, the closest point in the orbit that planet to that star is called periastron.

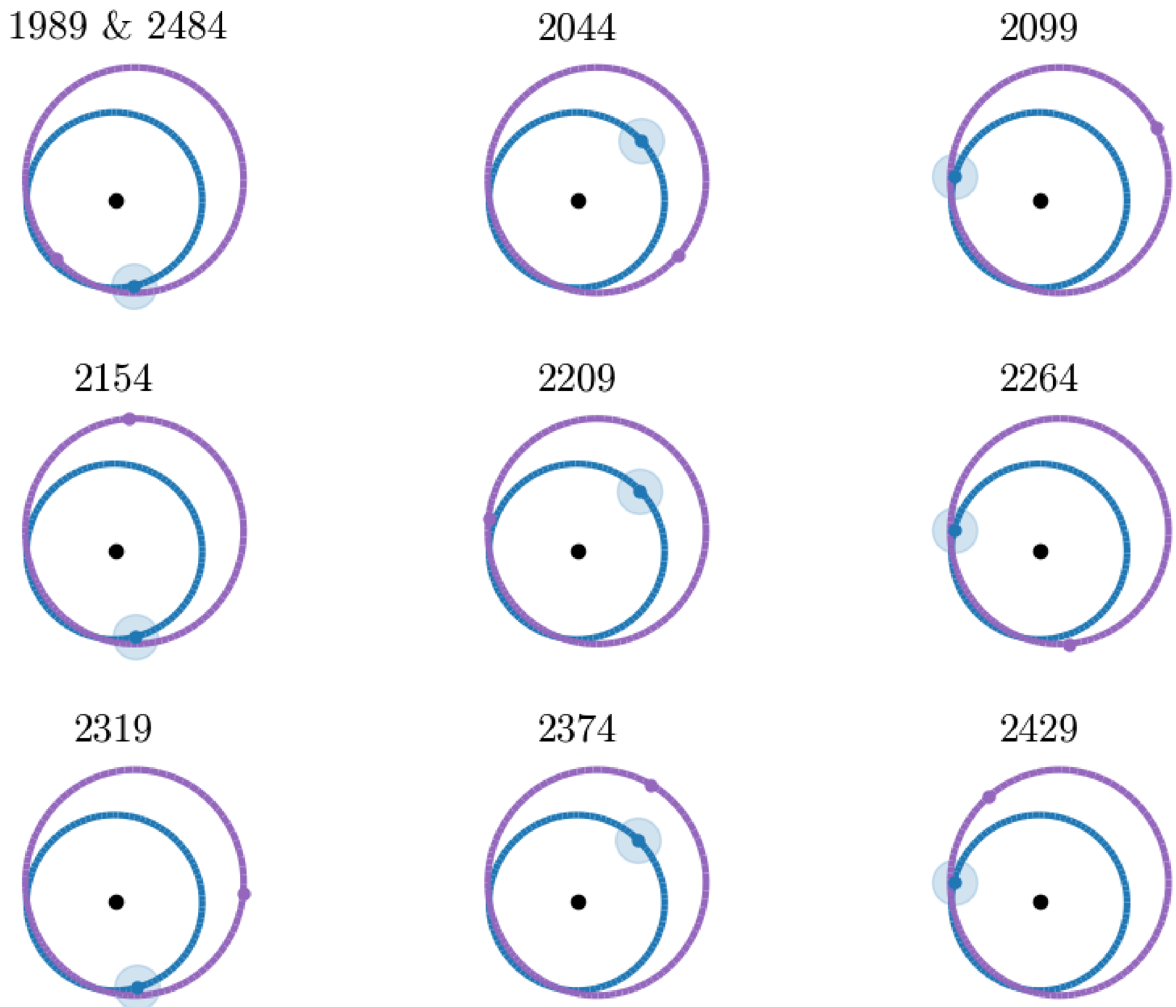


Figure 1.13 The orbits of Neptune (blue) and Pluto (purple) over one 2:3 Mean Motion Resonance cycle. At time $t = 0$ yrs (top left) Pluto is at perihelion (5 Sept 1989). The sphere of influence ($\sim$10 Hill radii, see Hill, 1878) of Neptune is shown as a transparent blue circle surrounding Neptune (orbits are to scale and the sphere of influence is approximately to scale, but the Sun, Neptune, and Pluto are not to scale). Objects coming within this region have their orbits perturbed. Both Neptune and Pluto orbit counterclockwise and their orbital configuration is shown every 55 years (1/3 of Neptune's orbital period). Neptune returns to its starting position every three frames (left column) because of this cadence. After the ninth frame ($t = 440$ yrs in 2429), Neptune and Pluto return to their original positions roughly 495 years later (in 2484). Because of the Mean Motion Resonance between Neptune and Pluto, where Neptune orbits three times in the same time Pluto orbits twice, the only time the two bodies approach each other on the same side of their orbits (same ecliptic longitude) is while Pluto is near aphelion and is far away from Neptune (bottom middle). This weakens the influence of Neptune on Pluto and allows Pluto to retain its Neptune-crossing orbit and remain stable over billions of years. Without this protective resonance, Pluto would have a close encounter with Neptune that would likely result in Pluto being ejected from the solar system.

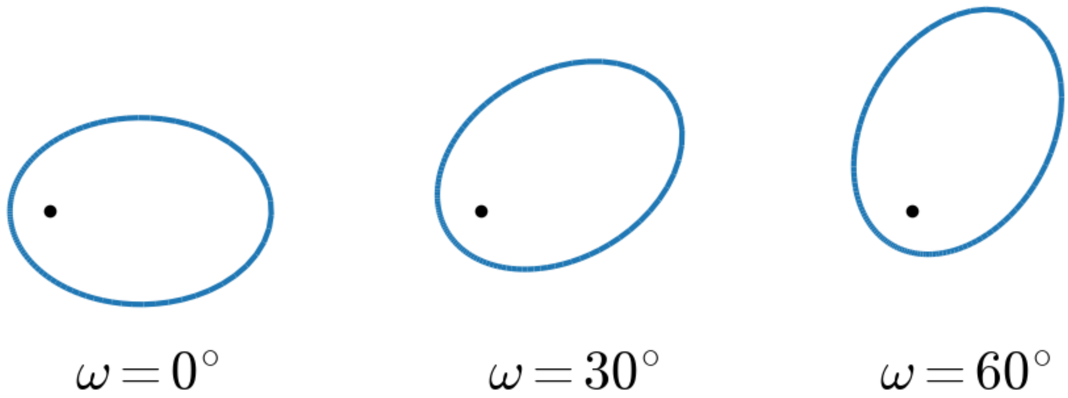


Figure 1.14 Precession of an orbit. The left panel shows the initial condition and, after some time (how long depends on the orbit and other factors), the argument of perihelion ω has shifted to its value in the middle panel. The orbit then continues to precess through the position shown in the panel on the right and then all the way around until it arrives at its original location. The process then repeats.

North will be used for convenience. The helicopter first moves forward then slowly transitions from moving forward to moving to the left all while still facing North. Then, the helicopter slowly transitions to moving backward, then to the right, then forward again, facing North the entire time. During this time, the blades of the helicopter "orbit" the rotor many times and the rotor itself makes one tilted revolution during the entire sequence. This change in the tilt angle of the rotor is analogous to the precession of an orbit. So, apsidal precession is where the argument of perihelion slowly changes and makes a complete revolution over a long timescale (112,000 years for the Earth[43]) In Figure 1.14, an example of precession is shown for an orbit at three different times. The orbit on the left is where it initially starts. After a long time (dependent on the orbit) and many orbital periods, the orbit slowly precesses and the argument of perihelion ω changes as is seen in the middle panel. This process continues and the orbit will continue to precess over 360° until it returns to its initial value, at which point it will begin another cycle. The longitude of the ascending node Ω can also precess in a similar manner.

The next type of orbital resonance is called a secular resonance. Here, the term "secular" means to occur over very long timescales[44], e.g., millions of years in some cases, and is typically contrasted with more regular periodic motion (such as MMRs). Rather than having a ratio between the orbital periods of planets, secular resonances are a ratio between precession periods (the time it takes to complete one precession cycle for either ω or Ω or both). These resonances, although they happen much slower than MMRs, can have a large effect on bodies in resonance. Over time, secular resonances can change the eccentricity e and inclination i

[43]This is one of the factors in long-term climate variation.

[44]Here the term secular is not referring to the non-religious. Rather, it is derived from the Latin word *Saeculum*, meaning a long time or something that would happen only once in a lifetime.

of a small body (like an asteroid) in resonance with a large body, such as a planet. The specifics of how e and i change is very dependent on the interaction. Some examples of secular resonances include a secular resonance with Saturn (ν_6) that defines some of the boundaries of the asteroid belt, long term interactions between Planet X and Extreme Trans-Neptunian Objects as discussed in Chapter 2, and Kozai resonances which are a type of long term interaction between three bodies that can dramatically change the orbit of the tertiary body[45]. There are several other kinds of resonances that apply in a wide array of astronomical phenomena from forming gaps in the rings of Saturn to causing ripples through spiral galaxies. Resonances can be and often are the dominant factor in the orbital structure of astrophysical systems. They play a key role in orbital stability and configuration and can be crucial for constraining planet location.

1.3 The Solar System and Beyond

In this section, an overview of relevant populations of solar system bodies is provided. The defining characteristics of these populations are highlighted and comparisons of various bodies provide a sense of relative scale.

1.3.1 The Planets

After the Sun, the planets are the largest bodies in the solar system. There are several ways in which the planets are ordered; the most common is by their distances from the Sun (semi-major axes a), but ordering them by their size and/or mass is also common. Table 1.2 gives some of these basic parameters for the planets and distance and size comparisons for the planets are shown in Figures 1.15 and 1.16 respectively. There are notable differences between the four inner planets and the four outer planets. The inner planets have orbits that are tightly spaced with just over 1 au of spacing between Mercury and Mars. They are also small (less than 10,000 km in radius) with relatively low mass (between 10^{23} and 10^{25} kg), while the outer planets are orders of magnitude larger and more massive, and they have greater orbital separation between them, with approximately 25 au separating the orbits of Jupiter and Neptune. Additionally, the inner and outer planets have differing compositions. The inner planets are considered to be terrestrial because of their rocky composition, whereas the outer planets are known as gas giants due to their extended gaseous atmospheres which contribute significantly to their large sizes and mass.

The dramatic differences between the terrestrial and gas giant planets arise largely from the formation process of these planets. The solar system formed from a giant cloud of unorganized gas and dust (on the

[45]Kozai resonances have been seen in a variety of systems such as the outer solar system, exoplanetary systems, and artificial satellites orbiting the Earth.

Table 1.2. Planet Parameters

Planet	Mean Radius	Mass	Semi-major Axis	Known Moons
Mercury	2,440 km	0.055 $M_\oplus$	0.387 au	0
Venus	6,052 km	0.815 $M_\oplus$	0.723 au	0
Earth	6,371 km	1.000 $M_\oplus$	1.000 au	1
Mars	3,390 km	0.107 $M_\oplus$	1.524 au	2
Jupiter	69,911 km	317.8 $M_\oplus$	5.204 au	80
Saturn	58,232 km	95.16 $M_\oplus$	9.583 au	83
Uranus	25,362 km	14.54 $M_\oplus$	19.19 au	27
Neptune	24,622 km	17.15 $M_\oplus$	30.07 au	14

Note. — $M_\oplus$ stands for Earth masses.

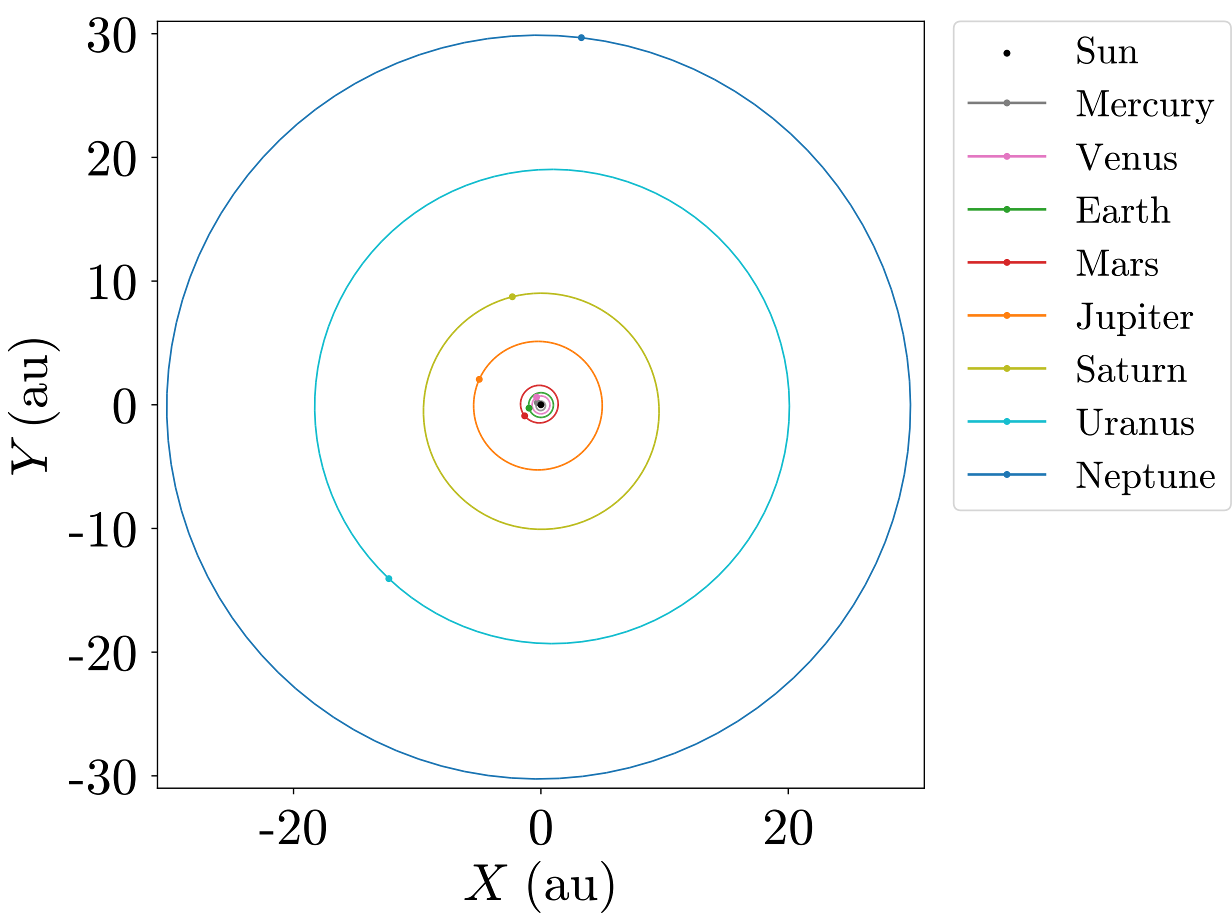


Figure 1.15 A plan view of the orbits of the planets. Average orbital distances can be found in Table 1.2. The locations of the planets on 5 April 2063 are also shown.

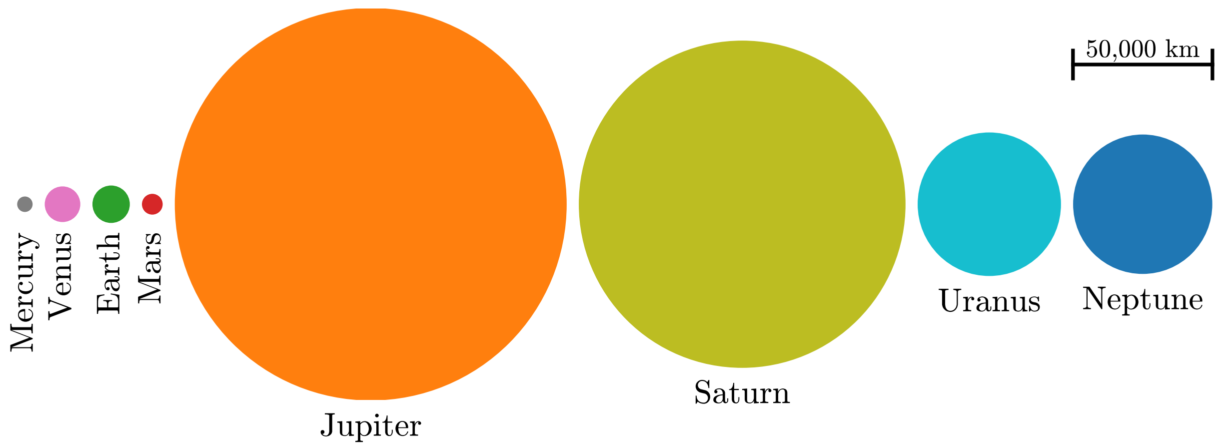


Figure 1.16 Relative sizes of the planets. The planets are shown in order of their distance from the Sun. The rings of Saturn (not shown) have an outer radius of ∼80,000 km. The Sun (also not shown) has a diameter roughly 10 times larger than that of Jupiter. Radii and masses for the planets can be found in Table 1.2. The colors of the planets are the same as in Figure 1.15.

order of 10^6 solar masses of material and extending over 10^7 au Petit & Morbidelli, 2005) which began to collapse when perturbed by an outside force (likely a supernova Cameron & Truran, 1977). The center of this collapse eventually became the Sun, surrounded by a spinning disk of planet-building material (similar to how a ball of pizza dough flattens out into a disk when it is spun). Through collisions and clumping caused by gravity and probably other forces (e.g., electrostatic Petit & Morbidelli, 2005), sub-micron sized pieces of material started to stick together. These pieces combined to form planetesimals, small bodies that form the building blocks of planets (Carrera et al., 2017). Some of these planetesimals grew into planets as they accreted material, including other planetesimals, while others formed the smaller bodies in the solar system or were ejected during periods of planet migration (see review by Montmerle et al., 2006).

Within the protoplanetary disk where the planets were growing, there was a thermal gradient with higher temperatures close to our young Sun and cooler temperatures farther out in the disk. Moving from the Sun outward, the temperature decreases with distance. The distance where the freezing point of water is reached is called the "ice line" ("frost line" and "snow line" are also used). Temperatures beyond the "ice line" are low enough that water freezes, becoming stable on planetary surfaces. This further enables the growth of the giant planets, which formed beyond the ice line. Conversely, the terrestrial planets formed too close to the Sun to acquire large amounts of water ice, hence, their growth was more limited (see Chambers, 2016, and references therein). There is still some water (including in solid form) in the inner solar system, however, and it has been proposed that some portion of this was likely transported to its present location from the outer solar system (Morbidelli et al., 2000).

In addition to these compositional differences between the terrestrial and giant planets, they also formed at different times. In the early solar system there was a large disk of planetesimals (on the order of 10-100 Earth masses) that were starting to grow into planet sized objects (Hahn & Malhotra, 1999; Levison et al., 2011). The majority of these planetesimals did not end up as part of one the planets we know today, but were instead ejected from the solar system by an interaction with one of the giant planets (Gomes et al., 2005). Because of Newton's third law, when a planetesimal interacts with a giant planet, the planetesimal exerts an equal and opposite force on the planet. Each of these gravitational encounters altered the orbits of the giant planets and, over time, this caused what is known as planetary migration, where the giant planets experienced a drift in their semi-major axes over millions of years (Hahn & Malhotra, 1999; Tsiganis et al., 2005).

At some point during this period of planetary migration, it is theorized that Jupiter and Saturn arrived at an unstable Mean Motion Resonance with each other. This caused dramatic jumps in the orbits of all of the giant planets and resulted in the ejection of most of the remaining material within the solar system that had not been swept up by the planets (Gomes et al., 2005). Throughout the process of solar system formation and evolution planets play a central role in determining the structure of the systems in which they reside. See Nesvorný (2018) for a detailed review of solar system evolution.

1.3.2 Small Bodies

Objects in the solar system that are not planets are collectively referred to as small bodies or minor planets. The term small bodies typically includes moons, comets, asteroids, Centaurs, Trans-Neptunian Objects, dwarf planets, and even interstellar objects[46]. Moons are occasionally excluded from this list because they primarily orbit their planet[47]. In the following sections, we will describe these bodies and provide size comparisons between several bodies in these categories.

1.3.2.1 Moons

A moon is a natural satellite[48] orbiting a planet. Moons are typically form from disks of material surrounding planets, similar to the rings around Saturn[49]. Small particles coalesce into moons orbiting their host planet much in a process similar to how planets form within a protoplanetary disk (Nesvorný, 2018). One source of this disk of material is large impacts on the planet. The resulting debris forms a disk from which moons

[46]Objects entering our solar system from the space between the stars, interstellar space. Only two of these have been discovered so far, but future large scale telescope surveys are expected to find several more in the next decade.

[47]Some of the moons in the solar system are larger than the smallest planet, Mercury, so calling them small bodies is somewhat ironic.

[48]As opposed to an artificial one. Starlink satellites are not moons.

[49]Rings can also be formed by material coming off of moons, such as the ring formed by geysers of water spraying into space off of Saturn's moon Enceladus (Spahn et al., 2006).

can form (Canup & Asphaug, 2001; Hesselbrock & Minton, 2017). A moon can also be another small body that was captured by the gravity of a planet and is now trapped in orbit around that planet, such as Triton, the largest moon of Neptune (Agnor & Hamilton, 2006).

There is a large variety of physical properties and surface features among the moons of the solar system. They range from little more than barren rocks — such as Phobos and Deimos, the moons of Mars — to intricate worlds full of geologic and hydrologic activity — Titan, the largest moon of Saturn, for example, has methane rain and rivers (Turtle et al., 2011). In Figure 1.17, we compare the sizes of several of the largest moons orbiting each planet in the solar system (excepting Mercury and Venus, which have no known moons). Note how Ganymede and Titan are both larger than the planet Mercury. The number of known moons orbiting each planet is given in Table 1.2.

It is possible to have moons orbiting a body other than a planet, such as an asteroid or Trans-Neptunian Object, several of which have known moons. Conversely, there is no requirement that planets have moons. Theoretically, a moon could have a moon[50], although there are no known examples of this (Kollmeier & Raymond, 2019).

1.3.2.2 Comets

Comets have been observed since ancient times[51] and are some of the longest-studied bodies in the solar system. The distinguishing features of a comet are tails and comae. A coma is a bright region around the nucleus (the solid center) of a comet and a tail is a long stream of material extending from the nucleus. Primary tails are made of dust that a comet leaves behind it as it moves in its orbit. Comets can also have secondary tails, called ion tails, that are composed of gases being impacted by the solar wind. Just as a flag would be blown in the wind, material from the comet is blown in the direction away from the Sun forming this additional tail. The coma is the brightest part of a comet since this is the region surrounding the nucleus, where material is actively being sublimated from the surface.

The material that makes up both the comet and the tails is primarily comprised of gas (ions) and dust, and volatile solids, such as water ice, in the case of the nucleus. Comets originate from the outer solar system, with different sub-groups of comets orbiting near Jupiter, between the orbits of Jupiter and Neptune, or in the most distant reaches of the solar system (Levison, 1996). When comet orbits bring them closer to the Sun, their temperature begins to rapidly increase. This temperature increase causes ices (water, as well as other volatile species, such as, carbon dioxide, methane, etc.) on and in the comet to begin to sublimate,

[50] These are sometimes called moon-moons. Astronomers are notoriously bad at naming things sometimes.

[51] In ancient times comets were viewed as signs of doom and destruction, but now when a bright comet comes it is typically seen as a reason to throw a party.

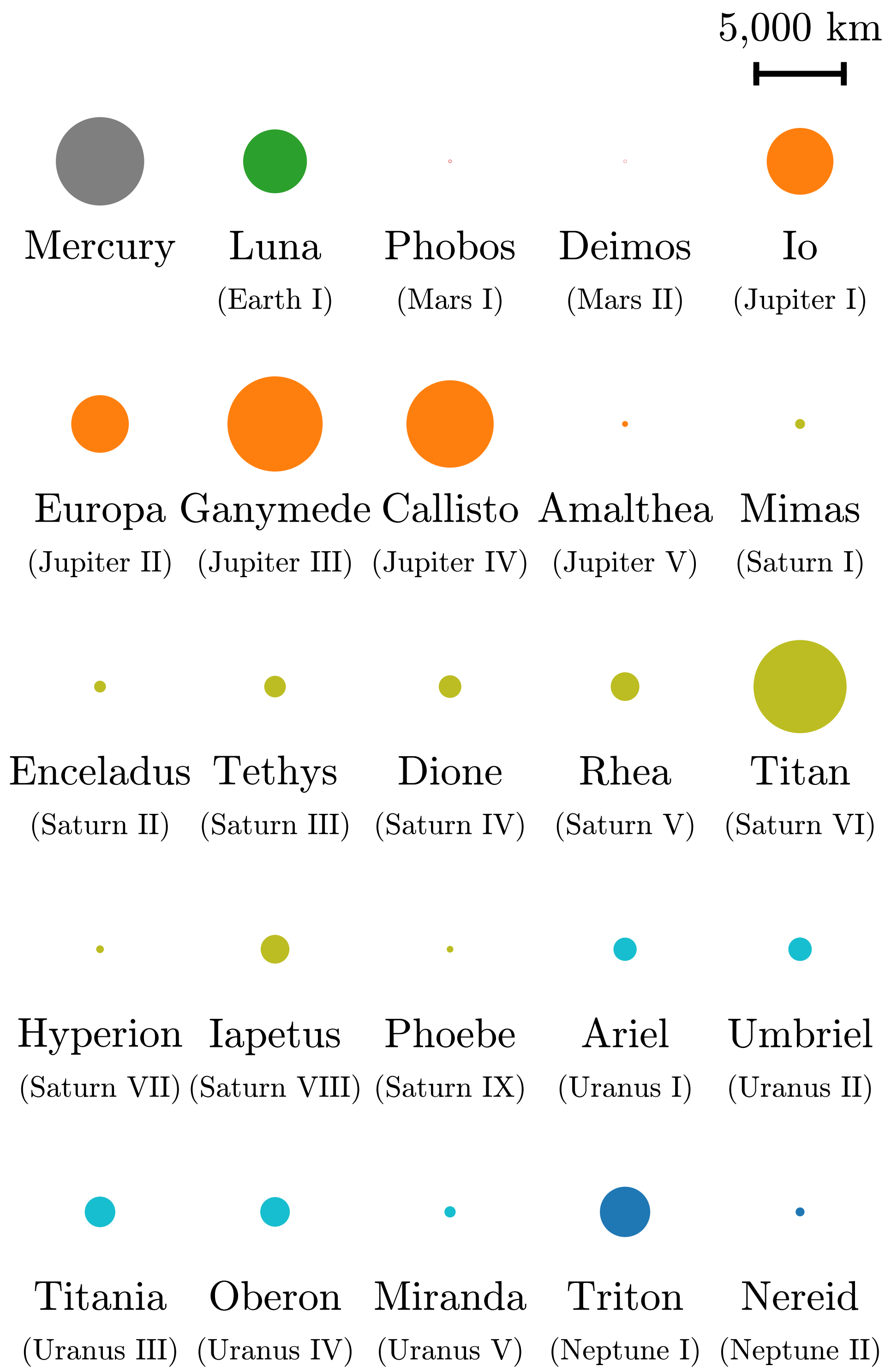


Figure 1.17 Relative sizes of some of the largest moons of each planet (with known moons) in the solar system. The planet Mercury, which is smaller than Ganymede and Titan, is shown for comparison. Colors correspond to the colors of the host planet shown in Figure 1.16. Moons are numbered using Roman numerals, with low numbered moons, such as those shown here, primarily ordered by their orbital distance from their host planet. New moons are now numbered in chronological order of their discovery.

changing from a solid state directly to a gas. This change occurs with enough force to eject these gasses, and some of the dust layer that was covering them, into space.

There are two main classes of comets — short period comets (such as Halley's comet) and long period comets (e.g., comet Hale-Bopp). The short period comets have aphelia Q among the giant planets and typically have orbital periods on the order of tens of years. The long period comets come from the distant reaches of the outer solar system, beyond Neptune, and have orbital periods of hundreds or thousands of years. One notable sub-class of short period comets are the Jupiter Family Comets (JFCs). These comets have orbital periods of roughly 10-20 years, and they reside near the orbit of Jupiter. The JFCs will be discussed more in Chapter 4. See Levison (1996) and Kaib & Volk (2022) for a reviews of comets in the solar system and their origins.

1.3.2.3 Asteroids

In contrast with comets, asteroids are mostly rocky bodies. They usually do not show signs of cometary activity, such as tails or comae, and they are, for the most part, located within the asteroid belt. The asteroid belt is located between the orbits of Mars and Jupiter and is divided into three main groups, the inner, middle, and outer belts. The divisions between these groups, called Kirkwood gaps after their discoverer, Daniel Kirkwood (Kirkwood, 1867), are caused by resonances with Jupiter (MMRs) and Saturn (secular resonances) which overlap to create unstable regions that occupy the Kirkwood gaps between the three asteroid groups (Dermott & Murray, 1983) (see Section 1.2.4 for a discussion of MMRs and secular resonances). Similar to the terrestrial planets, the asteroids formed closer to the Sun than the ice line in the inner regions of the solar system and, thus, suffered from a lack of water ice during the early period of rapid growth of bodies in the solar system. Additionally, the asteroids experienced extreme perturbations from the giant planets during planet formation, which limited their growth (Nesvorný, 2018).

Apart from the roughly 1 million known asteroids located in the asteroid belt, there are tens of thousands of asteroids with orbits that bring them close to the orbit of Earth. Near Earth Asteroids (NEAs)[52] form as main-belt asteroids drift[53] into the unstable regions where their orbits are perturbed by resonant interactions with Jupiter and Saturn. These asteroids have their orbits altered and some of them shift into orbits that intersect the orbit of Earth. NEAs are a major focus of study in astronomy, since they have a risk of impacting the earth. NASA currently has a spacecraft mission called DART (the Double Asteroid Redirection Test) which will intentionally crash into the moon of asteroid Didymos[54] in order to change its orbit slightly. The

[52] Sometimes called Near Earth Objects (NEOs), probably because the acronym sounds cooler, but also to include comets that come close to the Earth.

[53] The drifting is orbital diffusion caused primarily by collisions between asteroids.

[54] Didymos' moon is named Dimorphos.

goal of the DART mission is to show, as a proof of concept, that with sufficient lead time (~20-30 years), NASA would be able to alter the orbit of a NEA and change its course to prevent a disastrous impact on the Earth (Rivkin et al., 2021). I have done some work for this mission analyzing telescope data of Didymos, and, while not the focus of a chapter in this dissertation, the work resulted in a paper published in the Planetary Science Journal, Pravec et al. (2022).

Additionally, there are a few thousand known asteroids in the 1:1 MMR with Jupiter. This means that they share roughly the same orbit as Jupiter, but orbit 60° ahead of, or behind Jupiter. These 1:1 MMR asteroids are called Trojan[55] asteroids. There are a few known Trojan asteroids in resonance with other planets, but only a few of them have been discovered[56], so, the term Trojan asteroid usually refers to Jupiter Trojans. Trojan asteroids were captured into the 1:1 resonance during giant planet migration (Nesvorný, 2018).

Asteroids have been discovered in a wide range of sizes, from Ceres, the largest, at nearly 1,000 km in diameter, down to tiny asteroids only a couple of meters across. In Figure 1.18 a size comparison of some of the asteroids (selected because of the visits or planned visits of various spacecraft missions) with the Moon is shown. Note that only Ceres is actually close to spherical (although a few come relatively close) and the rest of the asteroids are a variety of different shapes[57].

There is a small subset of asteroids (~30) that occasionally show signs of comet-like activity. These active asteroids are enigmatic as they blur the traditional boundary of the definitions between asteroid and comet. There are several proposed mechanisms that may cause this activity including standard sublimation, outburst caused by impact, structural failure and breakup (as could be caused by the tidal effects of passing too near to a giant planet), and spinning at high enough speeds to lose material (Sheppard & Trujillo, 2015). Active asteroids will be discussed in greater detail in Chapter 4.

1.3.2.4 Centaurs

The Centaurs are a particularly interesting population of small bodies for orbital dynamics studies. This is because they are both few in number, with only a few hundred Centaurs known, and their orbits are unstable on the order of a few million years (Jewitt, 2009), short in comparison with the age of the solar system (4.5 billion years). Centaurs orbit in the giant planet region, between the orbits of Jupiter and Neptune. This location explains both the small number of known Centaurs, since they are distant and faint, and their short

[55] This is in reference to the Iliad (Homer, c. 630 BC) where Zeus (Jupiter) has a hand in a great war between the Trojans and the Greeks. The two groups of Trojan asteroids are called the Greek camp, those leading Jupiter in its orbit, and the Trojan camp, those trailing Jupiter in its orbit.

[56] 1-2 for Venus, Earth, and Uranus; 10-30 for Mars and Neptune; and none for Mercury and Saturn.

[57] Many of them are potato-shaped or top-shaped. A body needs to have a high enough mass to structural strength ratio in order to pull itself into a spherical shape.

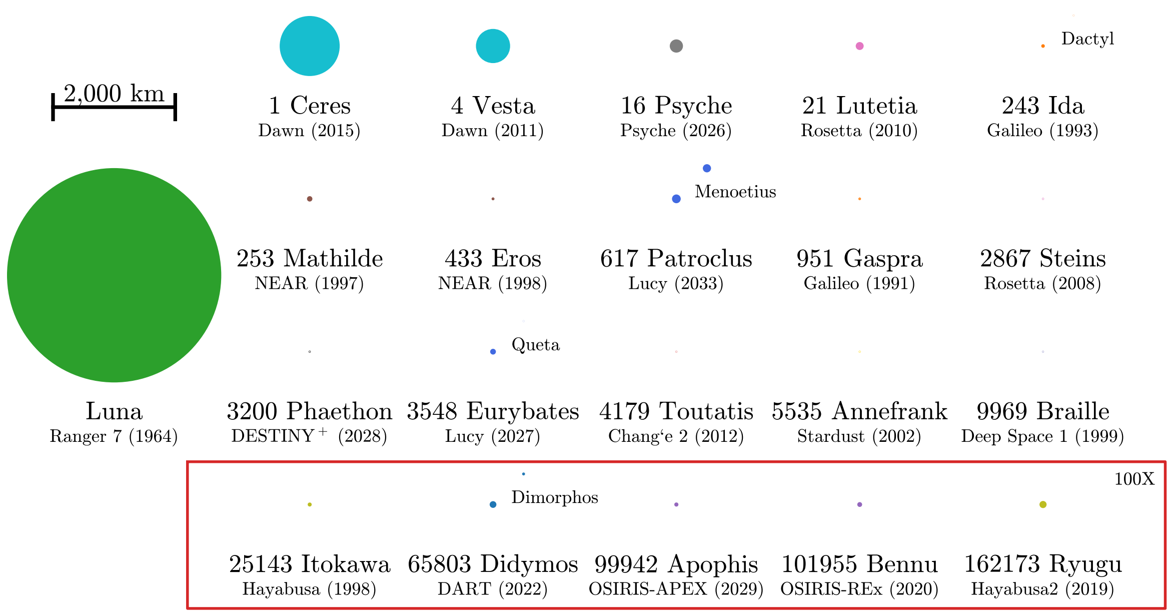


Figure 1.18 Relative sizes of several asteroids and their known moons. The Moon (Luna) is shown for comparison. Asteroids in the bottom row of the figure (red box) are shown with 100 times magnification. Asteroids were selected because they either have been or will be visited by spacecraft in the near future. Colors are somewhat arbitrary, but asteroids visited by the same spacecraft are shown in the same color (e.g., Ida and Gaspra) as are their moons. Asteroids are shown (left-to-right, top-to-bottom) in order of their discovery. Note that the majority of asteroids are not spherical in shape and only relative diameters are shown here.

dynamical lifetimes (how long they stay in a particular orbit), because they have frequent gravitational interactions with the giant planets.

Centaurs originate in the region beyond Neptune where they formed as Trans-Neptunian Objects (see Section 1.3.2.5). Then, through a gravitational interaction with Neptune, they migrate inwards to the giant planet region. While in this region dominated by the giant planets, close encounters with the giant planets further alter the orbits of the Centaurs. These encounters often bring Centaurs onto orbits closer to the Sun or eject them from the solar system. When they are close enough to Jupiter, gravitational interactions can cause Centaurs to migrate into the orbital region of the Jupiter Family Comets. Thus, Centaurs are transition objects migrating between the outer solar system beyond Neptune and the inner solar system interior to Jupiter's orbit (where they cease to be Centaurs). This migration is an important process for replenishing cometary populations (such as the Jupiter Family Comets) in the inner solar system (Jewitt, 2009).

The transition from the distant icy outer solar system to the much warmer inner solar system can cause sublimation driven cometary activity in Centaurs, even prior to fully migrating to the inner solar system (Jewitt, 2009). Because different volatiles (ices that will sublimate into gases) sublimate at different temperatures, multiple episodes of activity are expected for Centaurs based on their composition (water ice will sublimate at a different temperature, and, hence, a different heliocentric distance, than, for example, carbon dioxide ice or methane ice due to their material properties). Not all Centaurs have been observed to be active, but, due to their presumably icy compositions, Centaurs are likely to have undergone cometary activity in the past and/or are likely to show signs of activity in the future (see, for example Chandler et al., 2020).

1.3.2.5 Trans-Neptunian Objects

In the region beyond the giant planets, we find the Trans-Neptunian Objects (TNOs). The TNOs have semi-major axes greater than a_{Neptune} and only a few thousand have been discovered to date. They are typically divided into several sub-classes, the first being the Kuiper belt[58]. The Kuiper belt is disk of icy asteroids, somewhat similar to the asteroid belt, that extends from semi-major axes of roughly 30-47 au. Additionally, the Kuiper belt can be broken up into three groups, the cold classical Kuiper belt, the scattered disk, and the resonant objects.

[58]Named after 20th century Dutch astronomer Gerard Kuiper. There is, however, some naming controversy surrounding the Kuiper belt. Kuiper's contemporaries Kenneth Edgeworth, Fred Whipple, Julio Fernández, and Frederick Leonard have all been suggested as better namesakes for this asteroid-belt-like feature, particularly since Kuiper predicted that such a feature did not exist (but did during the formation of the solar system). However, because Kuiper was the most famous, the name stuck and is still commonly used, although some astronomers simply avoid the term whenever possible to avoid controversy and instead exclusively refer to objects beyond Neptune as TNOs.

The cold classicals[59], as they are often called, are presently of little interest dynamically, although this population does place important constraint on the migration of Neptune in the early solar system (Levison et al., 2008; Nesvorný, 2015). They have low eccentricities, low inclinations, and undergo practically no orbit-altering gravitational interactions. Sometimes the cold classicals are simply designated as the Kuiper belt, excluding the other sub-classes, and/or are called Kuiper Belt Objects (KBOs). The scattered disk is composed of objects that are on more eccentric (e between roughly 0.2 and 0.7)[60], more inclined (up to $\sim 30°$) orbits and have greater potential for interaction with Neptune (Volk & Malhotra, 2008). The resonant TNOs are in stable Mean Motion Resonances with Neptune and are further classified by the resonance in which they reside[61].

At eccentricities above that of objects in the scattered disk (which are called Scattered Disk Objects or SDOs), there is another sub-class of TNOs sometimes called Extreme Trans-Neptunian Objects (ETNOs). ETNOs are called extreme because their high eccentricities, which can sometimes reach above $e = 0.95$, resulting in very elliptical orbits (see Figure 1.7). The ETNOs can be further divided into extreme scattering ETNOs, which have perihelia $q < 40$ au, and extreme detached ETNOs, with $q > 40$ au (e.g., Sheppard et al., 2019). These definitions have only been used in the past decade and there is not a solid agreement between astronomers on the exact definitions of the ETNOs. For example, some astronomers only consider the extreme detached objects to be ETNOs (in which case the extreme scattering TNOs are classified as SDOs), and the line between populations is not always drawn at $q = 40$ au. Additionally, there are three known objects at very high perihelia, $q > 65$ au, Sedna, 2012 VP_{113}, and Leleākūhonua (Sheppard et al., 2019). These objects have been classified by several different designations including ETNOs, Inner Oort Cloud objects (IOCs)[62]. The Oort cloud is a theoretical cloud of comets surrounding the Sun at semi-major axes of thousands to hundreds of thousands of au and is thought to be the source of the long period comets (Oort, 1950; Duncan et al., 1987). In Figure 1.19, one set of definitions for these different object classes is given along with a plot of the perihelia and eccentricities of known objects in the outer solar system. Although astronomers have yet to agree on standard names for many of these populations of objects, they do represent the edge of our knowledge of the solar system. As this edge is expanded through increased telescope observation capability, a wealth of new objects providing greater insight into the structure of the outer solar system will doubtless be uncovered.

[59] Infrequently abbreviated as CCs and also sometimes called cubewanos (Q-B-1-os), after 1992 QB1, which is now named Albion, the first TNO discovered after Pluto.

[60] As more TNOs are discovered definitions of and divisions between the different sub-populations of TNOs will change. Hence, the outer solar system and the TNOs are arguably one of the most active areas of astronomy research in the solar system.

[61] For example, objects in the 2:3 MMR with Neptune, like Pluto, are called Plutinos and objects in the 5:7 resonance are called 5:7 resonant TNOs.

[62] Sometimes a capital O is used for Objects, in which case they are IOCOs. In older papers they are also sometimes called Sednoids, named after Sedna, the first of these objects to be discovered.

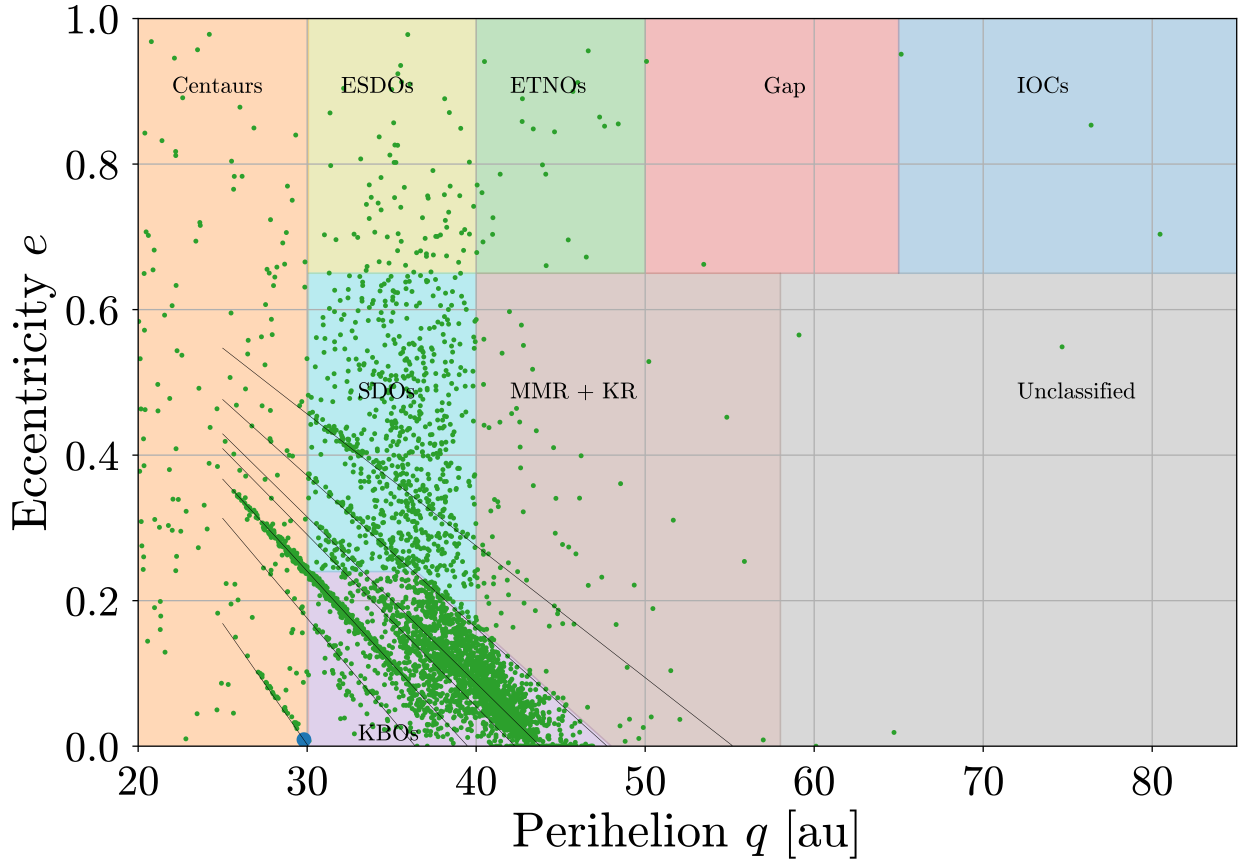


Figure 1.19 Orbital classes of the outer solar system. Neptune is shown as a blue filled circle at $q = 29.8$ au and $e = 0.01$. TNOs and Centaurs are shown as green points. Centaurs (orange region) orbit within the region of the giant planets and have perihelia interior to the orbit of Neptune (some definitions also require that their semi-major axes are also interior to the orbit of Neptune). The classical Kuiper Belt (KBOs; purple region) is located beyond Neptune and objects in this region have low eccentricities. Scattered disk objects (SDOs; cyan region) interact gravitationally with Neptune, which increases their eccentricities to moderate values. Resonant objects, in mean motion resonance with Neptune, reside along the black diagonal lines (not all MMRs with Neptune are shown). Extreme Scattered TNOs/Extreme Scatter Disk Objects (ESDOs; yellow region) have been scattered to high e though interactions with Neptune. ESDOs are often simply classified as SDOs and are sometimes considered ETNOs (however, this is usually only done in studies seeking to discount evidences for Planet X due to observational bias). ETNOs (green region) cannot form through interactions with Neptune alone and require an outside perturber (such as Planet X) in order to arrive at their locations. The perihelion gap (red region) is sparsely populated, especially in comparison to the more distant IOC population (blue region). This is enigmatic since objects should be much easier to detect within the gap than in the IOC region. This will be discussed in detail in Chapter 2. The brown region contains objects that are placed on their current orbits through a combination of MMRs and Kosai resonances. Objects in the gray region have yet to be classified with, or separate from, any of the other regions. Some of the objects shown do not have multiple oppositions of observations, and hence, will likely have different orbital parameters than are shown here once their orbits are determined more accurately. For discussions on the regions in this figure and examples of their usage, see, e.g., Gladman et al. (2008), Gomes et al. (2008), Sheppard et al. (2019), and Oldroyd & Trujillo (2021).

Formation mechanisms for the different populations of TNOs are an active area of study. The cold classicals likely formed near where they are now and the scattered disk and resonant objects were shepherded into their orbits by Neptune migration in the first few hundred million years of planet formation (Levison et al., 2008; Nesvorný, 2015). This is supported by the stark color difference between the KBOs, and the SDOs. KBOs are very red, indicating extremely old and unaltered surfaces (planetary surfaces redden as they age through a process called space weathering; see Barucci & Merlin, 2020), whereas the SDOs are bluer, meaning they have younger surfaces. The extreme scattering ETNOs form as SDOs and further interact with Neptune, increasing their eccentricities, and sometimes inclinations, to much higher values. The extreme detached ETNOs and the IOCs, however, cannot form from interactions with Neptune and require an outside force (Oldroyd & Trujillo, 2021). Their formation is discussed in detail in Chapter 2.

1.3.2.6 Dwarf Planets

Possibly the least popular astronomy event of the 21st century, the "demotion" of Pluto to the designation of dwarf planet remains a controversial topic with the public. In the decades after Pluto was discovered in 1930 (Tombaugh, 1961), several astronomers suggested that there may be more planets like Pluto orbiting beyond Neptune (Leonard, 1930; Kuiper, 1951; Edgeworth, 1949). Many astronomers searched for these planets, but none were found. Then, in 1992 David Jewitt and Jane Luu discovered a second Trans-Neptunian Object 1992 QB1 (Jewitt & Luu, 1993), which was later named Albion[63]. At this point, Pluto was known to be much smaller than the other eight planets (when it was discovered, Pluto was thought to be larger than Earth), but Albion was another story entirely. Measuring at just over 100 km in diameter, Albion is tiny in comparison to the planets and is even much smaller than the largest asteroids (although it is much larger than most asteroids; see Figure 1.20). Albion was clearly not a planet, but something else, more like an asteroid. In the following years, more similar objects were found and, thus, the Kuiper belt was discovered.

The discovery of the Kuiper belt opened up a new frontier for solar system discovery and many new TNOs have been discovered every year since. Then, in the early 2000's, Mike Brown, Chad Trujillo, and David Rabinowitz discovered several TNOs that approached and even rivaled Pluto in size (see, for example Brown et al., 2005). The largest of these was Eris, discovered in 2005, which, at the time, was thought to be larger than Pluto[64]. Eris was designated by some as the tenth planet, but this was controversial and brought up the question of whether Pluto may not merit the status as planet.

[63] In the time between the discoveries of Pluto and Albion, spaceflight was made possible, humans landed on the Moon, digital cameras were invented (a huge leap forward for astronomy), telescopes were launched into space, and spacecraft had visited all of the planets except Mercury and Pluto (which was still a planet at the time).

[64] Current estimates for the diameter of Eris place it just under Pluto by about 50 km.

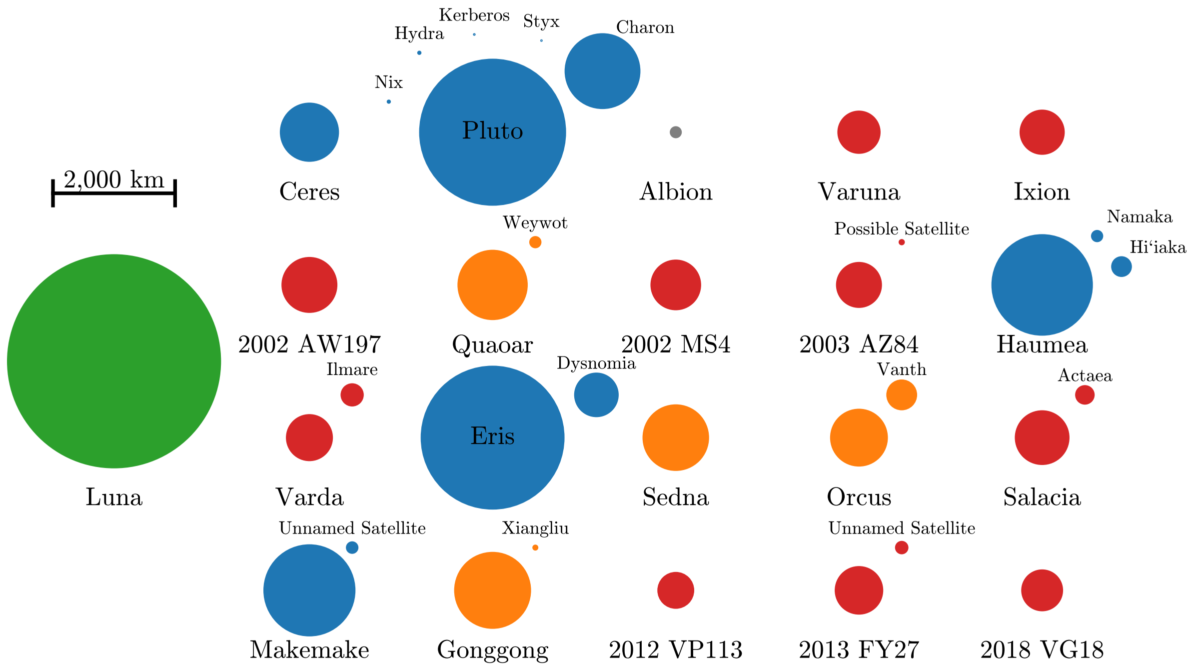


Figure 1.20 Relative sizes of official and unofficial dwarf planets. The Moon (Luna, green) is given for comparison. The scale of this figure is the same as the asteroid size comparison given in Figure 1.18. Objects in blue are officially recognized as dwarf planets by the IAU, and their moons. Orange objects are the most likely known candidates to become officially recognized dwarf planets in the near future, and their moons. Red objects are also similarly large bodies, and their moons, that are considered dwarf planets in this work. Albion, the first TNO to be discovered after Pluto, is also shown (gray) for comparison. Primary objects are shown (left-to-right, top-to-bottom) in order of their discovery. Note that not all of these object are spherical (e.g., Haumea) and only their relative diameters are shown.

In 2006, at a meeting of the International Astronomical Union (IAU), the formal definition of a planet was discussed and debated. In this debate there were two main arguments for what bodies should be considered as planets. The more lenient definition proposed focused on the ability of a body to gravitationally pull itself into a sphere[65]. Under this model, any body that was roughly spherical[66] would be considered a planet, provided it was orbiting the Sun (to exclude moons from also being planets), potentially resulting in dozens of planets[67]. The stricter definition of a planet involved a slightly more abstract concept, the ability of a planet to gravitationally dominate the region surrounding its orbit. In practice, this is calculated as a function of a body's mass and semi-major axis with more distant bodies requiring more mass to be dominant since there is more volume to cover the farther away the body is from the Sun; however, such calculations were only developed after this definition was in place (see, for example Margot, 2015). This criteria includes the eight bodies classified as planets today (see Table 1.2), but excludes Ceres, Pluto, Eris, and the other large asteroids, TNOs, and moons.

Although Pluto and similar bodies are unable to "clear their orbits," they are still sufficiently massive to be spherical, a trait that sets them apart from the myriad smaller objects in the solar system. In order to highlight this difference, a new class of object was defined, the dwarf planet. By IAU definition, a dwarf planet is a solar system body that is massive enough to be spherical, but that is not gravitationally dominant, and does not orbit another planet. This resulted in the demotion of Pluto from a planet to a dwarf planet. Much like the demotion of Ceres approximately 200 years earlier, as new objects are discovered, perspective and understanding of the structure of the solar system changes; hence, the way objects are classified changes to match that understanding. The debate over the planethood of Pluto has been continued in the years following the 2006 dwarf planet designation. While some astronomers agree with the new classification system, there are many who do not[68]. One of the most egregious issues with the current IAU definition of a planet is that it requires that a body orbit the Sun in order to be a planet, hence, exoplanets (discussed in Section 1.3.4) are not considered planets under this definition. As the rate of discovery in the outer solar system accelerates, more information will continue to shape how its structure and composition are viewed and classification systems will continue to evolve.

There are currently five bodies in the solar system with the formal IAU designation of dwarf planet. They are Ceres, Pluto, Eris, Haumea, and Makemake. There are several other large TNOs that have been

[65] This is called hydrostatic equilibrium and is a balance between the force of gravity, for which the state of lowest potential energy is for the mass to be distributed in a sphere, and internal pressure or structural strength of the body.

[66] Because bodies rotate, they are not perfectly spherical. The shape of the Earth is an oblate spheroid, a sphere that is bulged out along the equator because of its spin. The dwarf planet Haumea is shaped more like a rugby because it is spinning quite rapidly (with a rotational period of less than four hours), but it mathematically meets the hydrostatic equilibrium criteria and meets this formal requirement.

[67] Under this system, the eight planets would be called "classical planets."

[68] The term planet is derived from the Greek word planētai, meaning wanderers. So, the ancient Greeks would have considered any similar moving light in the sky, including Pluto, a planet.

suggested to be dwarf planets (most notably Sedna, Quaoar, Orcus, and Gonggong), but the IAU has not formally recognized them as such. It is difficult to distinguish whether a body is capable of pulling itself into a sphere, primarily because it not only depends on the mass of the body, but also on its internal structure, which, at best, can only be inferred based on several generalizations about these objects in many cases. This means that objects at the lower end of possibly being dwarf planets will likely remain off of the IAU's official list until much more detailed observations are able to reduce the generalizing assumptions used to classify these objects. The officially recognized dwarf planets are well beyond the minimum requirements and easily qualify. In this work, the term dwarf planet is used to refer to both the officially recognized dwarf planets and other similar bodies in the solar system.

One of the most pertinent questions regarding the dwarf planet definition and the demotion of Pluto is *Why does it matter whether Pluto is a planet or not?* In the general sense, it does not matter if Pluto is called a planet or a dwarf planet; it is a fascinating world full of dynamic and interesting features, it is historically significant, and it inspires millions of people world wide and, regardless of its name or designation, it has merit simply for what it is. On the other hand, one of the most fundamental ways scientists make sense of the universe is through classification. Being able to quantify the similarities and differences between two objects is often extremely useful and can provide insight into why those similarities and differences exist. Even if a classification is inadequate at fully representing an object, much can be learned about that object through the process of discovering why the classification is inadequate. So, even though it may not ultimately matter whether Pluto is a planet or dwarf planet, as trends and patterns that govern the outer solar system are discovered, Pluto and other bodies will be reorganized and reclassified with objects that exhibit similar orbital and physical characteristics.

1.3.3 Planet X

Since the the discovery of Neptune, astronomers have been searching for an additional, more distant planet in our solar system. Percival Lowell called this hypothetical world Planet X and this same term is adopted here, even though the planet discussed here has a very different mass and orbit than predicted by Lowell (Lowell, 1915; Brunini, 1992). In 2003, the dwarf planet Sedna (also an IOC) was discovered with extremely high semi-major axis (a = 506 au), eccentricity (e = 0.85), and perihelion (q = 76 au). This was an exceptional discovery in part because Sedna could not be placed on its current orbit through Neptune scattering (interactions with Neptune) as its high perihelion restricted Sedna from coming within 45 au of Neptune. Hence, Sedna was labeled as part of the inner Oort cloud (Trujillo et al., 2005). Additionally, the orbit of Sedna is very difficult to achieve without the influence of a massive object at a further distance from the Sun altering that orbit. One suggested external perturber is an additional planet in our solar system (see, for example, Trujillo & Sheppard, 2014).

Then, in 2014, Chad Trujillo and Scott Sheppard discovered a second IOC, 2012 VP_{113} (Trujillo & Sheppard, 2014). Along with this discovery, Trujillo and Sheppard made two important observations, (1) the arguments of perihelion ω were aligned for all objects with $a > 150$ au and (2) there was a gap in perihelion q between these two distant IOCs and other ETNOs. The clustering in ω was suggested as evidence for an additional planet in the solar system and a separate study by Mike Brown and Konstantine Batygin two years later further supported the existence of an additional solar system planet[69] (Batygin & Brown, 2016a). Since then, there have been over a hundred studies regarding the potential existence and implications of this hypothesized planet (see review by Trujillo, 2020, and references therein).

Despite the evidence in favor of Planet X, several astronomers assert that the alignment is only the result of observational bias (see, for example, Kavelaars et al., 2020). This means that objects were found along the alignment because that is where surveys searched for objects. While this is true for some surveys, there are others, such as a large survey being carried out by Sheppard and Trujillo (Sheppard et al., 2016) to search for Planet X and ETNOs, that search all values of ω and still have only found extreme objects along the direction of alignment. As more distant objects have been discovered, it has also been claimed that the alignment no longer appears, however, these claims typically involve the inclusion of many objects that are significantly less extreme[70] (e.g., in the extreme scattering population) (see, for example, Napier et al., 2021). In Figure 1.21 we show a plan view of the orbits of the known ETNOs and IOCs. We highlight several different extremity cutoffs for these objects, all of which are more strict than are used in observational bias arguments, and note that the tighter the limit placed on the selection criteria (with higher q being the main factor), the more aligned the population is overall. Hence, it is important to only consider objects that are primarily influenced by Planet X, and not by Neptune, when determining which objects may be aligned with such a planet.

In addition to the alignment in ω, other evidence has been shown to support the existence of Planet X. The most relevant for this work is the gap in perihelion q (shown in Figure 1.19), noted by Trujillo & Sheppard (2014) along with the discovery of 2012 VP_{113}. In 2019, Sheppard and Trujillo discovered another IOC (Sheppard et al., 2019), which has since been named Leleākūhonua. This third IOC strongly cements the existence of the gap along with implications for Planet X are the principle subject of Chapter 2. Since the gap is in perihelion rather than physical space, an easy way to visualize it is by plotting the perihelia of

[69] Brown and Batygin refer to this planet as Planet Nine. Brown also refers to himself as “the Pluto killer,” since he lead the team that discovered Eris, prompting the IAU to revisit the formal definition of a planet.

[70] It is difficult to define how extreme an object is, or, at least, it has not yet been done formally in the literature. Here, more extreme means higher e, a, and/or q.

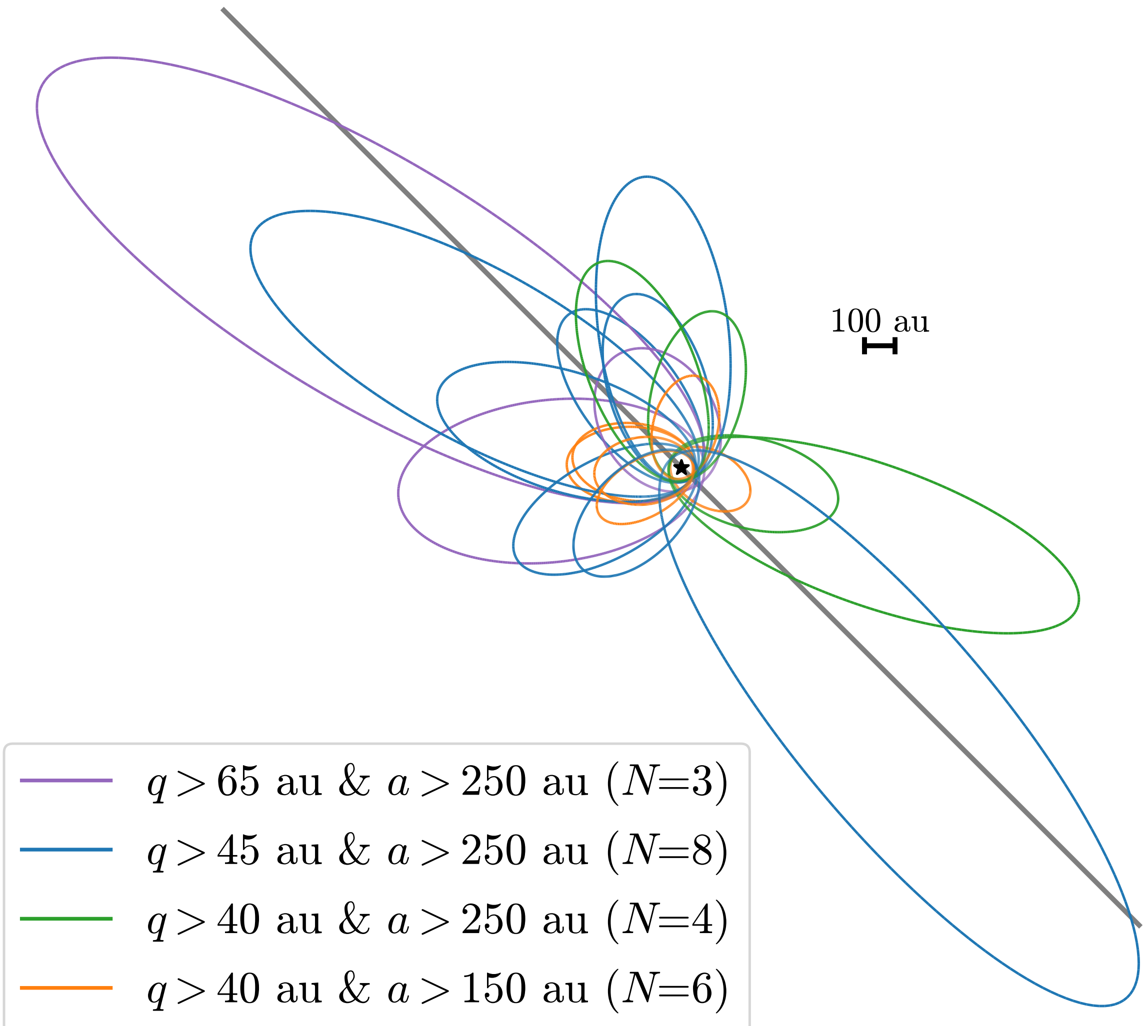


Figure 1.21 Plan view of IOC (purple) and ETNO (blue, green, and orange) orbits. The legend details progressively less strict cutoffs for the ETNOs; blue is a tighter limit than green and orange is a more lenient definition. The general direction of the observed orbital alignment is highlighted as a gray line. Overall, as more strict limits are used on the definition of what qualifies as an ETNO, the strength of the alignment increases.

the ETNOs and IOCs as is shown in Figure 1.22. Here the ETNOs all have a q interior to the gap, which is represented by the shaded region, and the q of the IOCs is always exterior to gap. One of the primary evidences that the perihelion gap is a real feature of the outer solar system is that more distant objects are fainter (brightness falls off as heliocentric distance to the fourth power) and, therefore, objects in the gap would be significantly (a factor of $\sim$5) easier to detect than IOCs.

1.3.4 Exoplanets

Around the same time that the first Kuiper Belt objects were being discovered, another major astronomical discovery was taking place. In 1995, the first planet orbiting another star, 51 Pegasi b, was discovered by a team of Swiss astronomers (Mayor & Queloz, 1995). These planets are called exoplanets and over 5,000 of them have been discovered so far, with nearly 6,000 more candidate exoplanets awaiting follow up and confirmation[71]. Exoplanets are discovered using a variety of methods and we will summarize three of the most common search techniques here.

By far the most prolific technique thus far in the search for exoplanets is the transit method. The basic concept behind the transit method is that as an exoplanet orbits a star, if it passes between the Earth and the star, then the light from the star will get slightly dimmer while the exoplanet is eclipsing part of it (e.g., blocking some of the light from the star from reaching the telescope). This method only works for exoplanets that are on orbits that allow the exoplanet to pass in front of the star from the perspective of Earth. The transit method also heavily favors planets that have very short orbital periods (days to weeks) and, consequently, are very close to their host stars. Despite these limitations, the transit method is relatively easy to employ as the only requirement is monitoring the brightness of stars and searching for a characteristic drop in brightness. Because this method favors exoplanets with short orbital periods, one of the more common types of exoplanets discovered via transit are called hot Jupiters. These exoplanets are typically at least as massive as Jupiter and, because they orbit so close to their host stars, they have very high temperatures. These high temperatures causes the atmospheres of hot Jupiters to become inflated, increasing their radii beyond typical values for planets of similar masses (Fortney et al., 2021).

Another effective detection technique for exoplanets is the radial velocity method. This is the method with which the first exoplanet (51 peg b) was discovered (Mayor & Queloz, 1995). The radial velocity method relies on spectroscopy (splitting light into its many components/colors in order to infer composition[72]) to detect exoplanets. This method is also based on a phenomenon called Doppler shift, when objects move towards an observer at high speeds, that object appears bluer and, conversely, objects moving quickly away

[71]For the most up to date statistics see NASA's Exoplanet Archive: `https://exoplanetarchive.ipac.caltech.edu/`

[72]Each element has a unique spectral signature, like a fingerprint. By measuring the spectrum of an object, information about its composition can be inferred.

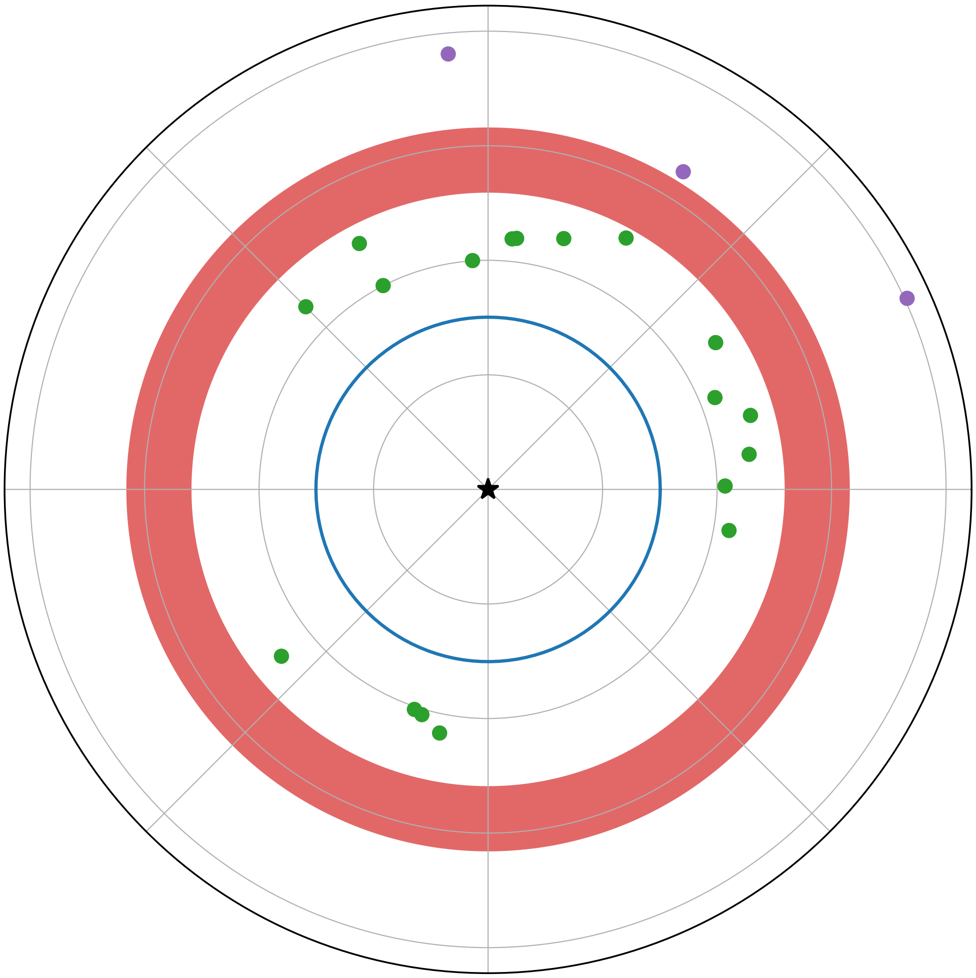


Figure 1.22 Perihelia of the ETNOs (green) and IOCs (purple). The orbit of Neptune (30 au) is shown in blue for scale. The perihelion gap (50 au – 65 au) is shown in red with the ETNOs coming to perihelion interior to this gap and the IOCs coming to perihelion exterior to the gap. Radial grid lines are 20 au apart. Because objects on distant orbits, such as ETNOs and IOCs, are most likely to be discovered near perihelion, where they are brightest, the uniform case would predict the existence of more objects within the gap than beyond it. However, the presence of the IOCs strongly suggests that the distribution of objects is not uniform. Hence, a mechanism (e.g., interactions with Planet X) is needed to explain the gap.

from an observer appear redder (Doppler, 1903). If a giant exoplanet is orbiting a star, it pulls the star closer to the Earth when it is on the near side of the star and pulls it further away from the Earth when the exoplanet is on the far side of the star. Measuring these pushes and pulls using spectra of the star reveals a shift to bluer wavelengths (colors) when the exoplanet is pulling the star towards the Earth. Reciprocally, the spectra are shifted to redder wavelengths when the exoplanet is pulling the star away from the Earth. This method is effective for Earth-sized exoplanets as well Jupiter-sized bodies and larger, but the signal is stronger and easier to detect the larger the exoplanet is. Exoplanets in close proximity to their host star are also more easily detected than distant ones using this method, and the closer a planet is to having an orbit that would cross in front of the star (from the perspective of Earth)[73], the more likely it is that the exoplanet will be detectable using its radial velocity. Effectively, exoplanets that are detectable using the transit method are also the best candidates for detection using the radial velocity method.

A third, and perhaps more straightforward, method for detecting exoplanets is through direct imaging. In this method, the area surrounding a star is observed at multiple times and exoplanets orbiting the star are directly imaged (the light coming from the star is reflected off the exoplanet and observed via telescope). The direct detection of exoplanets via this method is excellent for studying the physical properties of these planets. The direct imaging method is similar to the method Galileo used to discover the moons of Jupiter. In order for Galileo to observe these moons, however, the increased angular resolving power (ability to separate two closely spaced objects) of a telescope was required in order to be able to distinguish the moons from Jupiter. The same is true for finding exoplanets using the direct imaging method; high resolving power is required. In contrast with the other methods mentioned, direct imaging favors planets on large orbits that are widely separated from their host star. Exoplanets orbiting within a minimum angle[74] of their host star will be lost in the glare and be undetectable. Although direct imaging has resulted in relatively few ($\sim$300)[75] exoplanets being discovered to date, next generation space telescopes, such as HabEx and LUVOIR (see Chapter 3), will dramatically increase the number of exoplanets discovered using this method. For a review of exoplanet detection methods, see Wright & Gaudi (2013).

One of the most longstanding questions in astronomy is whether or not we are alone in the universe; *Is there life on other worlds?* In the search for the answer to this question, one of the most relevant concepts regarding exoplanets is the "habitable zone". A planet orbits in the habitable zone of a star if it is at the proper surface temperature to support liquid water on the surface of the planet (Kasting et al., 1993). Earth is the prime example of residing within the habitable zone, and rightfully so, since the only known examples

[73]If the exoplanet can eclipse the star, when it does it is pulling the star directly towards the Earth. If it is not directly in line, it only pulls the star partially towards the Earth and partially perpendicular to the Earth, across the sky.

[74]This angle is dependent on the brightness of the star, the brightness of the planet, and the resolving power of the telescope.

[75]https://exoplanetarchive.ipac.caltech.edu/

of life are here on Earth. Mars is near the outer edge of the habitable zone for the solar system and is, therefore, too cold. Venus, on the other hand, is interior to the inner boundary of the habitable zone and too hot to support liquid water on its surface[76] (Kopparapu et al., 2013). As all known examples of life require water to survive, water is the key element in determining the habitable zone around a star and is also an important indicator that an exoplanet may be suitable for life (Kasting et al., 1993).

1.4 Overview of Observing

Perhaps the most quintessential aspect of astronomy is observing the universe. Since the time of Galileo, telescopes have dramatically changed the way the heavens are viewed. This invention has enabled countless discoveries and, as new more powerful telescopes are built, they will pave the way for countless more. Throughout the course of this work, several professional research telescopes have been used to collect data on a variety of solar system objects; these telescopes are listed in Table 1.3. Telescopes used in this work are all ground-based, but space-based telescopes also provide essential insights. Two key advantages of space based telescopes are that they are free from the influence of Earth's atmosphere, which makes observations more difficult (and even unfeasible at certain wavelengths) and that they can more easily point away from the Sun, effectively doubling the amount of time they can spend observing (since most ground based optical telescopes cannot observe during the day[77]). However, space telescopes are typically vastly more expensive and much more difficult to repair or upgrade than ground based telescopes, hence, both ground-based and space-based telescopes play important roles in the study of the universe.

Recently, the James Webb Space Telescope (JWST) was launched and has just begun science observations. Its advanced optical and digital systems combined with its large (6.5 m) mirror will revolutionize astronomy by observing fainter and more distant objects than ever before and in greater detail. For the past several years, astronomers have been looking to the next step beyond JWST in space telescope technology. Since discovering exoplanets that could potentially harbor life (habitable worlds) is one of the primary science goals in astronomy going forward (National Academies of Sciences, 2021), two proposed next-generation space telescopes, the Habitable Exoplanet Observatory (HabEx) and the Large UV/Optical/IR Surveyor (LUVOIR), have been designed to dramatically improve on current capacity to detect exoplanets via the direct imaging method (see Section 1.3.4).

Advancement in exoplanet direct imaging capabilities comes through two major technologies, (1) the coronagraph and (2) for space telescopes, a secondary companion spacecraft called a starshade. A corona-

[76]The habitable zone is sometimes called the Goldilocks zone, since planets in the habitable zone are not too hot and not too cold, but just right for potentially supporting life.

[77]Radio telescopes and gravitational wave telescopes can still operate during the day.

Table 1.3. Telescopes Used

Telescope	Acronym	Diameter	Location
Large Binocular Telescope	LBT	2x8.4 m[a]	Mt. Graham, AZ
Gemini South Telescope	GS	8.1 m	Cerro Pachón, Chile
Walter Baade Magellan Telescope	Baade	6.5 m	Las Campanas, Chile
Lowell Discovery Telescope	LDT	4.3 m	Happy Jack, AZ
Víctor Blanco Telescope	DECam[b]	4 m	Cerro Tololo, Chile
Vatican Advanced Technology Telescope	VATT	1.8 m	Mt. Graham, AZ

Note. — The diameter of a telescope refers to the size of the primary mirror.
(a) The LBT has two 8.4 m mirrors that work in tandem making it effectively the largest optical telescope in the world.
(b) The Blanco Telescope is usually called DECam after the Dark Energy Camera, its signature instrument.

graph is a sophisticated instrument mounted within a telescope that blocks out the glare of light surrounding a star, called a stellar corona[78]. This allows a telescope using a coronagraph to observe exoplanets much closer to the star than they could otherwise detect (see, for example, Guyon et al., 2005). A starshade has a similar purpose to a coronagraph, but, rather than being mounted to the telescope, a starshade is a separate spacecraft that positions itself tens of thousands of kilometers away from the telescope. The starshade then maneuvers to position itself exactly between the telescope and the target star. This allows for extremely high precision measurements of the area around that star (even better than a coronagraph), although lining up each new target star is time consuming (days to weeks; see, e.g., Soto et al., 2019). HabEx and LUVOIR will nominally use both a coronagraph and a starshade (as they are good complements to one another) to search for habitable worlds around nearby stars. Additionally, coronagraphs are used in conjunction with adaptive optics systems — adaptive optics refers to fine adjustments to the telescope mirror, using actuators under the mirror, in order to correct for atmospheric variation in real time, thus, adaptive optics can greatly enhance image quality (e.g., as in Marois et al., 2008) — on ground-based telescopes for direct imaging of exoplanets and this technology is expected to continue to improve in the coming years (e.g., Carter et al., 2021).

The distance to nearby stars is often measured in a unit of distance called a parsec (abbreviated pc). For example, the closest star to our Sun, Proxima Centauri, is ~ 1.3 pc away[79]. Stars are considered to be nearby if they are within roughly 10 pc of the Sun and systems orbiting these stars will be primary targets for missions such as HabEx and LUVOIR in their search for habitable worlds (Gaudi et al., 2020; The LUVOIR Team, 2019). This 10 pc distance is close enough for both sufficient angular resolution (being

[78] *Corona* means crown in Latin.
[79] A parsec is equal to ~ 3.26 light-years or 206,265 au.

able to separate targets that are close together as distinct objects) and a reasonable Signal to Noise Ratio (SNR; e.g., 10). The Signal to Noise Ratio is a measure of how bright a source (star, planet, etc.) is in comparison with the background, which is designated as noise and consists mostly of scattered light from other stars, the atmosphere (for ground-based telescopes), and even inside the telescope. For example, an SNR of 10 would mean that the source is ten times brighter than the background noise.

Having a good SNR is important not just for directly imaged exoplanets, but also for measuring the brightness of any astronomical object. In astronomy, brightness is typically measured in magnitudes[80], where larger magnitudes are fainter than smaller magnitudes. The magnitude scale is set so that 5 magnitudes difference is a factor of 100 times brightness difference and a 1 magnitude difference is approximately a 2.5 times difference in brightness. For example, a magnitude 1 star would be 100 times brighter than a magnitude 6 star. An object is considered to be visible to the naked eye if it is 6th magnitude or brighter[81] (see reference to Hipparchus in work by Ptolemy, English translation: Ptolemy & Toomer, 1984).

The brightness of an object measured using a telescope is called the instrumental magnitude. A series of calibrations are then applied to this value to reduce the noise caused by the telescope, the camera taking the image, and the atmosphere. The result is referred to as the apparent magnitude, which is a measure of how bright the target actually appears. For small solar system bodies, the apparent magnitude is typically represented by a letter (e.g., R, V, or I) based on what filter was used in the telescope to take the image. Filters are specially designed glass that only allow a certain wavelength range of light to pass through them. For example, the R filter[82] allows mostly red light to pass through it to the telescope. Filters are standardized so measurements of the same object with the same filter will result in roughly the same apparent magnitude for that object (unless the objects is changing color). By using a sequence of filters, the color of an object can be measured by calculating the difference in brightness of the object between each different filter. One of the most commonly used filters for small solar system bodies is the V filter[83] (Johnson et al., 1966), hence, most apparent magnitude values reported in this work are given in the form $V =$, which stands for the magnitude in the V filter[84].

Another important magnitude often used for asteroids, Centaurs, and TNOs is the absolute magnitude H. The absolute magnitude is a way of comparing objects independently of their positions and, in essence, asks the question *How bright would this object be if it were 1 au away from both the Earth and the Sun while at opposition?* This configuration is physically impossible, opposition is where an object is on the

[80] Although some sub-disciplines of astronomy prefer flux which is a more direct measurement of brightness.

[81] In most populated areas, light pollution makes objects in the night sky much more difficult to see and even 3rd magnitude stars may be too faint to observe by eye in a big city.

[82] R stands for red.

[83] V stands for visual and primarily lets green light through.

[84] m_V is also commonly used to represent the magnitude of an object in the V filter.

opposite side of the Earth from the Sun, but H is still a useful metric for theoretically comparing small bodies throughout the solar system. Calculating H depends on two major factors, the brightness of the object and its phase angle, which is a measure of how illuminated the object should be from the point of view of the observer based on the angle between the Sun, the object, and the observer. The phase angle is analogous to time of day on the part of the object being observed. One of the most useful applications of H is the inference of physical properties of an object. In particular, two physical parameters, the size of the object and its albedo, can be calculated using H (see Section 2.2 and Equation (7) of Harris & Harris, 1997). Albedo p_V is a measure of how reflective a surface is and ranges in values from 0 to 1, where 0 is the blackest of black that absorbs all light and 1 is a perfect reflector[85]. The absolute magnitude can also be used to convert to apparent magnitude (and vice versa) if the heliocentric distance of the object and its distance from the Earth are known (see Section 2.2).

On a more procedural note, one interesting aspect of small solar system bodies is the process by which they are named and numbered[86]. When a new object is discovered, the discoverers submit the times, coordinates, magnitudes, and filters, and telescopes used in the discovery to the IAU's Minor Planet Center (MPC). The MPC is managed through the Harvard Center for Astrophysics and is charged with maintaining a detailed database of all small bodies in the solar system. Once the MPC gets the necessary information, they calculate an orbit for the object and give it a provisional designation[87] and announce its discovery along with anyone who assisted in the observations. The provisional designation is a four digit year followed by two letters and, in many cases, numbers, e.g. 2012 VP_{113}. The year is the year in which the object was first discovered[88]; the first letter is the half month[89] the object was discovered in — A for the first half of January, B for the next half, C for the first half of February and so on, excluding I and Z since only 24 letters are needed[90]. The second letter and the numbers designate the order in which objects are discovered during that period (excluding I, since it looks like l or 1 in some fonts), for example, A is the first, B is the second, Z is the 25th, then A1 is the 26th, and so on. It follows that 2012 VP_{113} was the 2,841st object discovered between November 1-15 of 2012[91].

[85] Technically values above 1 are allowed, but that would mean reflecting back more light than shines on the surface in the first place.

[86] Exoplanets are named in a separate process where they are simply given the name of the star they orbit and a letter indicating the order they were discovered in starting with b (a is the star itself).

[87] I like to call these licence plate numbers.

[88] 2012 VP_{113} was first discovered in 2012 although it was not announced until 2014, because more observations were needed in order to calculate an accurate orbit.

[89] I suspect that fortnights (2 week periods) were originally used and all 26 letters were included, but half months were easier to use in practice.

[90] Y and Z are the natural choice to exclude, but presumably because I looks ambiguously like 1 or l in some fonts, Y was swapped out with I.

[91] Based on the order it was reported in, so ETNOs that take a long time to get follow-up observations on to determine their orbit usually end up with big numbers since lots of easier to observe asteroids are reported more quickly and get all the smaller numbers.

Once an object has a provisional designation and its orbit is well-determined, the MPC assigns it a number[92]. This number is the order in which objects are found to have a sufficiently well-determined orbit[93] and for objects discovered early on, it also coincides with their order of discovery, e.g., Ceres, the first asteroid to be discovered is designated as 1 Ceres (Gould, 1852). Pluto, on the other hand, was not given a minor planet number until it was demoted so it is designated as 134340 Pluto[94]. Once an object has an official number, the discoverer can submit a proposal to name that object to the MPC, which sends to proposals to the IAU's Committee on Small Body Nomenclature, which decides whether to accept the name. Some types of objects require that a naming pattern is followed and some types of names are prohibited in general. A name can be in honor of a person, for example, asteroids 12101 Trujillo, 8063 Cristinathomas, and 39184 Willgrundy are all named after members of my graduate committee[95].

1.5 Statistical Methods

One of the most effective ways to search for quantifiable trends in data is through the use of statistics. The essence of statistics is determining the probability or likelihood of a given outcome. There are several probability functions (mathematical expressions) that are commonly used to model the behavior of astrophysical systems; and modeling is one of the most efficient ways to develop an understanding of the underlying mechanisms at work in these systems. The most frequently used probability function is the Gaussian function[96]. This is a standard symmetrical "bell curve" often seen when determining the probability of a given outcome. For example, a Gaussian can be used to estimate the probability of rolling a given number on a pair (or any number > 2) of dice. Gaussian curves can also represent the light from a star captured on a camera through a telescope[97], the area of uncertainty on a measurement, the distribution of an orbital element among a population of asteroids, and a host of other applications[98]. Figure 1.23 shows a Gaussian distribution along with two other distributions used in this work, the Lorentzian distribution and the Poisson distribution.

An essential aspect of creating a statistical model to represent a real system is verifying or at least comparing the model to the real system in order to determine if the model is indeed a good representation of the system. In this work, we utilize two main techniques for comparing statistical models with the systems

[92]The MPC has metrics that they monitor on how well known each object's orbit is. Once enough observations have been taken to bring these metrics below their threshold value, then they give the object a number.

[93]For moons, the number is the order of discovery or, for the first few moons in a system, they are typically ordered based on their distances from the planet and Roman numerals are used, e.g., Earth I is our moon, Luna.

[94]I like to call the official number of minor planets their phone number to go along with their licence plate number (the provisional designation).

[95]Basic rules for naming a numbered object can be found at https://minorplanetcenter.net/iau/info/Astrometry.html#name

[96]Named after the same Gauss that helped to rediscover Ceres.

[97]This actually requires a 2-dimensional Gaussian function.

[98]"Most things are more or less Gaussian." - Chad Trujillo

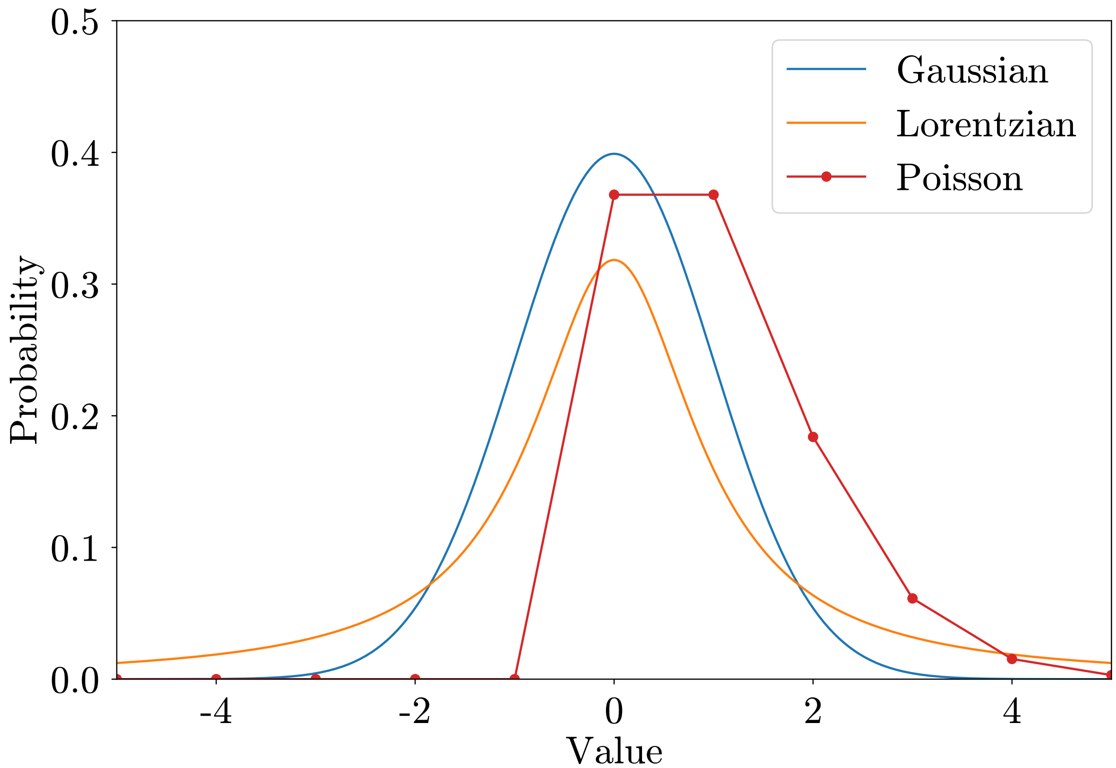


Figure 1.23 Example probability density functions, Gaussian (blue) and Lorentzian (orange), and Poisson (red) probability mass function. The Gaussian is the standard “bell curve” used widely across many disciplines. The Lorentzian is similar in many ways to the Gaussian with broader wings and a slightly sharper and shorter peak. The Poisson curve takes on discrete (integer) values, rather than being continuous, like the other two curves shown here, and is asymmetric.

we analyze, (1) statistical tests and (2) Monte Carlo methods. Statistical tests are techniques to numerically determine how similar two distributions are. For example, the Kolmogorov-Smirnov (K-S) test compares two distributions by finding where the sum of each distribution's values[99] are most different from each other (see Bonamente, 2017, p. 91). Statistical tests are fairly straightforward to implement and are widely used throughout the scientific community. Monte Carlo methods are ideal for systems with unknown variables (such as orbit determination). They rely on large numbers of random or pseudo-random (i.e., preferentially random, like a loaded die) samples to approximate real systems.

The basic Monte Carlo[100] method involves randomly generating points or samples. Each point is compared to a mathematical expression of the system in question. The ratio of points that do or do not satisfy a condition related to this expression is then used to approximate the state of the system. For example, the value of π could be approximated using a Monte Carlo simulation by sampling random points within some region that contains a circle. The ratio of points that fall inside vs. outside the circle (in conjunction with the area of the region and the radius of the circle) will form an estimate for the value of π. The larger the number of samples used in this method, the more accurate the approximate will be (see Metropolis (1987) for further discussion of the Monte Carlo approach).

Building on the foundation of the Monte Carlo method, Markov Chain Monte Carlo (MCMC) techniques excel at analyzing multidimensional systems, where efficiency in sampling is vital due to the large number of possible samples. In MCMC, the goal is to obtain the best possible match between model and data, but, rather than randomly and independently selecting each individual point, random points are connected through Markov chains, which are often referred to as "walkers". To conceptualize the MCMC process, one helpful analogy for MCMC is an expedition to find the highest peak in a mountain range. The walkers represent explorers who are all working together to find the highest peak, which represents the best fit or match between model and data. Typically, walkers all start near each other surrounding an initial estimate for the best-fit (highest peak).

Once walkers are assigned a starting point, they are all compared with the model using a probability function, which determines two things, (1) if the walker is within the priors (bounds) of the simulation[101] and (2) the probability score the walker should get based on how closely it matches the model to the data. Within the analogy, each explorer (walker) has a device that can measure their position and elevation, which correspond to the modeled parameters for the MCMC simulation, but there is a think fog that prevents them from visually inferring other information. The explorers know the coordinates of the edges of the mountain

[99] The location where the cumulative density functions of the two distributions have the largest separation.

[100] Monte Carlo methods were named after a casino in Monaco; a reference to the numerical connections between statistical methods and gambling (Metropolis, 1987).

[101] This can either be a hard "out of bounds" line, or a distribution that favors certain values over others.

range and want to stay within this region (priors). The elevation the explorers measure is analogous to the probability score of the walkers, with higher elevations corresponding to higher probabilities.

Next, the walkers all broadcast their scores to each other and take what is called a step in a pseudo-random direction. A step is simply changing a walker's position within the parameter space by a random, but small amount (so it could be a small step, or a tiny step, or anywhere in between). Steps are pseudo-random rather than completely random because each walker wants to improve its score, so it picks a direction, that is still random, but weighted towards other walkers that have higher scores[102]. In the analogy, a step could represent five minutes. The explorers all pick random directions to walk in with a divide-and-conquer approach to finding the peak. After walking for five minutes (one MCMC step) they all stop and send their positions and elevations to each other. Then, each explorer picks a new random direction to walk in, but they have some bias towards directions that they know are higher elevations (because other explorers are/were there). It is still useful for walkers to occasionally move towards lower scores as this helps to explore the area and find paths out of local maxima; otherwise the explorers might end up on the wrong peak.

This process is repeated for many steps — typically thousands or even millions of steps — until the walkers converge around a particular combination of parameters that provide a good match of the model to data (near the top of the mountain). The optimal answer is then typically taken to be either the average of these walkers after they have converged or the best matching walker's values. This would be akin to the explorers finding the general area near the top of the mountain, then taking either the average of their positions or the single highest measured position to be the peak. Both of these methods can be statistically informative and both are reported in this work where relevant (see MacKay, 2003; Foreman-Mackey et al., 2013; Bonamente, 2017, and similar works for further discussion of MCMC methods).

Typically, after completing an MCMC simulation, the results are visualized using a "corner plot" (sometimes called a "triangle plot" because of its distinctive shape). This type of figure shows the walker positions for each pair of variables in the simulation. It usually highlights the mean and/or peak values for each variable as well as uncertainties on these values (which are calculated based on the Gaussian curve). A particularly useful feature of a corner plot is that pairs of variables that strongly affect one another (correlated variables) are easy to notice as they will make diagonal lines through the corresponding section of the plot; this can provide additional insight into the system as a whole (Foreman-Mackey, 2016). A corner plot is used in Section 2.7.

Two additional plotting techniques that are particularly useful for interpreting statistical data are "heatmaps" and rolling averages. Heatmaps, or density plots, are a method of presenting 3-dimensional

[102] Think of rolling a six-sided die, but instead of the standard faces it has 2,3,4,5,6,6. This is similar to how a weighted random works, walkers are more likely to move towards higher scores, and less likely to move towards lower scores.

data using only 2 axes and color. For example, the temperature of a surface could be represented by plotting color on that surface (typically with hot regions being red and cold regions being blue) as could be seen with an infrared camera. In Chapter 2, heatmaps are used to show the duration over which particles in a simulation have certain combinations of orbital elements. A rolling average is a way to smooth out data that has a high level of variation. For this type of plot, each data point is the average of neighboring values, which effectively smooths the resulting distribution[103].

The main concern of statistics is how likely an event is to happen given certain conditions. This is often called likelihood or probability, with probability usually being on a scale from 0 to 1 or, in terms of percentage, from 0% to 100%. The process of changing a range from its values to a scale from 0 to 1 (min to max, and effectively identical to taking a percentage) is called normalization and is quite common in statistics and other fields that use statistics. Normalization allows values to be compared more easily, a fundamental technique in many forms of statistical analyses.

1.6 Chapter Overview

The remainder of this dissertation is broken up into 4 additional chapters. In Chapter 2, we discuss the outer solar system perihelion gap and its connection to the hypothetical distant giant planet, Planet X. We begin by outlining some of the recent history associated with the search for Planet X, highlighting the discoveries of the IOCs and their implications for the search. Next we define the gap and describe a suite of observational simulations. In these simulations, we make several synthetic (not directly corresponding to real objects) populations of ETNOs and IOCs in an attempt to match a population distribution with the real ETNOs/IOCs. No unimodal (using a single probability function) distributions accurately represent the real ETNOs and IOCs. However, two Gaussian distributions added together, with parameter fits derived using MCMC, are an excellent match to the real ETNOs/IOCs, which indicates that the perihelion gap is real and not simply caused by observational bias.

Next, we carry out a large suite of orbital dynamics simulations where we examine the effect Planet X would have on a group of synthetic TNOs. We find that without Planet X, the ETNOs and IOCs do not form from the scattered disk of the Kuiper belt. However, with Planet X, the ETNOs, IOCs and the perihelion gap can all form through secular resonances between the TNOs (which become ETNOs and IOCs) and Planet X. One particularly interesting finding from this chapter is that even though there is a gap, ETNOs and IOCs travel through this gap over time periods of hundreds of thousands of years. The secular resonances between Planet X and ETNOs/IOCs cause the ETNOs/IOCs to spend more time, on average, on either side of the

[103] This technique is commonly used to help predict stock market trends.

gap (up to a factor of roughly 5 times longer) than they spend in the gap, thus, making the gap appear empty because of the small number (roughly 20) of ETNOs and IOCs that have been discovered to date.

Chapter 3 focuses on a new method for finding the best revisit time for observing an exoplanet using direct imaging, thus, maximizing the constraint placed on its orbit. We provide background on next generation space telescopes, which will dramatically improve exoplanet direct imaging capabilities. Additionally, we discuss previously used and proposed techniques for determining directly imaged exoplanet orbits. We also describe three separate methods for computing a best revisit time for subsequent observations of a previously directly imaged exoplanet system, all of which utilize MCMC. We show these methods to be roughly comparable and that they all reduce the uncertainty in the orbit of the exoplanet in fewer observations than previously proposed methods.

Returning to the solar system, Chapter 4 focuses on the discovery of a new season or epoch of cometary activity on Quasi-Hilda (323137) 2003 BM_{80} (282P). This activity was detected through the NASA Partner Citizen Science program *Active Asteroids*. We then provide details on a suite of dynamical simulations which we use to examine the past and future orbital characteristics of 282P. We find that in the past 250 years, 282P migrated to its current Quasi-Hilda orbit from either the Centaur or Jupiter Family Comet regions due to several close encounters with Jupiter and/or Saturn which drastically alter its orbit. Future close encounters with Jupiter will cause 282P to transition from the Jupiter Family Comet region to the outer main asteroid belt (both of which overlap the Quasi-Hilda region) and back roughly 10 times in the next 250 years. In the next 350 years, 282P is expected to orbit solidly within the Jupiter Family Comet region, however there is an approximately 15% probability that it will instead become an active asteroid. This possibility hints at a potential pathway for Jupiter Family Comets to become active asteroids via migration through the Quasi-Hilda region.

In Chapter 5, we provide summaries of Chapters 2–4 along with discussion on their broader implications for the astronomy community and the general public. Additionally, we provide Tables 1.4 and 1.5 here with list of acronyms and variables, respectively, used throughout this dissertation.

Table 1.4. Acronyms Used

Acronym	Name	Description
AJ	Astronomical Journal	One of the foremost astronomy journals
au	Astronomical Unit	Average distance between Earth and the Sun
AZ	Arizona	State abbreviation
BYU	Brigham Young University	My undergraduate institution
CC	Cold Classical	A Kuiper belt object with low eccentricity and inclination
CPU	Central Processing Unit	The part of a computer that performs calculations and executes coded instructions
DART	Double Asteroid Redirect Test	A NASA mission focused on altering the orbit of an asteroid as a planetary defense proof of concept
DECam	Dark Energy Camera	The signature instrument of the 4 m Blanco Telescope atop Cerro Tololo, Chile
ETNO	Extreme Trans-Neptunian Object	An object orbiting beyond the orbit of Neptune with high eccentricity, semi-major axis, and perihelion
HabEx	Habitable Exoplanet Observatory	Future exoplanet-focused space telescope concept
IAU	International Astronomical Union	International group of astronomers providing guidelines on naming conventions
IOC	Inner Oort Cloud object	A Sedna-like outer solar system object with extremely high eccentricity, semi-major axis, and perihelion
IR	Infrared	Wavelengths of light just longer than the visible spectrum
JD	Julian Date	A unit of time
JFC	Jupiter Family Comet	A population of short period comets that orbit near the semi-major axis of Jupiter
JPL	Jet Propulsion Laboratory	Branch of NASA focused on robotic space missions
JWST	James Webb Space Telescope	Recently launched 6.5 m space telescope
KBO	Kuiper Belt Object	An object from an asteroid-belt-like disk beyond Neptune
K-S	Kolmogorov-Smirnov	Authors of a statistical test for determining similarity between two distributions
LBT	Large Binocular Telescope	Twin 8.4 m telescope on Mt. Graham, Arizona
LDT	Lowell Discovery Telescope	4.3 m telescope in Happy Jack, Arizona, formerly the Discovery Channel Telescope
LUVOIR	Large UV/Optical/IR Surveyor	Future exoplanet-focused space telescope concept
MCMC	Markov Chain Monte Carlo	A statistical approach, similar to machine learning, for finding optimum parameters to fit a model to data
MMR	Mean Motion Resonance	Orbital periods that are integer multiples of each other
MPC	Minor Planet Center	Hosts a database of all small bodies in the solar system
NASA	National Aeronautics and Space Administration	United States government agency focused on space exploration and research
NAU	Northern Arizona University	Where I am getting my PhD
NEA	Near Earth Asteroid	Asteroids with orbits that bring them close to Earth's orbit
NEO	Near Earth Object	Same as NEA, but could be a comet as well
OFTI	Orbits For The Impatient	An orbit fitting algorithm for directly imaged exoplanets
PB	Petabyte	One quadrillion (10^{15}) bytes of computer memory
PhD	*Philosophiae Doctor*	Doctor of Philosophy
PI	Principle Investigator	The lead astronomer on a telescope observing program
QH	Quasi-Hilda	Object near, but not in the 3:2 interior MMR with Jupiter
SDO	Scattered Disk Object	A KBO with moderate eccentricity and/or inclination
SNR	Signal-to-Noise Ratio	The amount of signal from a source in comparison to the background noise
TNO	Trans-Neptunian Object	Objects located beyond the orbit of Neptune
UV	Ultraviolet	Wavelengths of light just shorter than the visible spectrum
VATT	Vatican Advanced Technology Telescope	1.8 m telescope on Mt. Graham, Arizona

Table 1.5. Variables Used

Variable	Name	Description
a	Semi-major Axis	Average orbital distance
a_g	Gaussian Ratio	Fraction of objects belonging to one portion of a statistical fit
b	Ecliptic Latitude	Perpendicular distance of a body from the plane of the solar system
C	Opposition Phase Angle Constant	For calculating illumination factor for H
D	Telescope Diameter	The width of the primary mirror of a telescope
d	Distance	Positional separation between two objects
E	Eccentric Anomaly	Conversion factor between M and f of an orbit
e	Eccentricity	Parameter describing the ellipticity of an orbit
F	Force	Causes a change in momentum over a time
f	True Anomaly	The angle from ω of a body in orbit
G	Universal Gravitational Constant	Used to convert gravity parameters into standard units
$\mathcal{G}$	Gaussian	Standard "bell curve"
g	Lorentzian Power	Shape defining parameter for inclination distribution Lorentzian curve
H	Absolute Magnitude	V of an object at 1 au from the Earth and Sun while at opposition
I	Lorentzian Half Inclination	Parameter defining the shape of a Lorentzian curve
i	Inclination	Angle of orbital tilt away from a reference plane
$\mathcal{L}$	Lorentzian	Statistical curve similar to a Gaussian
$\mathcal{L}$	Likelihood	Probability of a set of parameters matching model to data
M	Mean Anomaly	Fraction of orbital period completed times 360°
M, m	Mass	the amount of matter in a body
$M_\oplus$	Earth Masses	A unit for measuring the mass of planets
P	Orbital period	Time taken to complete a full orbit
P	Poisson Probability	A statistical measure for distribution comparison
p_V	Geometric Albedo in V band	Brightness ratio between measurement and an ideal flat disk reflector
Q	Aphelion	The most distant point in an orbit from the Sun
q	Perihelion	The closest point in an orbit to the Sun
R	Heliocentric Distance	Distance between a body and the center of the Sun
$\mathbb{R}$	Real Numbers	Any non-imaginary number
r	Distance	Positional separation between two objects
r	Radius	Average distance from the center of a body to its surface
r	Diffraction Limit	The minimum angular separation between two bodies in order for them to appear separate in observations
r_H	Hill Radius	Metric of satellite orbital stability around a body
T_{J}	Tisserand Parameter with respect to Jupiter	Semi-constant measure of Jupiter's affect on a body's orbit
t	Time	Ordering of the sequence of all events
V	Visual Magnitude	Brightness of an object in the V Johnson-Kron-Cousins filter
v	Apparent Sky Velocity at Opposition	How fast a body appears to move when on the opposite side of the Earth from the Sun
V_X, V_Y, V_Z	Cartesian Velocities	Directional speeds along fixed axes in 3 dimensions
X, Y, Z	Cartesian Coordinates	Distance from a reference point along fixed axes in 3 dimensions
θ	Angular Error	Standard Gaussian error represented as an angle
λ	Wavelength	Peak-to-peak distance between light waves of a given frequency
$\hat{\lambda}_i$	Poisson Maximum Likelihood	Mean value of samples in a given bin
μ	Gaussian Mean	Center point of a "bell curve"
ν	Secular Resonance	Denotes the body the resonance is with e.g., ν_6 would be a resonance with Saturn
ρ	Momentum	Mass times velocity
σ	Standard Deviation	Portion of a Gaussian curve representing uncertainty on a value
χ^2	Chi-Squared	Used for determining how well a model matches data
Ω	Longitude of Ascending Node	Describes the rotation of an orbit away from a reference line about an axis perpendicular to a reference plane
ω	Argument of Perihelion	Angular direction of perihelion in an orbit relative to the ascending node
ϖ	Longitude of Perihelion	$\Omega + \omega$

Chapter 2

The Perihelion Gap and Planet X

This chapter contains my published work on the subject in its entirety. The formatting has been changed to comply with University requirements, but the text, tables, and figures are unaltered. The digital version of the original manuscript contains animations; if you would like to view these animations, please see the online version of the original (Oldroyd & Trujillo, 2021). *This is the Accepted Manuscript version of an article accepted for publication in the Astronomical Journal. IOP Publishing Ltd is not responsible for any errors or omissions in this version of the manuscript or any version derived from it. The Version of Record is available online at* https://iopscience.iop.org/article/10.3847/1538-3881/abfb6f.

Outer Solar System Perihelion Gap Formation Through Interactions with a Hypothetical Distant Giant Planet

William J. Oldroyd[1] and Chadwick A. Trujillo[1]

[1]*Department of Astronomy & Planetary Science*

Northern Arizona University

PO Box 6010, Flagstaff, AZ 86011, USA

(Received 2021 February 10; Revised 2021 April 13; Accepted 2021 April 22; Published 2021 July 1)

Submitted to AJ

ABSTRACT

Among the outer solar system minor planet orbits there is an observed gap in perihelion between roughly 50 and 65 au at eccentricities $e \gtrsim 0.65$. Through a suite of observational simulations, we show that the gap

Corresponding author: William J. Oldroyd
woldroyd@nau.edu

arises from two separate populations, the Extreme Trans-Neptunian Objects (ETNOs; perihelia $q \gtrsim 40$ au and semimajor axes $a \gtrsim 150$ au) and the Inner Oort Cloud objects (IOCs; $q \gtrsim 65$ au and $a \gtrsim 250$ au), and is very unlikely to result from a realistic single, continuous distribution of objects. We also explore the connection between the perihelion gap and a hypothetical distant giant planet, often referred to as Planet 9 or Planet X, using dynamical simulations. Some simulations containing Planet X produce the ETNOs, the IOCs, and the perihelion gap from a simple Kuiper-Belt-like initial particle distribution over the age of the solar system. The gap forms as particles scattered to high eccentricity by Neptune are captured into secular resonances with Planet X where they cross the gap and oscillate in perihelion and eccentricity over hundreds of kiloyears. Many of these objects reach a minimum perihelia in their oscillation cycle within the IOC region increasing the mean residence time of the IOC region by a factor of approximately five over the gap region. Our findings imply that, in the presence of a massive external perturber, objects within the perihelion gap will be discovered, but that they will be only $\sim 20\%$ as numerous as the nearby IOC population (65 au $\lesssim q \lesssim 100$ au).

Keywords: Trans-Neptunian objects (1705) – Planetary theory (1258) – Detached objects (376) – Oort cloud objects (1158) – Celestial mechanics (211) – Solar system astronomy (1529)

2.1 Introduction

The discovery of (90377) Sedna suggested the existence of a population of objects well beyond the known extent of the Kuiper Belt (Brown et al., 2004). As the outer edge of the Kuiper Belt was measured to be at ~ 47 au (Trujillo & Brown, 2001), a mechanism for scattering Sedna outward from the Kuiper Belt or inward from the Oort cloud was needed to explain the distant location ($q = 76.26$ au) of this new object. Some of the hypotheses proposed included single stellar encounters, migration through the stellar birth cluster, or perturbations from a distant undiscovered planet, and all of these scenarios predict the presence of additional objects at similarly high perihelia (Brown et al., 2004; Kenyon & Bromley, 2004; Morbidelli & Levison, 2004; Gladman & Chan, 2006; Levison et al., 2010; Schwamb et al., 2010).

The idea that an undiscovered planet could be responsible for the high perihelia of Sedna and other similar objects was revisited shortly after the discovery of another distant object, 2012 VP_{113} (Trujillo & Sheppard, 2014; Batygin & Brown, 2016a). Along with announcing its discovery, Trujillo & Sheppard (2014) note that a distant planet could constrain the arguments of perihelion ω of 2012 VP_{113} and other objects with similarly high semimajor axes and perihelia; these objects have been classified as Extreme Trans-Neptunian Objects (ETNOs) for semimajor axes $a \gtrsim 150$ au and perihelia $q \gtrsim 40$ au, and Inner Oort Cloud objects (IOCs) for $a \gtrsim 250$ au and $q \gtrsim 65$ au (Trujillo, 2020). In addition to highlighting this apparent orbital alignment, Trujillo & Sheppard (2014) noted that there was strong indication for an inner edge to the IOC distribution. The inner Oort cloud has also been shown to have an inner edge near the location of Sedna in Oort cloud formation simulations through interactions within a dense stellar birth cluster (Brasser et al., 2012).

Since the discovery of 2012 VP_{113}, there have been over a hundred studies conducted exploring the possible orbital parameters of the planet (Trujillo & Sheppard, 2014; Batygin & Brown, 2016a; Becker et al., 2017; Sheppard et al., 2019; Clement & Kaib, 2020), its location in the sky (Brown & Batygin, 2016; Holman & Payne, 2016; Malhotra et al., 2016; Sheppard & Trujillo, 2016; Millholland & Laughlin, 2017; Fienga et al., 2020), and the implications an undiscovered distant giant planet would have for our solar system (Bailey et al., 2016; Batygin & Brown, 2016b; Batygin & Morbidelli, 2017; Nesvorný et al., 2017; Volk & Malhotra, 2017; Li et al., 2018; Siraj & Loeb, 2020). This planet is often referred to as Planet 9 or Planet X; for more information, see review articles by Batygin et al. (2019) and Trujillo (2020) and references therein.

Recently, a third IOC has been discovered, (541132) Leleākūhonua (provisional designation: 2015 TG_{387}). Along with this discovery Sheppard et al. (2019) find that the lack of objects discovered with 50 au $\lesssim q \lesssim$ 65 au approaches the 3σ significance threshold—as objects within this region would be roughly five times easier to detect than the IOCs. They term this the "gap region" (the red box in Figure 2.1), and state that it may be formed through resonant interactions with Planet X. However, Kavelaars et al. (2020) propose that the perihelion gap may instead be a way to rule out the presence of Planet X as simulations that include Planet X scatter objects throughout the gap region rather than preserving this observed structure. Additionally, Zderic & Madigan (2020) show that a feature resembling the perihelion gap can be formed through a self-gravitating disk of primordial planetesimals which undergo an inclination instability if they collectively exceed a mass threshold of 20 Earth masses.

In this work, we examine the agreement between the observed perihelion gap and various distributions for the ETNO/IOC population through observational simulations. We also explore the relationship between the gap and the hypothesised distant giant planet using dynamical simulations and discuss the implications of our results; namely (1) the perihelion gap appears to be the result of a separation between two distinct populations, the ETNOs and the IOCs, (2) some Planet X orbits can form the ETNOs, the IOCs, and the perihelion gap from an initial Kuiper-Belt-like distribution of objects, and (3) this connection between the perihelion gap and Planet X may provide a way to constrain the mass and orbital properties of Planet X.

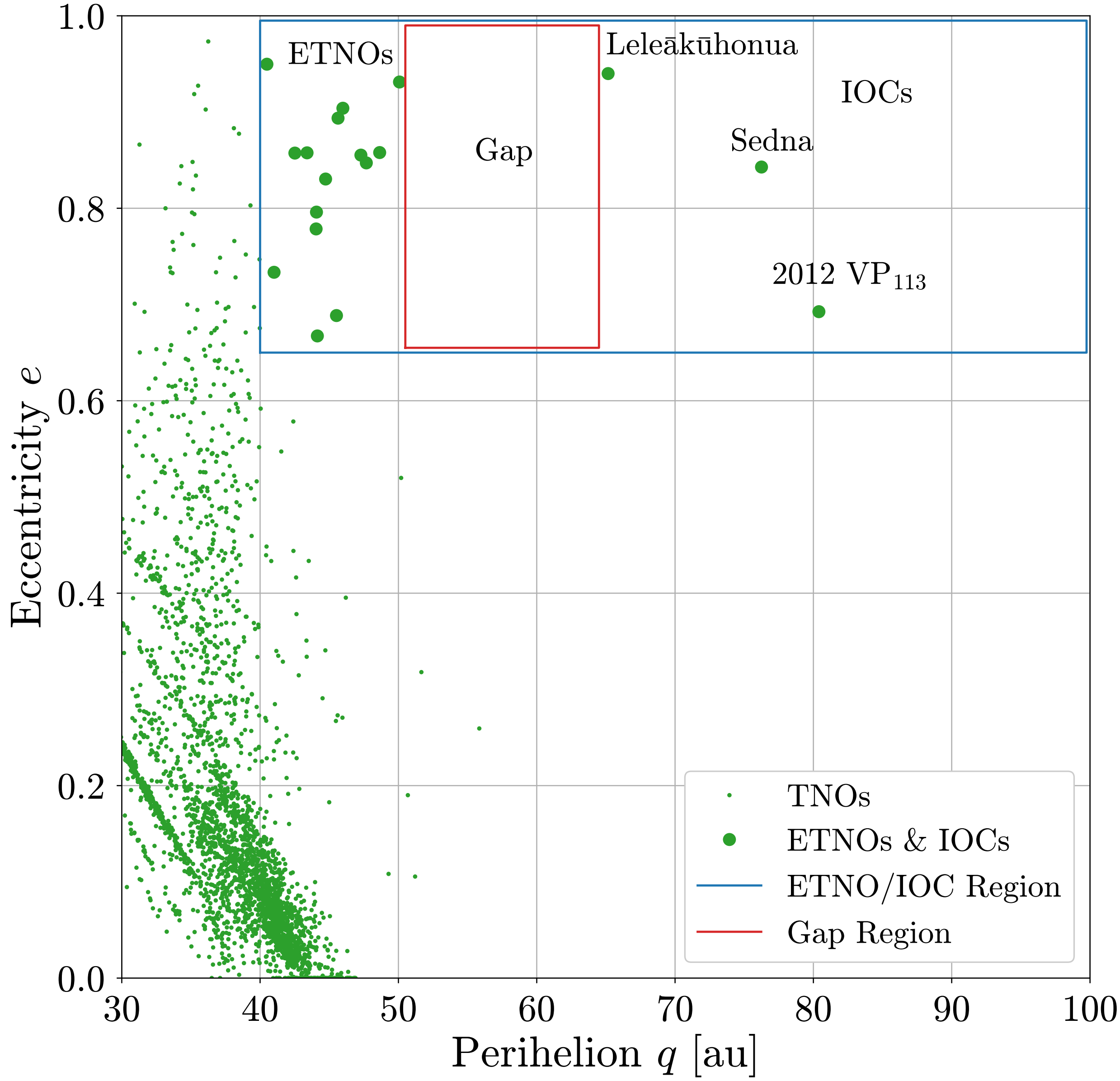


Figure 2.1 Eccentricity vs. perihelion for all known Trans-Neptunian Objects (TNOs) with perihelion $q > 30$ au (green points), circles representing the ETNOs and IOCs are enlarged for emphasis. The perihelion gap (50 au $\lesssim q \lesssim$ 65 au and $0.65 \lesssim e < 1$) is outlined in red and the ETNO/IOC region (40 au $\lesssim q \lesssim$ 100 au and $0.65 \lesssim e < 1$) is outlined in blue.

2.2 Observational Simulations

2.2.1 Parameters and Limits

We examine the statistical agreement between several outer solar system distribution models (power laws and unimodal and bimodal Gaussian distributions in a and q) and the observed distribution of ETNOs and IOCs using a suite of observational simulations. In each simulation, we assume a semimajor axis or perihelion distribution for synthetic objects within the region near the gap (the blue box in Figure 2.1; 40 au $\leq q \leq$ 100 au and $0.65 \leq e < 1$) along with parameter distributions for other orbital and physical characteristics as outlined in Table 2.1. The inclination distribution was calculated using the best-fit results from the Deep Ecliptic Survey in the form of a Gaussian, $\mathcal{G}(\mu, \sigma)$, and Lorentzian, $\mathcal{L}(I, g)$, distribution, both multiplied by the sine of the inclination, i:

$$f(i) = \{(0.75)\mathcal{G}(i, \mu = 0, \sigma = 2.3) + (0.25)\mathcal{L}(i, I = 11.2, g = 7.0)\} \times \sin i, \tag{2.1}$$

where μ is the mean of the Gaussian, σ is the Gaussian standard deviation, I is the Lorentzian half inclination, and g is the Lorentzian power (see Gulbis et al., 2010, Equation (32) and Table 2).

The size distribution was calculated as:

$$r = r_{\min} + (r_{\max} - r_{\min})(1 - \mathbb{R}[0, 1]^4), \tag{2.2}$$

where $\mathbb{R}[0, 1]$ is the range of real numbers from 0 to 1; the minimum radius $r_{\min}$ was the smallest detectable size for a spherical object in our synthetic survey corresponding to a heliocentric distance $R = 40$ au, a geometric albedo $p_V = 0.1$, and a visual magnitude $V = 26$; and the maximum radius $r_{\max}$ was the mean radius of Pluto (see Table 2.1). This radius distribution results in few large synthetic objects and many small ones. Using different power laws for the size distribution had negligible effects on the results.

When particles were drawn from a semimajor axis a distribution, perihelion q was calculated as $q = a(1 - e)$. When particles were instead drawn from a perihelion distribution, their semimajor axes were calculated using $a = q/(1 - e)$. Power-law distributions in q were calculated as:

$$q = q_{\min} + (q_{\max} - q_{\min})(1 - \mathbb{R}[0, 1]^x), \tag{2.3}$$

where $q_{\min}$ and $q_{\max}$ are the perihelion limits defined in Table 2.1 and x is a real valued exponent ranging from 1 to 4. Power-law distributions in a were calculated using Equation (2.3) as well, but with a replacing q in each instance.

Table 2.1. Observational Simulation Parameters

Parameter	Range	Distribution
eccentricity e	$0.65 \leq e < 1$	uniform
inclination i	$0^\circ \leq i \leq 90^\circ$	composite[a]
argument of perihelion ω	$0^\circ \leq \omega < 360^\circ$	uniform
longitude of ascending node Ω	$0^\circ \leq \Omega < 360^\circ$	uniform
mean anomaly M	$0^\circ \leq M < 360^\circ$	uniform
object radius r	$25 \leq r \leq 1188$ [km]	power-law[b]
geometric albedo p_V	$p_V = 0.1$	constant
perihelion q	$40 \leq q \leq 100$ [au]	varies[c]
semimajor axis a	$40 \leq a \leq 2000$ [au]	varies[c]

Note. —
[a] Determined using Equation (2.1), (see Gulbis et al., 2010).
[b] The size distribution was calculated using Equation (2.2) and limits correspond to the smallest observable object within the simulated survey limits (see Table 2.2) and the radius of Pluto.
[c] When particles were drawn from a distribution in a, q was calculated and vice versa.

Table 2.2. Observational Simulation Detection Limits

Parameter	Limits
ecliptic latitude b	$-20^\circ < b < 20^\circ$
heliocentric distance R	40 au $< R <$ 100 au
apparent sky velocity at opposition v	$0.''04\ \mathrm{hr}^{-1} < v < 4''\mathrm{hr}^{-1}$
visual magnitude V	$V < 26$

Note. — These parameters were calculated using Equations (2.4)-(2.10). The limits are derived from constraints typical of modern outer solar system observational surveys.

Synthetic objects were "detected" in the simulation if they were within the observationally derived limits outlined in Table 2.2. These constraints were selected to simulate the detection limits of a realistic outer solar system survey, including on-sky velocity v, heliocentric distance R, ecliptic latitude b, and faintest observable magnitude V. Once the detectability of these objects was determined, the distribution of detected synthetic objects was compared statistically with the true observed distribution.

The ecliptic latitude b for a given synthetic object sampled from the above distributions was calculated as

$$b = \arcsin(\sin(i)\sin(\omega + f)), \tag{2.4}$$

(see Karttunen et al., 2007, Example (6.5)) where i is the inclination, ω is the argument of perihelion, and f is the true anomaly

$$f = 2\arctan\left(\sqrt{\frac{1+e}{1-e}}\tan\left(\frac{E}{2}\right)\right), \tag{2.5}$$

(Murray & Dermott, 1999, Equation (2.46)) with E being the eccentric anomaly. To determine E, we utilized the methods outlined in Murison (2006) with Kepler's equation

$$E - e\sin(E) = M, \tag{2.6}$$

where M is the mean anomaly (Murray & Dermott, 1999, Equation (2.52)). Synthetic objects with $|b| > 20°$ were excluded from detections similar to the majority of survey fields observed by Sheppard et al. (2019).

Limits for apparent on-sky velocity and heliocentric distance overlap, particularly for the lower limit for v which is much slower that objects at the upper limit for R. However, the upper limit for v was useful for removing fast-moving synthetic objects within the heliocentric distance range. Heliocentric distance, R, was computed using the equation

$$R = \frac{a(1-e^2)}{1+e\cos(f)} \tag{2.7}$$

(Murray & Dermott, 1999, Equation (2.19)). We placed a minimum bound on heliocentric distance at 40 au and set a maximum value for R at 100 au as nearly all TNOs have been discovered at $R < 100$ au.[1]

The apparent on-sky velocity (in arcseconds per hour) at opposition, v, was calculated following the method of Luu & Jewitt (1988):

$$v = 148\frac{1-R^{-0.5}}{R-1}, \tag{2.8}$$

where $R > 1$. This makes the assumption that synthetic objects are discovered at opposition. We set observational limits such that synthetic objects are detected by our simulated survey only if they exhibit on-sky velocities between $0.''04$ and $4''$ per hour to rule out objects that require multiple nights between observations to detect movement or objects moving too quickly to be ETNOs.

A final key limit employed in our model is the limiting magnitude of the survey. The apparent magnitude V for synthetic objects is calculated by first determining the absolute magnitude, H, of a given synthetic object which we do by rearranging Equation (7) from Harris & Harris (1997)

$$H = 5\log_{10}\left(\frac{1329\,\mathrm{km}}{2r\sqrt{p_V}}\right), \tag{2.9}$$

[1]With two notable exceptions (Sheppard et al., 2018, 2021).

where r is the object radius and p_V is the geometric albedo of the object (see Table 2.1). Using the H magnitude, we can transform to visual magnitude

$$V = H + 5\log_{10}(R(R-1)) + C, \tag{2.10}$$

where the constant $C = -2.5\log_{10}(2/3) = 0.44$ is the phase angle term at opposition (see Karttunen et al. (2007), Equations (7.43) & (7.48) and Whitmell (1907), Equation (2)). We chose a magnitude cutoff of $V < 26$ as a faintness limit on our simulated survey (similar to the limiting magnitude obtained by Sheppard et al., 2016).

Synthetic objects were considered "observed" if they were within all observational limits (summarized in Table 2.2). Each simulation was run until the number of synthetic objects detected in the simulation was equal to the number of known ETNOs and IOCs within the considered region ($N = 18$) which are outlined in Table 2.3. For most distributions in a and q, the vast majority of particles were not observable, thus 10^2 to 10^5 synthetic objects were generated for each distribution in order to be able to detect the prescribed number.

2.2.2 Statistical Analysis

To determine the probability that the apparent gap is the result of a random draw from a single continuous distribution, we utilize rolling histograms of perihelion to improve continuity of analysis given the relatively small sample size (see Figure 2.2). A variety of bin widths and spacings were tested, all resulting in similar statistical trends. For the analyses presented here, a bin width of 6 au and a bin spacing of 2 au were used, as this combination highlights the features within the gap region. Both real and synthetic objects were binned (separately) and the model was run 10^4 times for each distribution in semimajor axis and perihelion.

From the results of the model runs, a Poisson maximum likelihood (see Bonamente, 2017, p. 91) was derived for each bin

$$\hat{\lambda}_i = \frac{1}{n}\sum_{j=1}^{n} N_{i,j}, \tag{2.11}$$

where n is the number of sample runs (10^4), and $N_{i,j}$ is the number of particles in the ith bin of the jth sample run. Using this maximum likelihood and the results from the model runs, we employed two statistical metrics to determine the agreement between the model distributions and the observed distribution: the Poisson Probability, and the Kolmogorov-Smirnov (K-S) Test, as described below.

Table 2.3. Real ETNOs and IOCs Used

Name/Designation	q (au)	e	a (au)
(90377) Sedna *	76.26	0.843	484.5
(148209) 2000 CR_{105}	44.08	0.796	216.1
(474640) 2004 VN_{112}	47.29	0.856	327.5
(496315) 2013 GP_{136}	41.01	0.734	153.9
(505478) 2013 UT_{15}	44.05	0.779	198.9
(541132) Leleākūhonua *	65.16	0.940	1085
2010 GB_{174}	48.65	0.858	342.3
2012 VP_{113} *	80.40	0.693	261.5
2013 FT_{28}	43.39	0.858	304.8
2013 RA_{109}	45.98	0.904	479.5
2013 SY_{99}	50.08	0.931	728
2014 SR_{349}	47.68	0.847	311.9
2014 SS_{349} †	45.29	0.697	149.5
2014 WB_{556}	42.65	0.852	288.1
2015 KG_{163}	40.49	0.950	805
2015 KE_{172} †	44.14	0.667	132.7
2015 RX_{245}	45.69	0.891	420.9
2018 VM_{35}	44.96	0.822	251.9

Note. — *: Real ETNOs and IOCs used in this work. †: these objects technically do not qualify as ETNOs under the definition of Trujillo (2020), as they have semimajor axes slightly less than 150 au, however, they both lie within the ETNO/IOC region we define in Figure 2.1 (40 au $\leq q \leq$ 100 au, 0.65 $\leq e <$ 1), so we include them for the purpose of these analyses. These values were retrieved from the JPL Small-Body Database (Giorgini et al., 1996) on 2020 December 22.

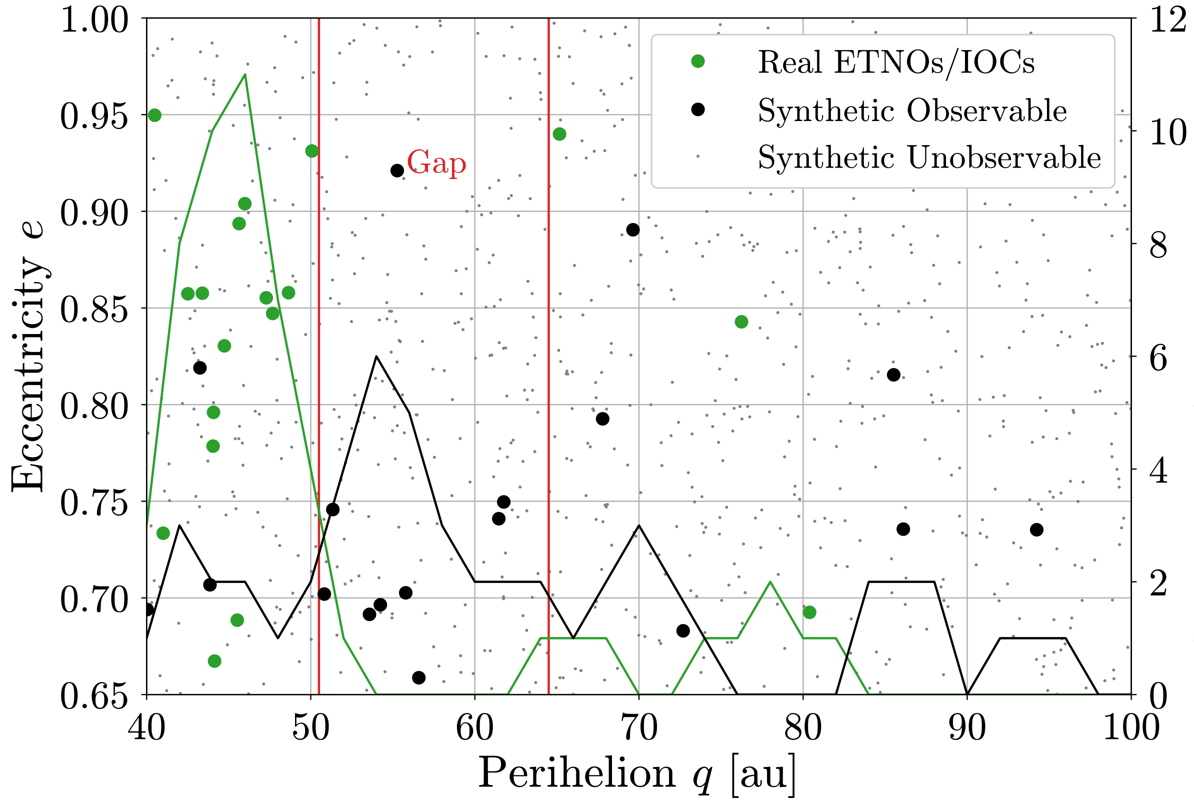


Figure 2.2 Example rolling histograms of binned perihelion. Real ETNOs and IOCs are shown as green filled circles. Black filled circles indicate synthetic objects that were within the observational limits outlined in Table 2.2. Synthetic objects outside these limits are “unobservable,” are shown as gray points, and are not included in further analysis. The perihelion gap is outlined in red. A bin width of 6 au and a bin spacing of 2 au are used. This example is typical of our observational simulations using a single continuous distribution in a or q and synthetic objects appear in the gap region with similar frequency compared to the surrounding regions.

The Poisson Probability is a measure of how likely the observed distribution is to be drawn from the Poisson maximum likelihood distribution and is computed as

$$P = \exp \sum_{i=1}^{n_{\text{bins}}} \left[\ln e^{-\hat{\lambda}_i} + \ln \frac{1}{N_{\text{real},i}!} + \ln \hat{\lambda}_i^{N_{\text{real},i}} \right], \tag{2.12}$$

where n_{bins} is the number of histogram bins, $\hat{\lambda}_i$ is the Poisson maximum likelihood for a given bin, and $N_{\text{real},i}$ is the number of real objects in a given histogram bin (see Bonamente, 2017, p. 91).

The K-S Test is a metric for determining the likelihood that two data samples are drawn from the same distribution (see Press et al., 2007, Chapter 14). We implement the `ks_2samp` function from the `scipy.stats` `python` package (Virtanen et al., 2020) to compare the synthetic objects from all 10^4 model runs, collectively, with the real objects. The reported metric for these tests is the p value obtained from the analysis.

We ran the model 10^4 times (sufficient to get a statistical representation; additional runs have minimal effect on the results) for each of 15 distributions (5 in semimajor axis and 10 in perihelion) and calculated the statistical distribution agreement metrics for the results. To compare these single-population distributions with a two-population distribution, we fit the observed distribution with the sum of two Gaussian probability density functions using Markov Chain Monte Carlo (MCMC) statistical sampling described in Section 2.2.3. The model was then applied to the maximum likelihood fit from this analysis 10^4 times and the same statistical metrics were calculated for the model results. The results of the statistical analyses are presented in Table 2.4.

2.2.3 Best-fit Distribution

To compare the effectiveness of a bimodal model, which most resembles the ETNO and IOC populations, with the unimodal distributions described in Section 2.2.1, we implemented an MCMC statistical sampling fit for the observed object distribution using a two-Gaussian model. To carry out this statistical fit, we utilized the Affine-Invariant Ensemble Sampler from the `emcee` `python` package (Foreman-Mackey et al., 2013). As a forward model (used for simulating results from the varied parameters at each step that can later be compared with the data), we employed our observational simulation procedure of drawing particles from the distributions given in Table 2.1, with a two-Gaussian perihelion distribution (see Table 2.5) and the observational limits from Table 2.2 until the number of synthetic objects detected in the simulation was equal to the number of ETNOs/IOCs in the considered region ($N = 18$, see Table 2.3). This process was repeated 25 times, resulting in averaging over hundreds of particles, for each model step before calculating the Poisson maximum likelihood of the rolling histogram bins corresponding to those used in Section 2.2.2.

Table 2.4. Distribution Agreement

Parameter	Distribution	P	K-S Test
a	Uniform	6.08 e-24	0.0006
a	Power-law $x = 1$	1.80 e-27	<0.0001
a	Power-law $x = 2$	4.67 e-30	<0.0001
a	Power-law $x = 2.7$	5.86 e-31	<0.0001
a	Power-law $x = 4$	2.31 e-31	<0.0001
q	Uniform	4.99 e-24	0.0005
q	Power-law $x = 1$	2.55 e-20	0.0214
q	Power-law $x = 2$	2.14 e-19	0.1697
q	Power-law $x = 3$	1.33 e-19	0.5484
q	Power-law $x = 4$	9.18 e-21	0.7150
q	Gaussian $\sigma = 5$	0	0.0053
q	Gaussian $\sigma = 10$	6.10 e-29	0.7385
q	Gaussian $\sigma = 15$	3.32 e-21	0.3230
q	Gaussian $\sigma = 20$	2.70 e-20	0.0735
q	Gaussian $\sigma = 25$	1.13 e-20	0.0228
$\boldsymbol{q}$	**Two-Gaussian fit**	**6.91 e-12**	**0.9901**

Note. — Agreement between the observed ETNO/IOC distribution and modeled synthetic object distributions measured by the Poisson Probability P (Equation (2.12)) and the K-S test. The semimajor axis distribution with a power-law exponent of 2.7 was chosen to match the distribution used in Sheppard et al. (2019) for observational bias calculations. Distributions for other orbital and physical parameters used in these models are given in Table 2.1. Single Gaussian distributions all have a mean $\mu = 40$ au. The two-Gaussian-fit distribution is the maximum likelihood from the MCMC fit (see Section 2.2.3 and Table 2.5). Larger values for the Poisson Probability and K-S test correspond to better distribution agreement between the observed and synthetic objects. Values of 0 indicate that the probability was too low to measure in 10^4 model runs. Note that the two-Gaussian-fit distribution has significantly better ($\sim 10^7$) agreement with the observed distribution than any other distribution tested.

Table 2.5. MCMC Parameters and Priors

Parameter	Priors	Model Fit	Max $\mathcal{L}$
Gaussian ratio a_g	$0-1$	$0.59^{+0.10}_{-0.11}$	0.62
mean 1 μ_1	$40-50$ au	$44.9^{+0.71}_{-1.25}$ au	45.0 au
standard deviation 1 σ_1	$0.1-30$ au	$3.1^{+1.21}_{-0.74}$ au	2.8 au
mean 2 μ_2	$60-90$ au	$76.9^{+5.83}_{-3.66}$ au	75.3 au
standard deviation 2 σ_2	$0.1-30$ au	$9.4^{+9.08}_{-3.40}$ au	4.9 au

Note. — Parameters, priors, and results from the two-Gaussian MCMC fit for the observed ETNO/IOC distribution. The Gaussian ratio is a weight defining the fraction of particles drawn from the first Gaussian versus the second. Priors for Gaussian means were chosen to generously encompass the ETNOs and the IOCs, respectively. The standard deviation priors were also chosen generously to ensure the most likely values would fit comfortably within the range. The model fit is the median of the sample chain distribution with 1σ quantiles reported (as is typical for MCMC analyses). The maximum likelihood values $\mathcal{L}$ are the parameters from the single MCMC sample that had the best fit to the data, hence, $\mathcal{L}$ may vary between iterations, but serves as an excellent parameter combination to use for our observational model. A corner plot of the sample distributions reported here is given in Figure 2.9.

The natural log of the Poisson Probability (Equation (2.12)) was used at each step as a likelihood function and priors (see Table 2.5) were incorporated to give a measurement of the goodness of fit for each model at each step. Short preliminary runs to obtain approximate best-fit values were carried out. The best fit from the short runs was then used as the initialization point for 50 walkers; walkers are statistical sampling chains that pseudo-randomly change their parameters, with weight toward higher likelihoods, at each model step. Each walker was then given a small random offset from the central initialization position to improve the rate at which they explore the parameter space. The model was then run for 15,000 steps which allowed ample time for them to converge (see Appendix 2.7.1).

Model fit and maximum likelihood parameters from the MCMC sampling are outlined in Table 2.5. Maximum likelihood values obtained through the MCMC fit were run through the original observational model 10^4 times and the statistical tests outlined in Section 2.2.2 were applied as they were for the other distributions. The two-Gaussian fit resulted in significantly higher distribution agreement than any of the other tested distributions. The results are shown in Table 2.4, where we see that the two-Gaussian fit is a better fit to the data by ~ 7 orders of magnitude. Figure 2.3 shows the two-Gaussian fit overlaid on the ETNO/IOC distribution. Here we see a steep reduction in the number of objects at the inner edge of the gap ($\sim$50 au), the perihelion gap with few objects between $\sim$50–65 au, and the IOCs increasing in number beyond $\sim$65 au.

The improvement in fit of the two-Gaussian fit over the unimodal models indicates that the ETNOs and the IOCs are statistically more than seven orders of magnitude more likely to be two separate populations separated by the gap region rather than the observed distribution developing from a single continuous population. The orbital alignment of the ETNOs and the IOCs has been used as some of the primary evidence for the existence of Planet X (Trujillo, 2020). As the separation between these populations defines the perihelion gap, we examine the relationship between Planet X and the ETNOs, IOCs, and "gap objects" to ascertain whether the Planet X hypothesis is compatible with the presence of the perihelion gap.

2.3 Dynamical Simulations

To explore the dynamical cause of the separation in q between the ETNOs and IOCs, and the effect an additional distant perturber would have on the perihelion gap, we ran a suite of computational N-body simulations. In these simulations, a range of initial test particle distributions interacted with the Sun, the known giant planets, and Planet X.

In each simulation between 10^2 and 10^4 test particles were drawn from distributions in q and e to explore the effect of initial particle distribution on the gap region. The test particle distributions in our observational simulations were used for i, ω, Ω, and M in our dynamical simulations (see Table 2.1). All particles were integrated for 4.5 Gyr with a 0.2 yr time step (roughly 1/60th of the orbital period of Jupiter, the innermost orbit in the simulations), which we found to give an appropriate balance between low amounts of relative energy error (between $\sim 10^{-8}$ and $\sim 10^{-7}$) and acceptable run times (generally between 24 and 48 hr per simulation, see Rein & Spiegel, 2015). All integrations were performed using the REBOUND N-body integration `python` package with the hybrid symplectic `mercurius` integrator designed for both speed and accuracy in complex planetary dynamics scenarios (Rein et al., 2019).

In order to reduce compute time, dynamical simulations were run in parallel on the Northern Arizona University high-performance computing cluster, *Monsoon*, which has over 2800 cores, peak CPU performance of over 100 teraflops, and is free to use for NAU faculty, staff, and students.[2] During integration, particles exceeding a distance threshold (2×10^5 au) were considered ejected and were removed from the simulation to further improve computation time; this is 10 times greater than the limit beyond which Kaib et al. (2011) noted that $\sim$90% of objects are ejected over the age of the solar system through galactic tides and stellar perturbations, hence, very few particles beyond this limit would survive and can be ignored.

[2] `https://in.nau.edu/hpc/details/`

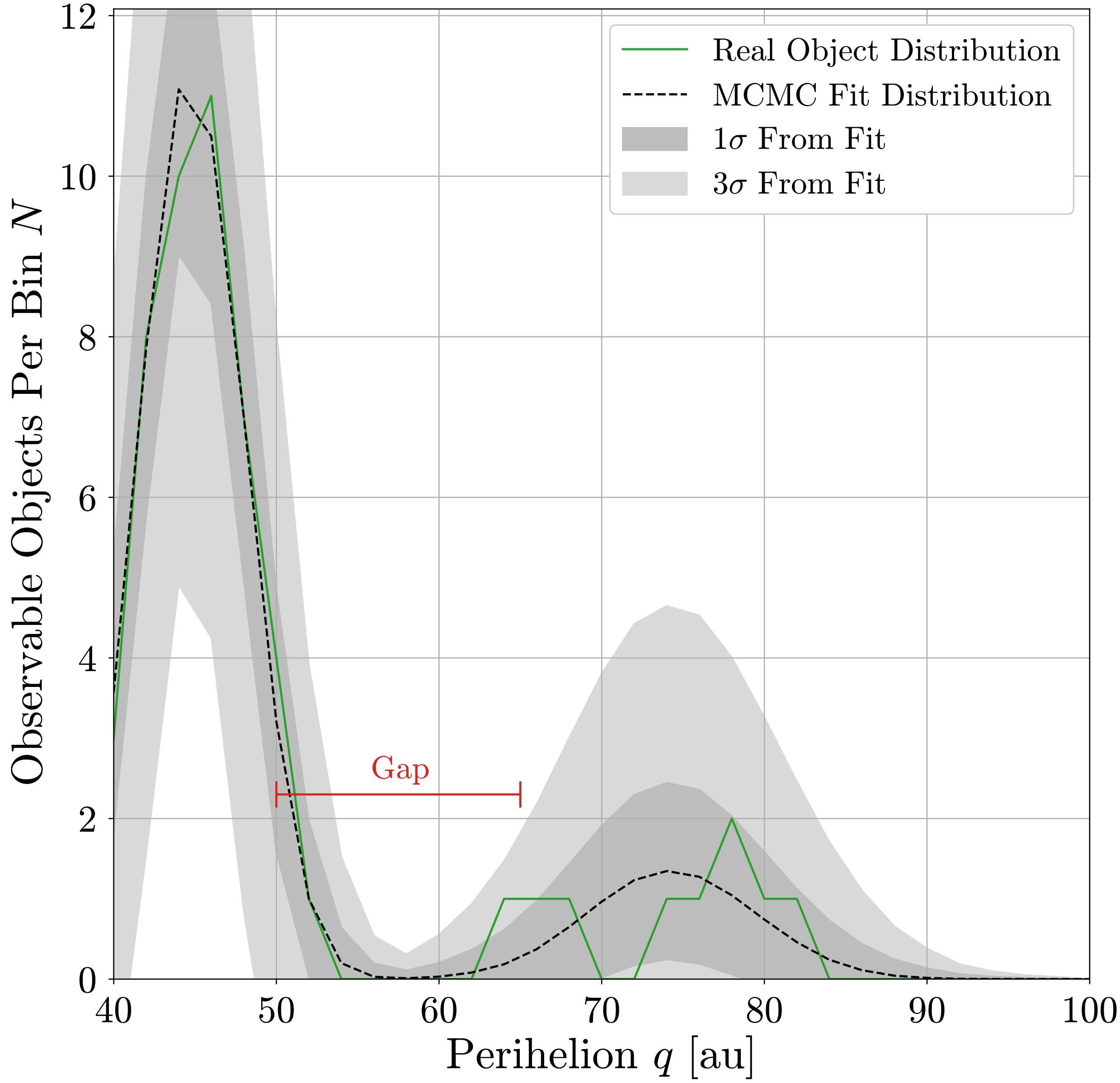


Figure 2.3 Rolling histogram in perihelion of the observed ETNO/IOC distribution (green) and the two-Gaussian MCMC fit (dashed black line). The dark shaded area indicates the 1σ error on the Poisson maximum likelihood and the light shaded area is the 3σ error. The extent of the observed perihelion gap is shown for emphasis. Note the excellent agreement between these two distributions.

For each simulation, we divided the results into 10 Myr time steps and computed the average number of particles per region per time step on a grid in perihelion and eccentricity. These results were then plotted as "heatmaps" (particle density plots) with the $\log_{10}$ of test particle density shown as a color gradient with the real objects overplotted. This facilitates qualitative comparison of the time particles spend in each region between the simulated and observed distributions.

The simulations that were most successful at reproducing the observed features of the ETNOs, the perihelion gap, and the IOCs were initialized with a Kuiper-Belt-like distribution of test particles with 25 au $\leq q \leq$ 40 au and $0 \leq e \leq 0.4$. These simulations not only do a better job than other initial distributions we tested at forming a gap-like structure, they also resemble the observed Kuiper Belt. Because our focus is on the region near the gap—the ETNOs, and the IOCs—it is not essential that our simulations exactly reproduce all features of the less extreme TNO distributions (such as the number of objects populating various resonances). Thus, we assume that a simple approximation of a realistic initial TNO distribution, such as the one used here, is sufficient for our investigation.

When no Planet X is included in our simulations, a simplified overall Kuiper-Belt-like structure is preserved with the main part of the belt concentrated between roughly 35 au $\lesssim q \lesssim$ 40 au. Several test particles are scattered to high e by Neptune, but there is no scattering to high q in this system, thus, no ETNOs or IOCs are formed and no perihelion gap is present (see Figure 2.4). These simulation without Planet X, therefore, do not explain the observed features of the distant outer solar system.

Including Planet X in the simulations produced markedly different results. We tested several different candidate parameter sets for Planet X from the literature (e.g., Batygin et al., 2019; Brown & Batygin, 2019; Trujillo, 2020, see Table 2.6). We find that, with Planet X, an initial test particle distribution uniform in q does not produce a perihelion gap. For Kuiper-Belt-like initial distributions a roughly Kuiper-Belt-like structure is preserved, many Neptune resonances are more readily populated than in the case without Planet X, and particles scattered to high e by Neptune are subsequently scattered to high q by Planet X (see Figure 2.5). This outward scattering effect happens along lines of roughly constant a as particles are captured into resonance with Planet X. These particles oscillate in eccentricity, in some cases reaching $e < 0.1$ (see Figure 2.6 and the associated animation in Figure 2.12). We also note that many of the particles in the ETNO and IOC region in our simulations clearly exhibit "resonance hopping," as discussed in Bailey et al. (2018) and Khain et al. (2020), which accounts for their slight semimajor axis variations seen over secular timescales.

To explore the effectiveness in which Planet X orbits can produce a perihelion gap, we use the methods outlined in Section 2.2, and compare the number densities of objects for the three Planet X parameter combinations outlined in Table 2.6 over the ETNO region, the gap region, and the IOC region (40 au $\leq q \leq$ 100 au and $0 \leq e < 0.65$; blue box in Figure 2.1). The perihelion distribution of the simulated particles was

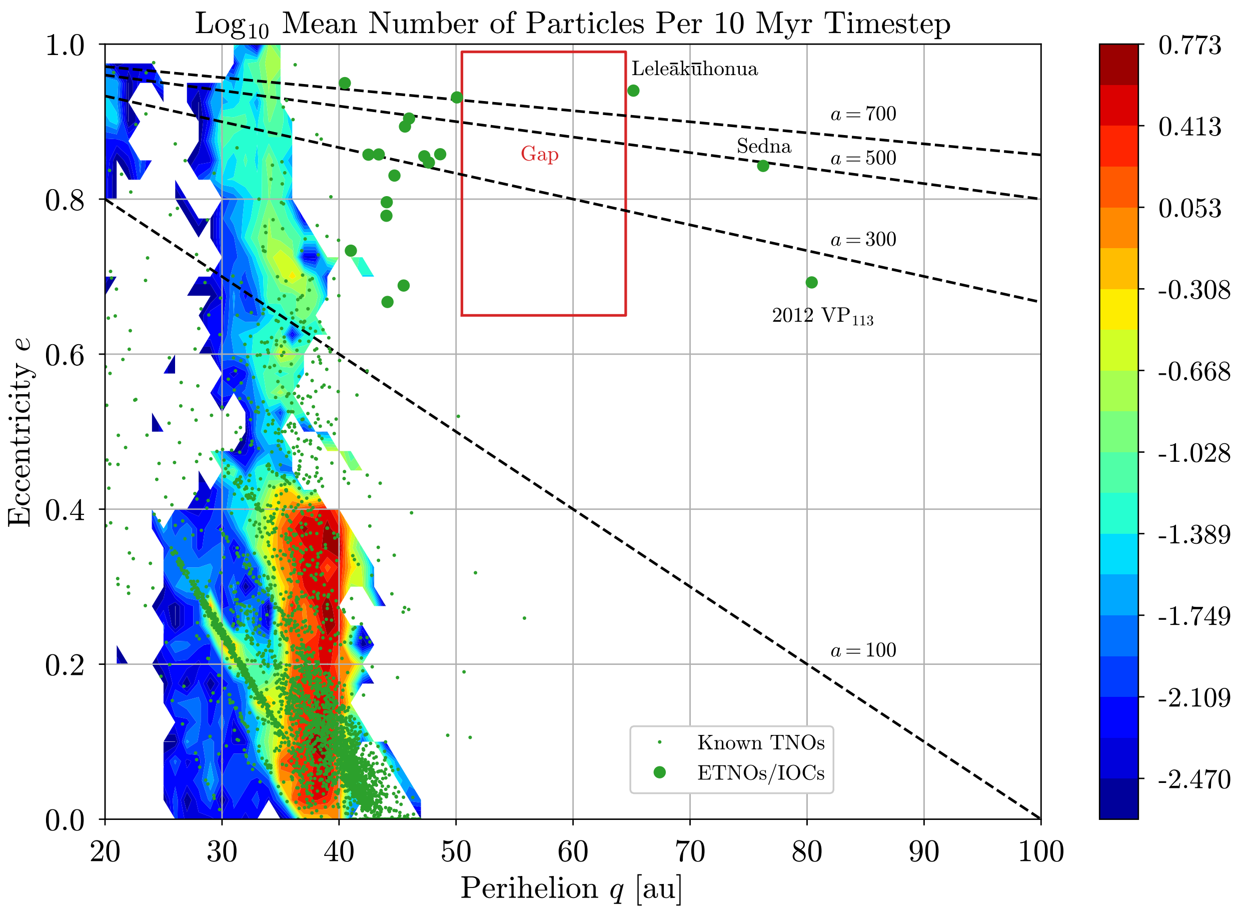


Figure 2.4 Heatmap in perihelion and eccentricity of the $\log_{10}$ mean number of particles per 10 Myr time step. Known objects are overplotted as green filled circles and ETNOs and IOCs are enlarged for emphasis. Deep reds indicate a constant high particle density and deep blues show that few particles passed through this region over the course of the simulation. White regions had no particle enter the area during the 4.5 Gyr of integration. The gap region and lines of constant a are drawn for reference. This simulation contained the Sun, the giant planets, and an initial Kuiper-Belt-like distribution of 10^3 test particles with 25 au $\leq q \leq 40$ au and $0 \leq e \leq 0.4$ and no Planet X. Here we see that the majority of particles remain in a Kuiper-Belt-like region with a small number being scattered to high e by Neptune. No objects, however, are scattered outward to high q near the region of the ETNOs, the perihelion gap, or the IOCs. An animation of the synthetic particles used to create this figure is available in the electronic version of this manuscript (Figure 2.10).

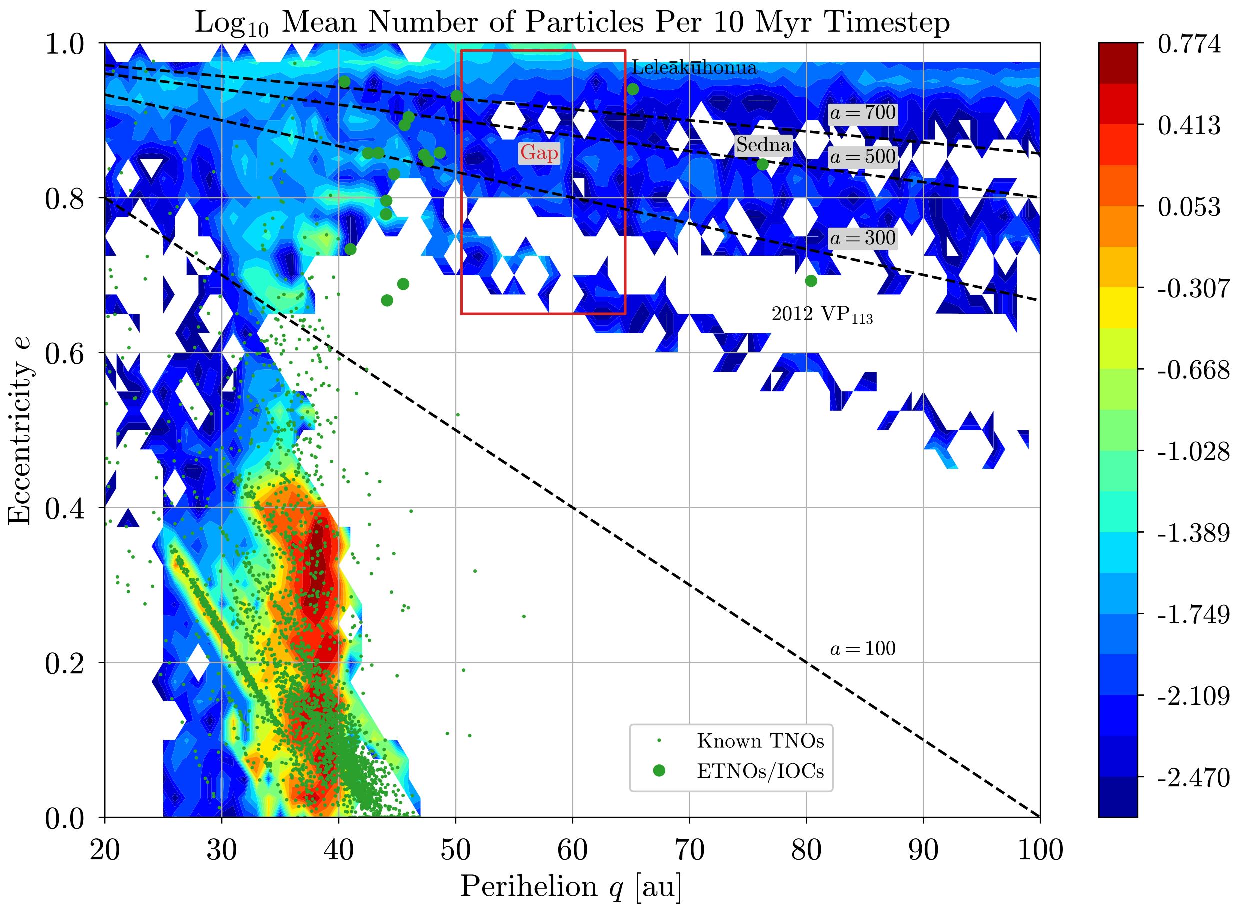


Figure 2.5 Heatmap in perihelion and eccentricity of the $\log_{10}$ mean number of particles per 10 Myr time step. As Figure 2.4 with the addition of Planet X B19 from Table 2.6 (Batygin et al., 2019). Here we see a similar pattern at low e as we observed in the no Planet X case where the majority of particles remain in a Kuiper-Belt-like distribution (see Figure 2.4). However, with a planet we find that particles are scattered outward to high q by Planet X after they are scattered to high e by Neptune. Additionally, we see a gap-like underdensity, particularly for objects with $0.65 \lesssim e \lesssim 0.8$ passing through the perihelion gap. An animation of the synthetic particles used to create this figure is available in the electronic version of this manuscript (Figure 2.11).

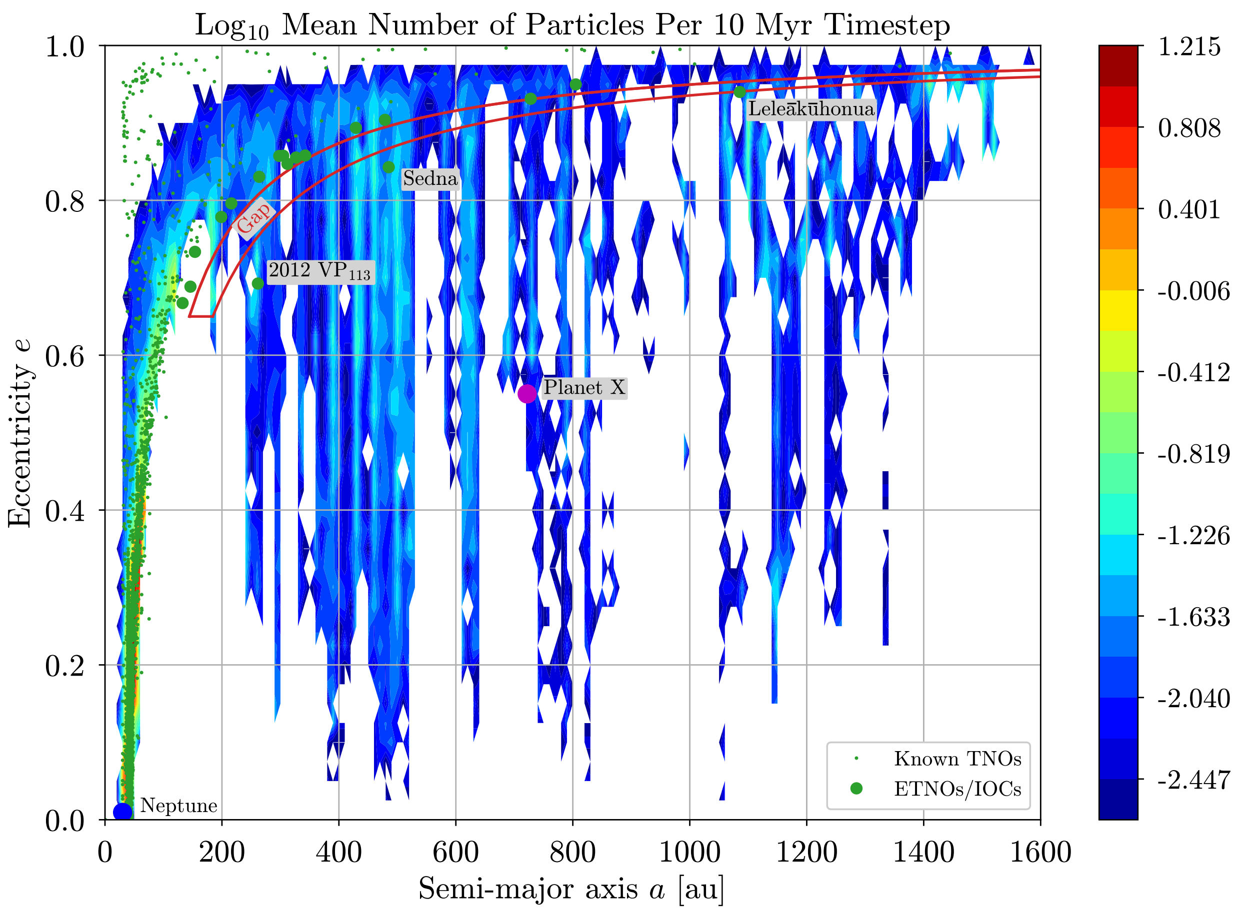


Figure 2.6 Heatmap in semimajor axis and eccentricity of the log$_{10}$ mean number of particles per 10 Myr time step. As Figure 2.4 (in e vs. a rather than e vs. q) with the addition of Planet X T20 from Table 2.6 (Trujillo, 2020). Neptune and Planet X are shown as large blue and magenta circles respectively. Here we see that objects have their eccentricities dampened as they are scattered to higher perihelia across the gap region along lines of roughly constant semimajor axis by Planet X (the vertical bars). These objects then oscillate between high and low e and many exhibit "resonance hopping" (Bailey et al., 2018; Khain et al., 2020) inducing a slight shift in their semimajor axes. We note that in simulations without Planet X, no objects cross the perihelion gap to low eccentricities (high perihelia). An animation of the synthetic particles used to create this figure is available in the electronic version of this manuscript (Figure 2.12).

Table 2.6. Sample Planet X Parameters

Parameter	BB19[a]	B19[b]	T20[c]
mass m	6 $m_{\oplus}$	5 $m_{\oplus}$	10 $m_{\oplus}$
semimajor axis a	300 au	500 au	722 au
eccentricity e	0.15	0.25	0.55
inclination i	17°	20°	29°
argument of perihelion ω	...	...	142°
longitude of ascending node Ω	...	...	93°
initial true anomaly f	...	...	174°
perihelion q	255 au	375 au	325 au

Note. — These were the three parameter sets for Planet X that we tested most extensively. $m_{\oplus}$ = Earth masses. When values for ω, Ω, and f were not given, we used 0°.
[a] Brown & Batygin (2019).
[b] Batygin et al. (2019).
[c] Trujillo (2020).

sampled every 500 Myr and aggregated into an overall distribution for each simulation. We use a similar rolling histogram technique to that outlined in Section 2.2, but here, simulated objects were each given an observability score by assuming a range of sizes with a distribution matching the radius limits in Table 2.1. For each simulated object, the observability limits were computed for 10^3 radii following an r^4 distribution between $r_{\min}$ and $r_{\max}$. Each simulated object was then given a weight based on the smallest radius it could have while still satisfying the observational survey limits outlined in Table 2.2. These weighted observability scores of the synthetic objects were then binned and scaled to the peak value of the real ETNO and IOC data.

The results from this analysis, shown in Figure 2.7, indicate that, for some parameter combinations for Planet X (see Table 2.6), a feature resembling the perihelion gap is formed. The Planet X BB19 curve has a deep gap-like feature, however, it is offset from and is a poor fit to the observed ETNO/IOC population distribution. Only a weak gap feature is seen in the Planet X B19 curve, but in the Planet X T20 curve, there is a moderate gap feature close to the location of the observed perihelion gap. Hence, there are parameter combinations for Planet X that can form the ETNOs, the IOCs, and features resembling the perihelion gap from simple Kuiper-Belt-like initial distributions. Additionally, we note that we do not favor any particular set of Planet X parameters based on these analyses as we have not evaluated the effect which small differences in parameters may have on the ability of a given Planet X orbit to form a perihelion gap.

Because of the difference in ability of Planet X orbits to produce the perihelion gap given different parameter combinations, the extent and location of the gap may provide valuable constraints on the mass

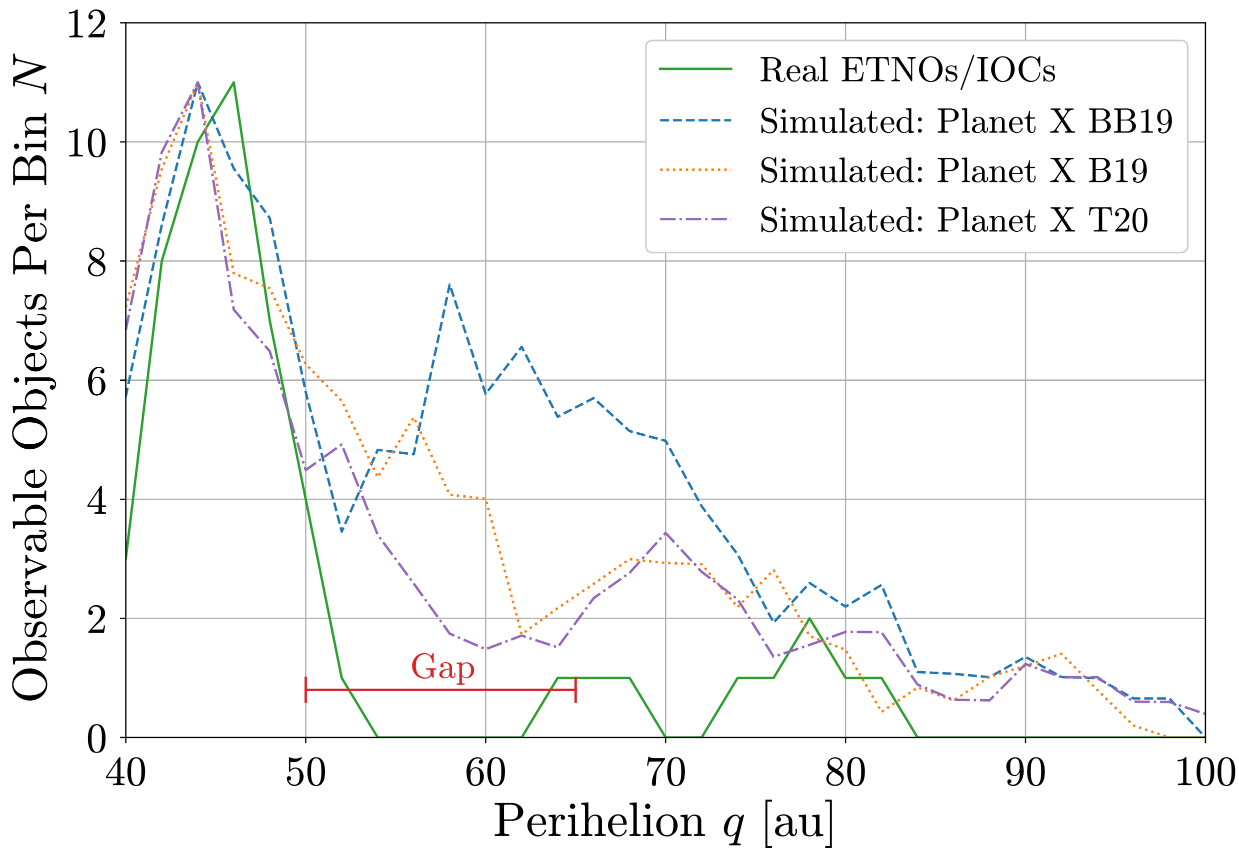


Figure 2.7 Rolling histogram in q of the real ETNOs and IOCs (solid green) and synthetic objects (dashed blue, dotted orange, and dotted-dashed purple) from dynamical simulations containing Planet X with parameters corresponding to those outlined in Table 2.6. Note that the simulated data all exhibit gap-like features. Planet X BB19 causes a deep gap-like feature between ∼50 and 56 au, Planet X B19 causes a small dip resembling a gap at roughly 62 au, and Planet X T20 causes a broad gap-like feature from ∼56 to 68 au. Additionally, none of these features is the result of an absence of objects within the perihelion gap, but rather is marked by a local minimum in the number of observable objects. We note that we do not favor any particular parameter combination for Planet X given our analyses, but emphasize that many different Planet X orbits may result in the observed perihelion gap. Additionally, simulations that do not include Planet X do not produce ETNOs, Gap Objects, or IOCs and result in zero high eccentricity observable objects for every perihelion value shown in this figure.

and orbit of Planet X upon further examination. In particular, the perihelion gap may be used to rule out some, but not all, parameter combinations for Planet X (see Kavelaars et al., 2020).

2.4 Discussion

From the results in Table 2.4, we find that a bimodal distribution (the two-Gaussian fit) is a better fit, over a single continuous distribution, to the observed ETNOs and IOCs by several orders of magnitude. This is in agreement with previous estimates for the statistical significance of the gap region of nearly 3σ (Sheppard et al., 2019). We find all continuous distributions we tested (a wide range that encompasses most realistic possibilities) to be poor fits to the observed distribution of ETNOs and IOCs, and, thus, we find it unlikely that the perihelion gap is the result of a continuous population of distant objects in perihelion or semimajor axis and eccentricity. As more IOCs are discovered, the extent and magnitude of the perihelion gap will become more clear; roughly 10 ETNOs/IOCs discovered in a single well characterized survey (such as the Legacy Survey of Space and Time) would raise the significance above 3σ for that survey alone (Trujillo, 2020).

By examining our dynamical simulations that do not contain Planet X, we find that neither the ETNO population nor the IOC population can be formed from a simplified Kuiper-Belt-like initial configuration solely by Neptune scattering. An alternative method for forming these populations is Oort cloud diffusion caused by galactic tides and close stellar passages. This mechanism has been shown to have a $\sim$20%-30% likelihood of placing objects onto orbits within the ETNO and IOC regions, particularly if the Sun experienced outward radial migration in the galactic disk (Kaib et al., 2011). However, this method does not appear to produce a perihelion gap, but rather a single population of Oort cloud objects that extends from the ETNO region outward (see Figure 11 of Kaib et al., 2011). In our dynamical simulations, we focus on the ability of Planet X to cause the perihelion gap through secular interactions and, thus, do not further explore the effects of stellar encounters or galactic tides on the system. Additionally, we assume that interactions with Planet X take place after planet migration and, hence, initial conditions resemble a Kuiper-Belt-like structure.

Initial configurations that instead begin with substantial ETNO and IOC populations separated by a perihelion gap largely retain this structure and preserve the gap over the age of the solar system in the absence of Planet X. This is expected since Neptune only weakly influences objects with $a > 60$ au even accounting for mean-motion and Kozai resonances (Gomes et al., 2008). Objects in these regions are unlikely to have formed in situ—they would be ejected during planet migration if they formed early on and at later times there was insufficient material to form efficiently at these distances—(Kenyon & Luu, 1999; Morbidelli

& Nesvorný, 2020) and we find it is difficult to scatter objects to these regions without the help of a massive external perturber, such as Planet X (or a similarly massive collection of distant objects, Zderic & Madigan, 2020). In our dynamical simulations that began with an initial Kuiper-Belt-like particle distribution and did not contain Planet X, we found that some objects were scattered to high eccentricities by Neptune, but no objects had their perihelia raised to become ETNOs, "gap objects," or IOCs.

When Planet X is present in these simulations, it is the dominant gravitational perturber for objects with orbits in the ETNO and IOC regions. Interactions with Planet X cause objects scattered to high eccentricities by Neptune to migrate into the ETNO region ($0.65 \lesssim e < 1$ and 40 au $\lesssim q \lesssim$ 50 au) by raising their perihelia. As ETNO perihelia increase, the influence of Neptune on their orbits is diminished, though this also depends on the eccentricities of the objects. Once objects reach a perihelion of roughly 50 au, their perihelia are increased much more quickly by Planet X. It is not surprising, then, that the perihelion gap begins at $q \sim 50$ au, since objects crossing this transition boundary appear to be completely detached from the influence of Neptune and tend to be captured into secular and/or mean-motion resonances with Planet X (e.g., the vertical patterns in Figure 2.6).

The time it takes objects to cross the perihelion gap depends on the parameters selected for Planet X, with mass being a primary factor. On average, objects spend $\sim$1-2$\times 10^4$ yr within the gap region in the presence of a 10 Earth mass Planet X and for a 5 Earth mass Planet X they take longer to traverse the gap ($\sim$2-4$\times 10^4$ yr; see Figure 2.8). After crossing the gap, objects continue to have their perihelia increased by Planet X until they reach their maximum perihelion, which is different for each object and each secular resonance cycle. This increase in perihelion is accompanied by a decrease in eccentricity, hence, the maximum perihelion within a secular cycle for an object corresponds to a minimum eccentricity for that object. These long-term resonant librations in q and e happen along lines of roughly constant semimajor axis (see Figure 2.6 and the associated animation, Figure 2.12).

As these now very distant objects reach their maximum q (minimum e) in a libration cycle, their rate of change in q and e slows and then reverses. These objects ("returning IOCs") now decrease in perihelion and increase in eccentricity and migrate back toward the gap. Some of these objects pass back through the IOC region and the perihelion gap and return to the ETNO region. Here, they are again subject to perturbations from Neptune (Khain et al., 2020), and some are ejected from the solar system, but most again reverse direction and begin a new libration cycle as they cross the perihelion gap again and migrate through the IOC region toward a new, and often quite different, maximum q (minimum e).

There are several "returning IOCs," however, that end their libration cycle before crossing back over the gap. These IOCs reach their minimum q within the IOC region (near the location of Sedna and 2012 VP_{113} in q-e space) and subsequently begin to increase in q again without reentering the gap region (see Figure

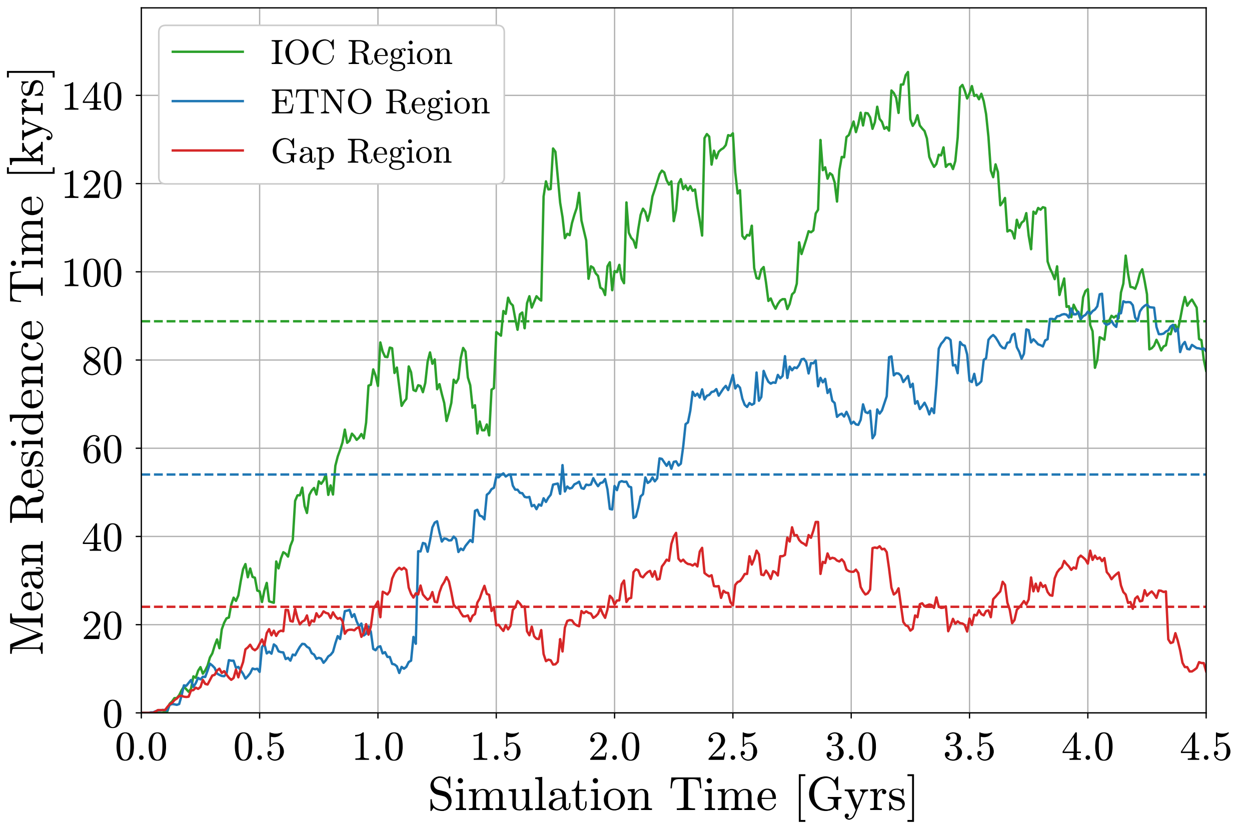


Figure 2.8 Mean residence time (the time particles spend in a particular region) for the ETNO region (blue), the IOC region (green), and the gap region (red) as a function of simulation time. Horizontal dashed lines show the average residence times of each population over the entire simulation. These data are from a simulation containing Planet X B19 outlined in Table 2.6. Note that the average residence time of objects in the IOC region is roughly five times higher than that of particles within the perihelion gap. This difference in mean residence time is the reason for the underdensity of particles within the gap.

2.11, an animation of Figure 2.5). This process keeps this subset of objects in the IOC region for $\sim 10^5$ yr, much longer than the time it takes for objects to traverse the gap. The difference in the time it takes for objects to cross the perihelion gap and how long objects spend at the peak of a libration cycle within the IOC region causes a measurable difference in the number density of objects in the gap region and the IOC region. This difference in mean residence time of objects within the IOC and gap regions is evident in Figure 2.8 where particles spend roughly five times longer on average in the IOC region than they do in the perihelion gap.

The mean residence times plotted in Figure 2.8 are computed by calculating continuous periods of time that each object spends in each region (e.g., a particle may spend 50 kyr in the ETNO region, followed by 20 kyr in the gap region, then 100 kyr in the IOC region, and subsequently the next 500 kyr beyond the IOC region, then 80 kyr in the IOC region again, and so on). Then, for each time in the simulation, the time particles spend in a particular region is averaged over the number of particles within that region at that simulation time. This metric allows for easy visualization of the difference in the amount of time particles spend in the gap compared to the surrounding regions, which is the reason for the underdensity of particles within the gap.

The slow passage of very high eccentricity objects ($0.9 \lesssim e < 1$) through the gap may make it difficult to determine whether or not these objects are temporarily in the gap, or if they are in the IOC region (see high e objects within the gap in Figure 2.11, an animation of Figure 2.5). Combined with the small sample size of known IOCs (Table 2.3), this slow movement makes determining the outer edge of the perihelion gap quite challenging. Additionally, the ratio of objects in the IOC region to objects within the gap is unknown and depends on the parameters of Planet X and the initial distribution of objects.

As more ETNOs and especially IOCs are discovered through outer solar system surveys, constraints on the perihelion gap and, consequently, Planet X will improve. Our dynamical simulations highlight that the IOCs are indeed the inner edge of a large distant population of objects. Many of these objects are unobservable with current ground and space based telescopes, but advances in observational facilities will continue to expand our capacity to probe the far reaches of our solar system. We predict as these new objects are discovered that some will be discovered within the perihelion gap, however, these "gap objects" will only be on the order of ~20% as numerous as the IOCs within 65 au $\lesssim q \lesssim$ 100 au given the majority of probable parameter combinations for Planet X.

2.5 Summary

In this work, we explore the outer solar system perihelion gap (see Figure 2.1) and its connection to the hypothetical Planet X. We find that (1) the gap is very unlikely to result from a realistic single, continuous distribution but is rather a transition region between two separate populations in the outer solar system, the ETNOs and the IOCs; (2) Neptune cannot form the perihelion gap on its own; and (3) Planet X can form the ETNOs, the IOCs, and the perihelion gap from a simple Kuiper-Belt-like initial distribution through a difference in the mean residence time of "gap objects" and "returning IOCs." We predict that, in the presence of Planet X, "gap objects" will be discovered, but that there will be roughly five times more IOCs discovered with 65 au $\lesssim q \lesssim$ 100 au than "gap objects."

We began our study by conducting a series of observational simulations. These simulations contain synthetic objects drawn from several different distributions in perihelion q or semimajor axis a within the ETNO/IOC (Extreme Trans-Neptunian Object/Inner Oort Cloud object) region (see Table 2.1). Using observational detection limits typical of modern outer solar system observational surveys, we determine which synthetic objects would be "observable" given these limits (see Table 2.2 and Figure 2.2). Next, we use the Poisson Probability and the K-S test to compare the likelihood of observing the perihelion gap given each distribution. Additionally, we fit a two-Gaussian model to the real ETNO/IOC distribution with Markov Chain Monte Carlo statistical fitting techniques (see Figure 2.3 and Table 2.5). We find that a bimodal distribution, i.e., a separation between the ETNO and IOC populations in the form of a perihelion gap between them, is a much better fit to the observed ETNO/IOC distribution compared to a unimodal distribution by several orders of magnitude (see Table 2.4). This suggests that the perihelion gap is not the result of observational bias and that the ETNOs and IOCs are two separate but related populations of objects.

With the assurance that the perihelion gap is a real feature of the outer solar system, we examine its relationship with Planet X through a series of dynamical simulations. In these simulations, we draw particles from similar parameter distributions to those used in the observational simulations, but initialized them in a Kuiper-Belt-like disk. These simulations were then integrated for 4.5 Gyr. When Planet X was not present, particles were scattered to high e by Neptune, but never moved into the ETNO/IOC region (see Figure 2.4). When Planet X was present, particles scattered to high e by Neptune are subsequently captured into resonance with Planet X and transported across the gap to high q along lines of roughly constant semimajor axis (see Figure 2.5). This process begins near a perihelion of 50 au coinciding with the inner edge of the perihelion gap.

Once objects cross the perihelion gap from the ETNO region to the IOC region and beyond, they begin secular oscillation in eccentricity and perihelion (see Figure 2.6). This oscillation is somewhat stochastic as the extremities of this motion are not fixed, but change each cycle. When these objects reach their minimum eccentricity in their oscillation cycle, their eccentricities begin to increase and their perihelia are reduced bringing them back toward the gap. A subset of these "returning IOCs," however, do not cross back over the perihelion gap, but rather reach their maximum eccentricity and minimum perihelion within the IOC region resulting in a difference in population density within the gap region and in the IOC region (see Figure 2.7). This difference in population density is augmented by the slow reversal of the direction of eccentricity-perihelion oscillation of the "returning IOCs" which causes them to have a much longer mean residence time within the IOC region than objects transitioning through the gap region (see Figure 2.8).

From these observations we find that Planet X neatly accounts for the outer solar system perihelion gap through secular resonance librations of "returning IOCs" and their larger mean residence times over objects within the gap region. Additionally, the connection between Planet X and the perihelion gap may serve as a useful constraint on the mass and orbital properties of Planet X, especially as more ETNOs and IOCs are discovered in the next several years.

2.6 Acknowledgments

We thank Renu Malhotra (UArizona) for insights regarding dynamical interactions, Ben Proudfoot (BYU) for advice on animation aesthetics, Garett Brown (UToronto) for suggestions on dynamical simulation integration energy errors, Patrick Tribbett (NAU) and Ty Robinson (NAU) for recommendations regarding MCMC fit probabilities, and our anonymous reviewer who provided helpful feedback which contributed to the improvement of this manuscript. This work was funded through NASA Planetary Astronomy grant 80NSSC18K1006. This work was made possible in part through the State of Arizona Technology & Research Initiative Program. Some of the computational analyses were run on Northern Arizona University's Monsoon computing cluster, funded by Arizona's Technology and Research Initiative Fund. This research has made use of NASA's Astrophysics Data System. This research has made use of data and/or services provided by the International Astronomical Union's Minor Planet Center.

Software: `astropy` (Price-Whelan et al., 2018), `corner` (Foreman-Mackey, 2016), `emcee` (Foreman-Mackey et al., 2013), `matplotlib` (Hunter, 2007), `numpy` (Harris et al., 2020), `pandas` (McKinney, 2010), REBOUND (Rein & Liu, 2012; Rein et al., 2019), `scipy` (Virtanen et al., 2020), `tqdm` (da Costa-Luis et al., 2020).

2.7 Appendix

2.7.1 MCMC Convergence and Distributions

To measure the convergence of our MCMC sampling run of the two-Gaussian distribution, we utilized autocorrelation times provided by the built-in functionality of the `emcee python` package (Foreman-Mackey et al., 2013). Over 15,000 steps, the maximum autocorrelation time for our five free parameters was 172 steps. So, our model was run for roughly 87 autocorrelation times and was sufficiently converged (runs $\gtrsim 50$ autocorrelation times have usually converged adequately; see `emcee` documentation[3]).

The first 344 steps in the sample chains were discarded as a burn-in period—time set aside to allow the walkers to approach a solution prior to sampling their distribution. This burn-in value was selected to be twice the maximum autocorrelation time (see `emcee` documentation[4]). Examination of the trace of the walker paths shows that this is sufficient time to approach a solution.

We compute the reduced χ^2 metric, with values near 1 indicating good fits, for the fit to be 0.27 (slightly over-fit, which is sufficient for this analysis) using the Kraft, Burrows, and Nousek method of calculating a Poisson confidence interval (Kraft et al., 1991) for the rolling histogram error on the real object distribution. The `poisson_conf_interval` function from the `astropy.stats python` package (Price-Whelan et al., 2018) was used to calculate these confidence values with a background noise of zero and a confidence level of 0.68 (1σ assuming Gaussian statistics).

In Figure 2.9 we provide a corner plot of the sample distributions of each of the free parameters used in our MCMC run. The values shown correspond to those reported in Table 2.5. The solid red lines indicate the parameter values from the correlated link with the highest likelihood value and all of these values are near the peaks of their corresponding distribution; however, the maximum likelihood value for σ_2 lies just outside of the lower quantile of its distribution. To confirm that this would not have an adverse effect on our results, we tested other values near these peaks using the observational model outlined in Sections 2.2.1 and 2.2.2. We find that nearby values (within the given uncertainties) also result in fits that are $\sim$7 orders of magnitude better than the single component continuous distributions we tested (see Table 2.4).

2.7.2 Dynamical Simulation Animations

Figure 2.10 is an animation of the particles mapped in Figure 2.4 and highlights how Neptune is unable to form ETNOs, IOCs, or "gap objects" from a simple Kuiper-Belt-like initial distribution. Figure 2.11 shows a similar animation, but for the particles mapped in Figure 2.5. This simulation includes Planet X B19 from

[3] https://emcee.readthedocs.io/en/stable/tutorials/autocorr/
[4] https://emcee.readthedocs.io/en/stable/tutorials/monitor/

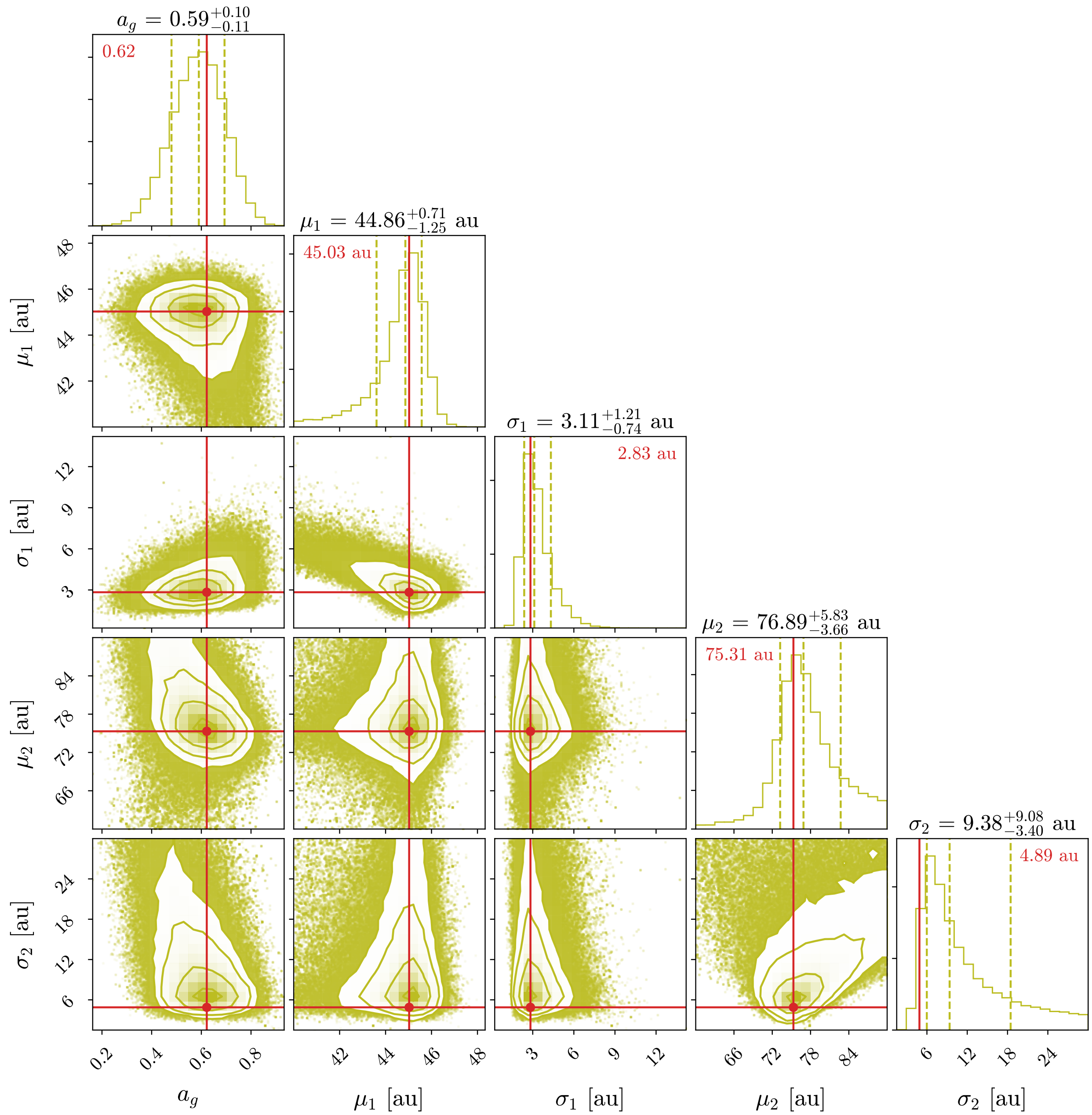


Figure 2.9 Corner plot of the MCMC sample parameter distributions. One-dimensional histograms for the sample chains of each free parameter are located at the top of each column. Two-dimensional histograms of each possible parameter combination are also shown. Values shown above each column are the model fit reported in Table 2.5 and correspond to the median of the distribution with uncertainties at the upper and lower 1σ quantiles (dashed vertical lines). Solid red lines and their corresponding values within the one-dimensional histogram subplots are the maximum likelihood $\mathcal{L}$ values in Table 2.5 and result from a single correlated link.

Table 2.6 (Batygin et al., 2019) and shows how objects transition from the ETNO region to the IOC region and beyond via the perihelion gap. "Returning IOCs" are also apparent in this animation. Figure 2.12 is an animation of the particles mapped in Figure 2.6 and contains Planet X T20 from Table 2.6 (Trujillo, 2020). This animation highlights secular resonances of objects with Planet X and their movement along nearly constant lines of semimajor axis.

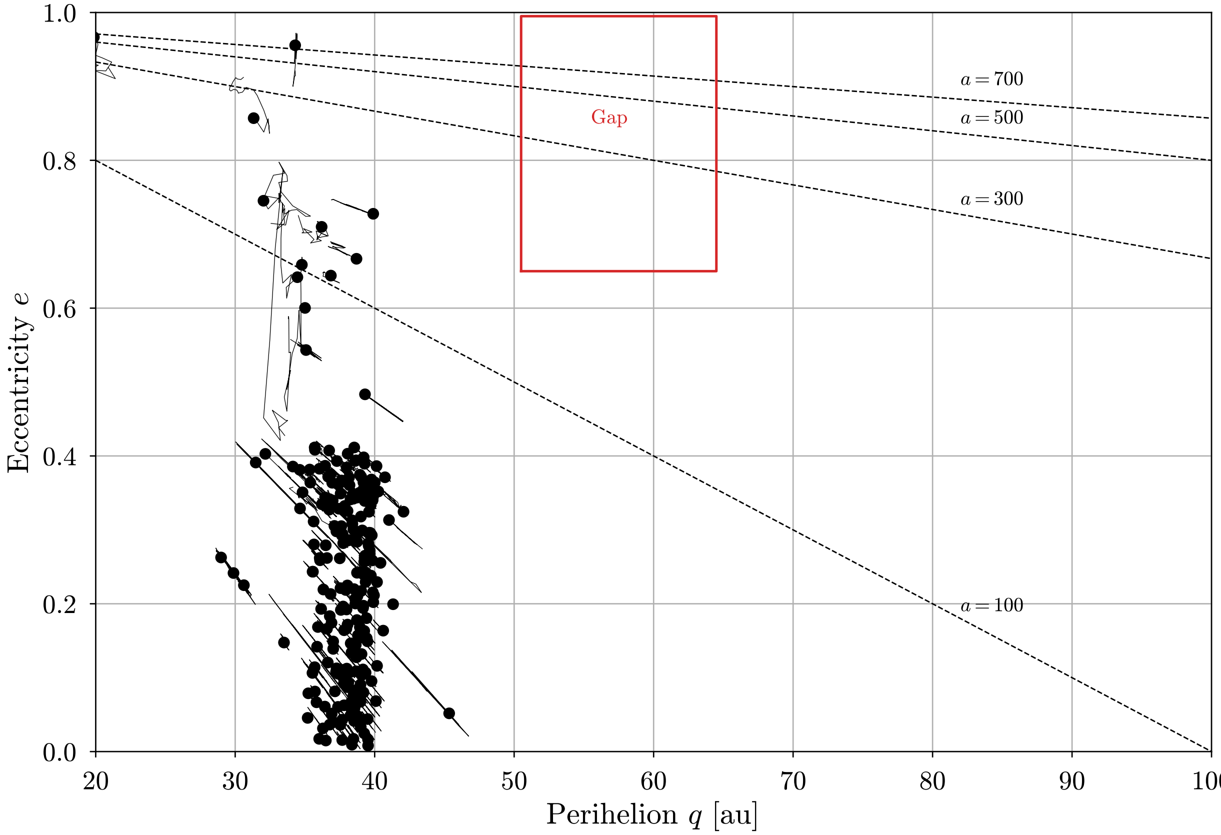


Figure 2.10 Animation over 4.5 Gyr of integration of test particles (black points) in eccentricity and perihelion with a Kuiper-Belt-like initial distribution in the presence of the Sun, the known giant planets, and no Planet X. The animation time step is 10 Myr and the trails connected to the test particles track their motion over the previous 250 Myr in the simulation. As the simulation progresses, particles are scattered to high eccentricities ($e > 0.6$) through interactions with Neptune. Many particles are captured into mean-motion resonance with Neptune, as expected, and oscillate along diagonal lines of constant semimajor axis. Note that no particles migrate to the ETNO, gap, or IOC regions. The retained Kuiper-Belt-like initial distribution, particles scattered to high e, particles in mean-motion resonance, and the lack of particles near the gap are all shown in the single-frame version of the figure. The average particle densities over this simulation are shown as a heatmap in Figure 2.4.

(An animation of this figure is available.)

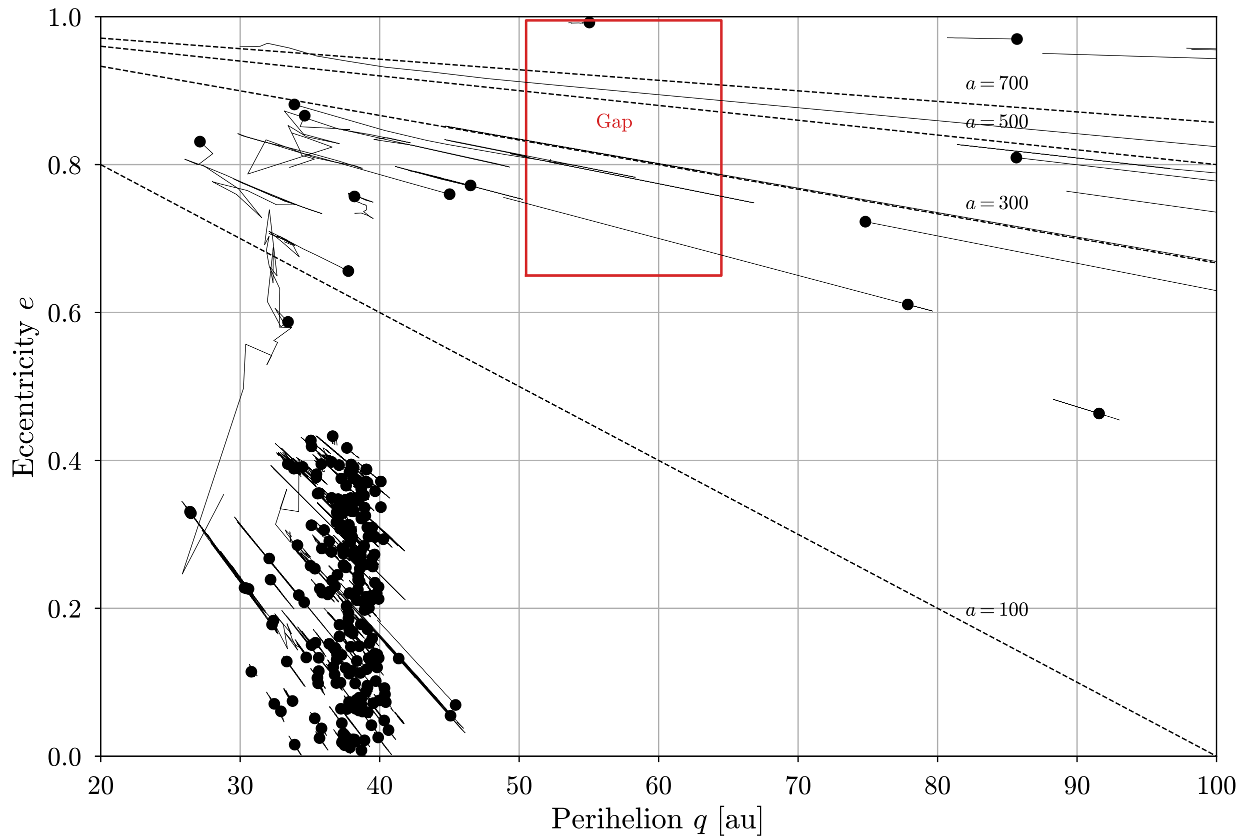


Figure 2.11 Animation over 4.5 Gyr of integration of test particles (black points) in eccentricity and perihelion. As Figure 2.10, with the addition of Planet X B19 from Table 2.6 (Batygin et al., 2019). The animation time step is 10 Myr and the trails connected to the test particles track their motion over the previous 250 Myr in the simulation. Throughout the simulation, particles are scattered to high eccentricity by Neptune (as in Figure 2.10), however, many of these particles are now captured into resonance with Planet X and migrate through the ETNO region. Once these particles approach the gap, they migrate more quickly and rapidly transition through the gap into the IOC region and continue beyond $q = 100$ au. Particles with eccentricities $e \gtrsim 0.95$ migrate through the gap much more slowly. Additionally, many particles return to the IOC region from beyond $q = 100$ au. Some of these particles cross back over the perihelion gap into the ETNO region, but others reach the peak of their oscillation cycle within the IOC region. The retained Kuiper-Belt-like initial distribution, particles scattered to high e, quickly migrating particles in resonance with Planet X, slowly migrating high e particles in the gap region, and "returning IOCs" are all shown in the single-frame version of this figure. The average particle densities over the simulation are shown as a heatmap in Figure 2.5.

(An animation of this figure is available.)

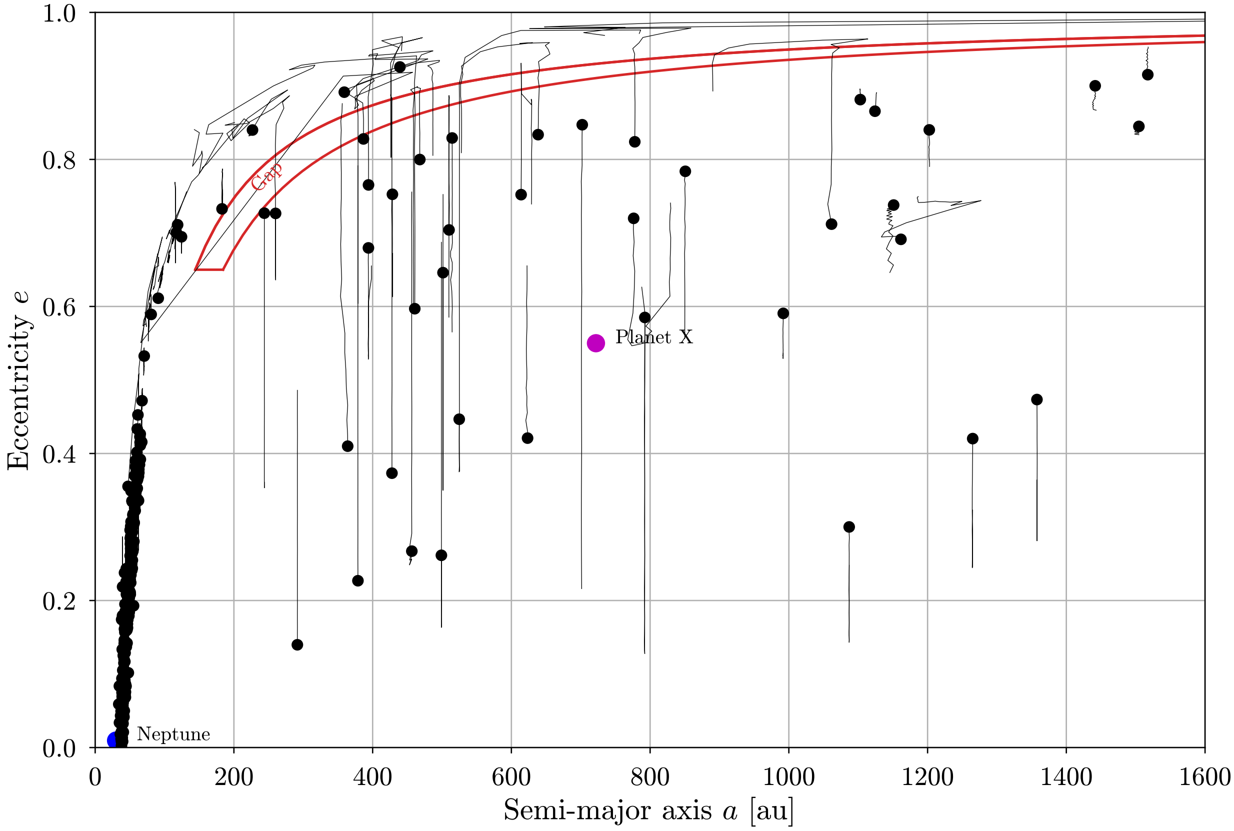


Figure 2.12 Animation over 4.5 Gyr of integration of test particles (black points) in eccentricity and semimajor axis with a Kuiper-Belt-like initial distribution in the presence of the Sun, the known giant planets, and Planet X T20 from Table 2.6 (Trujillo, 2020). The animation time step is 10 Myr and the trails connected to the test particles track their motion over the previous 250 Myr in the simulation. Over the course of the simulation, particles are scattered to high eccentricity by Neptune which also increases their semimajor axes. Many of these particles are captured into secular resonances with Planet X, causing them to oscillate in eccentricity along lines of roughly constant semimajor axis. Occasionally, particles undergo large shifts in semimajor axis during this oscillation period, particularly upon re-entry of the ETNO region where they are again subject to the gravitational influence of Neptune. Additionally, "returning IOCs" are visible in the simulation, with some crossing the gap toward high e before reaching their peak eccentricity, and others reaching their peak e within the IOC region. The retained Kuiper-Belt-like initial distribution, particles scattered to high e, particles in secular resonance with Planet X, particles undergoing "resonance hopping," and "returning IOCs" are all shown in the single-frame version of this figure. The average particle densities over this simulation are shown as a heatmap in Figure 2.6.

(An animation of this figure is available.)

Chapter 3

Exoastrometry Optimization

This chapter contains work to be published in the Astronomical Journal as soon as final analyses are complete. As such, it is presented in manuscript form and the submitted manuscript will be similar to this version.

Orbit Determination Optimization for Directly Imaged Exoplanets

William J. Oldroyd[1] and Tyler D. Robinson[1]

[1]*Department of Astronomy & Planetary Science*

Northern Arizona University

PO Box 6010, Flagstaff, AZ 86011, USA

To be submitted to AJ

ABSTRACT

Constraints on orbital parameters are essential for understanding the architecture of exoplanetary systems and for constraining the potential habitability of rocky exoplanets. However, the most efficient observing strategy for constraining orbital parameters is often unclear. This presents a major challenge for space-based exoplanet direct imaging surveys where significant mission time may be spent revisiting systems for astrometric observations. In this work, we explore several methods of improving constraints on the orbits of directly imaged exoplanets by optimizing the timing of revisit observations. Optimization is achieved through the use of orbital integration in conjunction with Markov Chain Monte Carlo statistical sampling. The sampling tool predicts the time of observation along a best-fit orbit to maximize constraints on key orbital parameters. The performance of this approach is studied via application to synthetic observations.

Corresponding author: William J. Oldroyd
woldroyd@nau.edu

We find that these methods greatly enhance the precision of orbital fits over previously explored fixed cadence observations, saving roughly 20% to 40% of total revisit observation time for a simple face-on Earth-like planet. In the future published version of this work, we will also explore a wide variety of orbital configurations to examine the effectiveness of these methods in the general case.

Keywords: Exoplanets (498) – Exoplanet astronomy (486) – Exoplanet detection methods (489) – Direct imaging (387) – Irregular cadence (1953)

3.1 Introduction

Orbits are key to understanding systems and habitability. Studying the architecture of exoplanet systems has provided constraints on planet formation models and has revealed whole categories of worlds not found in the solar system, such as hot Jupiters and Superearths (Hatzes, 2016; Winn & Fabrycky, 2015; Kopparapu et al., 2013). Furthermore, the potential for surface habitability on rocky exoplanets is thought to be strongly controlled by insolation, hence, one of the most fundamental parameters for Habitable Zone models is orbital distance (Kopparapu et al., 2013).

The 2020 Decadal Survey on Astronomy and Astrophysics[1] prioritized the development of a 6-meter class space telescope capable of detecting and characterizing potentially Earth-like exoplanets orbiting within the Habitable Zone of nearby stars. A major challenge for these space-based exoplanet direct imaging surveys is target revisits. After an initial detection of a planetary system, revisits ensue to confirm the planetary nature of a target (through co-moving trajectory with the central star) and/or to constrain planetary orbital parameters. Critically, the determination of planetary orbits then identifies potential targets for prioritized spectral follow-up, such as would be the case for a world orbiting in its host star's habitable zone. As explored by the relevant precursor studies of the Habitable Exoplanet Observatory (HabEx) (Gaudi et al., 2018) and Large UltraViolet-Optical-InfraRed Surveyor (LUVOIR) (Roberge & Moustakas, 2018) concepts, the number of revisits for a given system can be as large as 5–6, implying the detection and revisit survey could consume roughly 25% of total mission time (Gaudi et al., 2020; The LUVOIR Team, 2019).

Recently, several developments in orbit characterization efficiency of directly imaged exoplanets have been made. One of the most groundbreaking developments is a Bayesian Monte Carlo method for determining orbital elements from astrometric observations of long-period orbits titled "Orbits for the Impatient" (OFTI) Blunt et al. (2017) (and building off of the work of Konopacky et al., 2016). OFTI has been shown to reproduce equivalent results to Markov Chain Monte Carlo (MCMC) approaches, but with orders of magnitude greater computational efficiency. This approach is particularly useful for exoplanets with low fractional

[1] `https://www.nationalacademies.org/our-work/decadal-survey-on-astronomy-and-astrophysics-2020-astro2020`

orbit coverage, as is the case for many, and possibly most recently-imaged wide-separation exoplanets. Additionally, Blunt et al. (2020) have integrated OFTI together with MCMC sampling and incorporation of radial velocity measurements into the orbit fitting `orbitize! python` package, which greatly enhances the accessibility and effectiveness (Morgan et al., 2021) of these techniques. Similar methods are also used to determine the orbits of binary solar system objects (Grundy et al., 2008) and stars orbiting black holes (Ghez et al., 2008).

In addition to advances in computational tools for orbit characterization, it has been shown that Earth-like exoplanets (terrestrial worlds orbiting within the habitable zone of their host stars) could have their orbits constrained to within 5 to 10% uncertainty in semi-major axis in just 3 to 4 observations using a fixed cadence with sufficient time between observations (Guimond & Cowan, 2019). Preliminary results in the HabEx concept study also noted that three "well-spaced" detections can constrain orbital semi-major axis to within 10% (Gaudi et al., 2020). One potential method that has been proposed as a way to improve orbital constraint per observation is dynamical cadence scheduling which would allow for mission constraints, such as preferentially observing nearby systems, to be more carefully considered when determining revisit times (Morgan et al., 2021).

In this work, we outline three methods for determining the optimal or most efficient revisit time for constraining the orbit of a directly imaged exoplanet. We then compare the effectiveness of these methods with a fixed cadence approach, as has been commonly explored in previous modeling studies. In the final version of this manuscript, we will explore the limiting cases of these models over several example exoplanetary systems in various sky plane orientations. Finally, we discuss the implications for applications of these techniques and ways to improve upon and integrate the models into existing and future computational tools.

3.2 Methods

In order to determine the optimal revisit timing for orbital constraint of a directly imaged exoplanet, we generated synthetic exoplanet observations, used Markov Chain Monte Carlo (MCMC) statistical fitting coupled with N-body integration to find a distribution of orbits that fit the observations, and then applied three separate techniques for calculating the optimal revisit timing. These methods are (1) the swath method, (2) the cluster method, and (3) the comprehensive method. Here, we keep our presentation of our optimization approach general, and later applications will explore specific case studies.

Synthetic observations were generated using provided values for the host star mass, planet mass, planetary orbital elements, and the number and timing of initial observations. We then used the `python` implementation of the `REBOUND` `mercurius` N-body integrator (Rein & Liu, 2012; Rein et al., 2019) with a time step of 0.01

yr (sufficient for most orbits with $a \geq 0.16$ au, including Earth-like exoplanets; see Hernandez et al., 2022) to integrate the system over the selected time span. At each observation time, we calculated the X and Y positions of the planet relative to the star and added error bars on the measurements as given by the Rayleigh criterion (see Rayleigh, 1879) and accounting for the Signal to Noise Ratio

$$\sigma_{X,Y} = 1.22\frac{\lambda d}{D}\frac{1}{\mathrm{SNR}}(206,265 \text{ au pc}^{-1}), \tag{3.1}$$

where $\sigma_{X,Y}$ is the positional error of the planet in au, λ is the observation wavelength in meters, d is the distance to the system in parsecs, D is the diameter of the telescope in meters and SNR is the Signal to Noise Ratio for which we adopted a value of 10. For examples in this section, we used an error of 0.035 au or[2] 3.5 mas, corresponding to a distance of 10 pc, a telescope diameter of 4 m, and an observation wavelength of 550 nm; appropriate values for HabEx (see Gaudi et al., 2020).

Using these synthetic observations, we derived a distribution of orbits consistent with these observations using the MCMC `emcee python` package (Foreman-Mackey et al., 2013) by calculating a posterior probability distribution of the exoplanet orbital parameters. As a forward model through which the orbital element combinations generated through the MCMC sampler were run, we again used the `REBOUND mercurius` integrator with a 0.01 yr time step. The X and Y positions of the planet given the projected orbit of each MCMC sample were then derived for each observation time. The likelihood, a measure of how well a given sample fits the data (or more formally, how likely a particular measurement is given that your model accurately describes the system, see Bonamente (2017) Chapter 1), was calculated as

$$\mathcal{L} = -\frac{1}{2}\sum_{t=0}^{t_f}\frac{(X_s - X_d)^2 + (Y_s - Y_d)^2}{\sigma^2_{X,Y}}, \tag{3.2}$$

where t is the observation time, t_f is the final observation time, X_s and Y_s are the coordinates from the MCMC sample, X_d and Y_d are the coordinates from the synthetic data in au, and the error $\sigma_{X,Y}$ as calculated in Equation 3.1.

Priors for sample runs were selected to be broad in order to encompass known exoplanetary systems and are outlined in Table 3.1. For the sample runs, 60 walkers—pseudo-random sampler chains that preferentially change their parameters towards higher likelihoods—were used. The walkers were initialized at values near the known orbital parameters of the synthetic data and each had a small random offset to these initial values. The model was then run for 50,000 steps. Although this was insufficient time for the model to fully converge, we detected no significant differences between runs with 50,000 steps and runs with over

[2] $\theta = \arctan(\sigma_{X,Y}/(d(206,265 \text{ au pc}^{-1})))(3.6 \times 10^6 \text{mas}/^\circ)$

Table 3.1. MCMC Parameters and Priors

Parameter	Priors
semi-major axis a	0.002 au $\leq a \leq 650$ au
eccentricity e	$0 \leq e < 1$
inclination i	$0^\circ \leq \sin(i) < 180^\circ$
argument of perihelion ω	$0^\circ \leq \omega < 360^\circ$
longitude of ascending node Ω	$0^\circ \leq \Omega < 360^\circ$
initial true anomaly f	$0^\circ \leq f < 360^\circ$

Note. — Semi-major axis priors were selected based on known exoplanetary systems with extreme a values (Bailes et al., 2011; Bailey et al., 2014). Inclination priors use $\sin(i)$ rather than i due to observing geometry.

10^6 steps other than run time, which was significantly shorter with fewer steps. A burn-in period—time for the walkers to sufficiently converge before counting them towards the resulting posterior probability parameter distributions—were taken to be twice the maximum autocorrelation time of the walkers (see `emcee` documentation[3]).

After the MCMC run was completed, we calculated the maximum likelihood correlated sample from the posterior probability distributions and took that to be our best-fit orbit. The synthetic observations, best-fit orbit, and the distribution of sample orbits for a single MCMC run with an Earth-mass planet orbiting a solar-mass star with $a = 1$ au, $e = 0$, and $i = 0^\circ$ in a face-on configuration are shown in Figure 3.1. Using the results from the MCMC fit, we then evaluated the time at which we could optimize the orbital constraint of the system using three separate techniques.

3.2.1 The Swath Method

The first technique we employed for determining optimal revisit observing times involves a comparison of the swath of sample orbits that are interior to the best-fit orbit with those that are exterior to the best-fit orbit. The hypothesis for this method is that when these sample orbit groups diverge the most, in on-sky separation, from each other over one orbital period (of the best-fit orbit), that is the time where the true orbit of the planet is least constrained. Hence, follow-up observations at that time will provide the best orbital constraint.

To implement this method, we randomly selected 1,000 sample orbits from our posterior probability distributions and computed the X-Y distance between each orbit and the best-fit orbit at 100 points along the orbits, sampled in mean anomaly M. Sample orbits were then sorted into two groups, on average

[3] `https://emcee.readthedocs.io/en/stable/tutorials/monitor/`

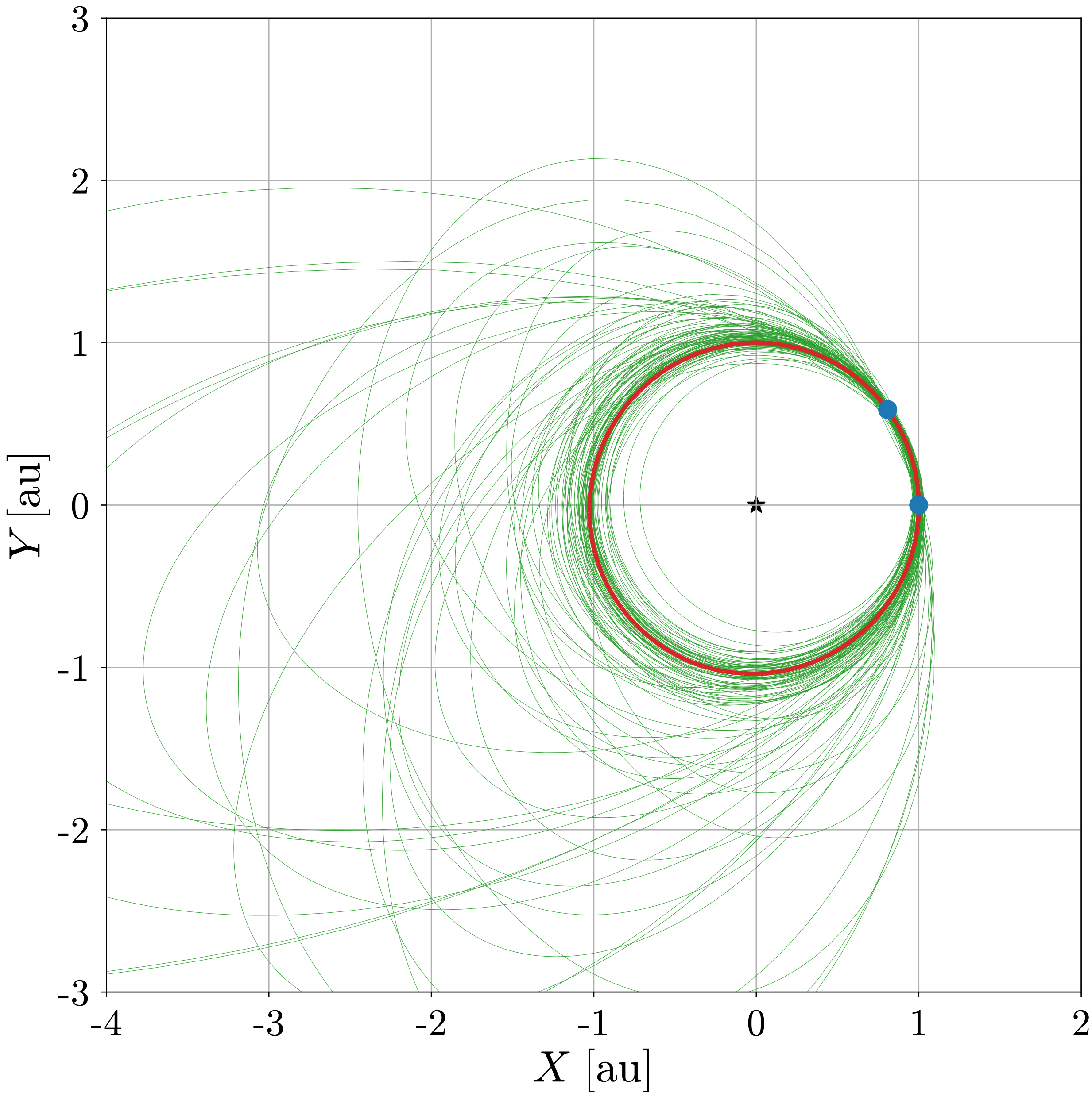


Figure 3.1 Example best-fit orbit (thick red) and sample orbits (thin green) pulled from the posterior probability distribution of an MCMC run for an Earth-mass planet orbiting a solar-mass star with $a = 1$ au, $e = 0$, and $i = 0°$ in a face-on configuration. The host star is represented by a black star at the origin. Synthetic observations of the planet are shown as blue points. The errorbars on these measurements are smaller than the points.

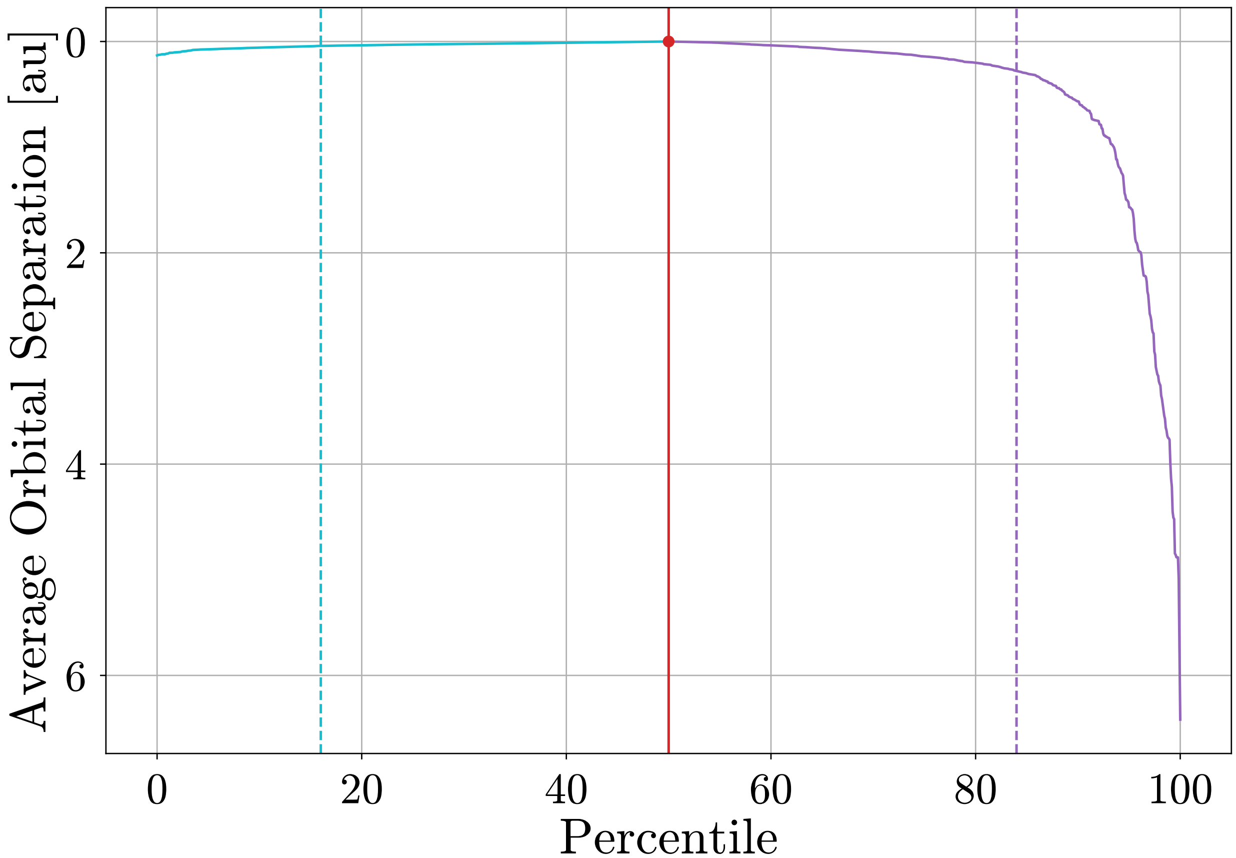


Figure 3.2 Sorted average orbital separation of the sample orbits from the best-fit orbit shown in Figure 3.1. Orbits that are on average interior to the best-fit orbit (red point) are shown in cyan on the left and orbits that are on average exterior to the best-fit orbit are in purple on the right. The dashed vertical lines are 1σ from the best-fit orbit and indicate the orbits that are selected as representative for the inner and outer sample orbit distributions.

interior or exterior to the best-fit orbit, forming a orbit swath distribution. Representative orbits located at 1σ (assuming Gaussian statistics) from the peak of this distribution (the best-fit orbit) were then defined (see Figure 3.2).

By calculating the X-Y separation between the inner and outer representative sample orbits, we determined the mean anomaly at which they had maximum divergence within one orbital period of the best-fit orbit (see Figure 3.3). Because the sample orbits were randomly selected, we repeated this process 100 times and constructed a distribution of times at which pairs of representative sample orbits achieved maximum separation (see Figure 3.4). The peak of this distribution is the returned optimum timing for this method, with values 1σ below the peak representing the uncertainty on this timing (see Figure 3.3).

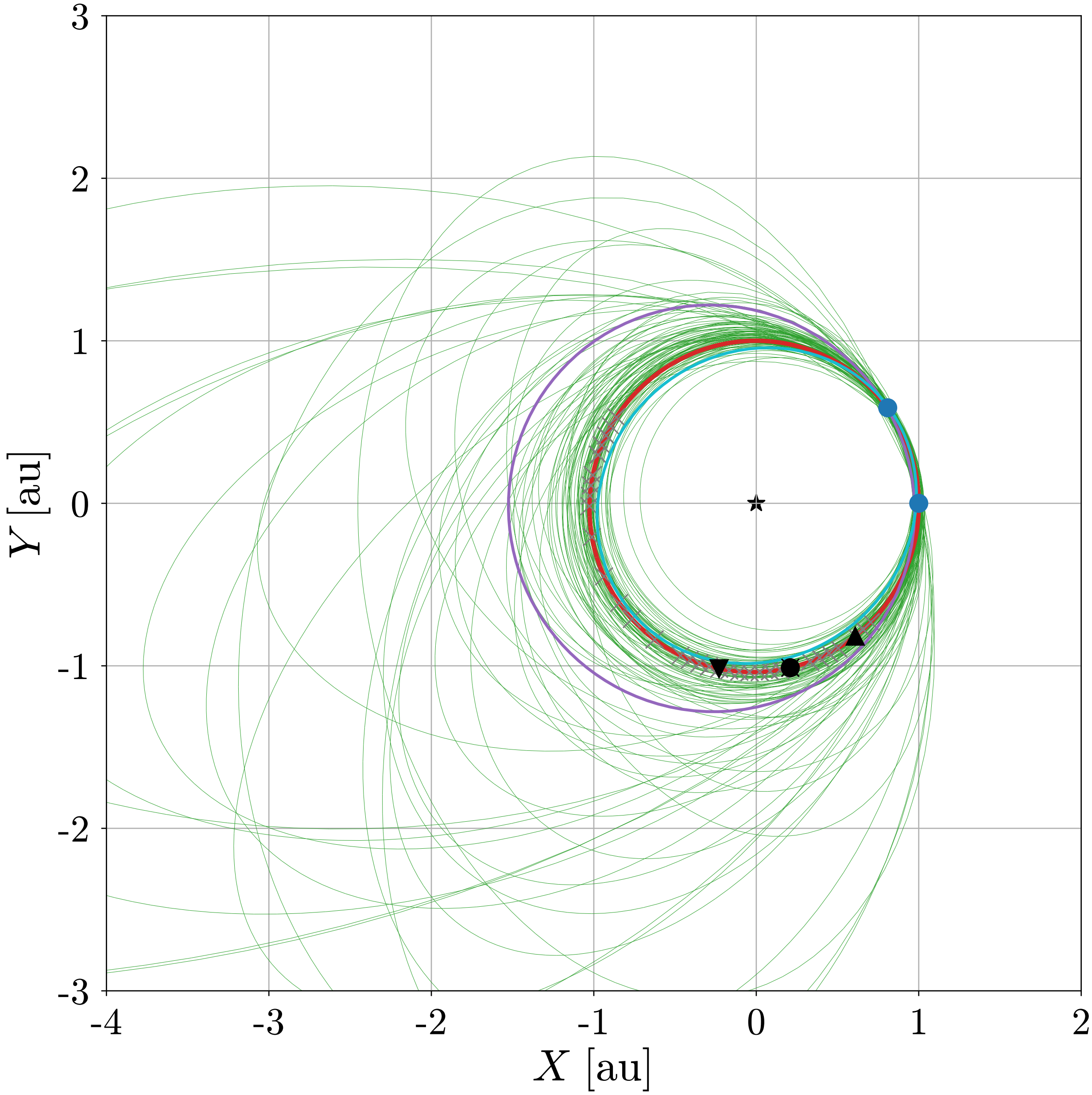


Figure 3.3 The swath method. Best-fit orbit, data points, and sample orbits shown as in Figure 3.1. Representative sample orbits, shown in cyan and purple, correspond to the vertical dashed lines in Figure 3.2. The point along the best-fit orbit corresponding to the time at which these representative sample orbits have maximum separation is shown with a black X (overlapping the black filled circle). Locations of maximum separation for other iterations of the model are shown with grey Xs. The black filled circle indicates the peak of this best observing time distribution, generated by repeating the swath method 100 times, and the black triangles represent the 1σ uncertainties on this value (see Figure 3.4).

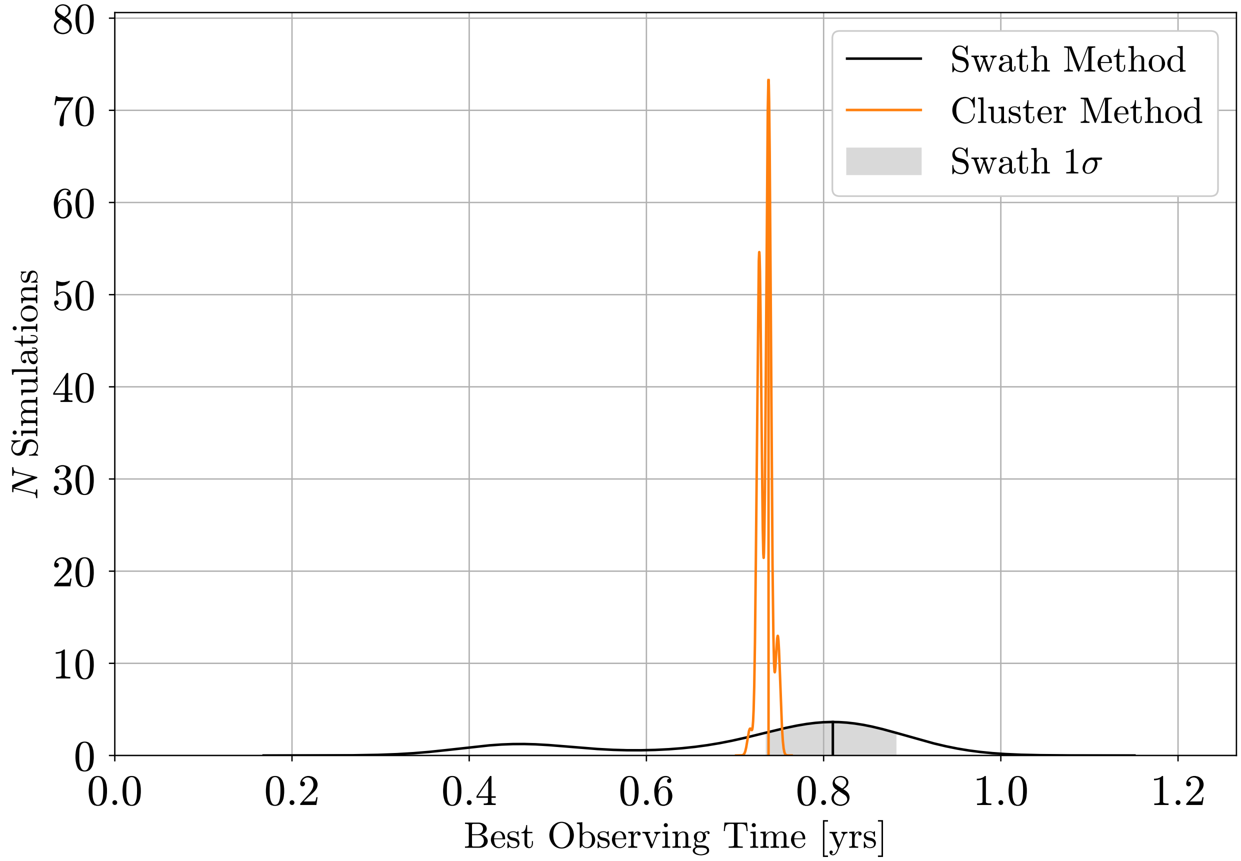


Figure 3.4 Distribution of best times for making the next observation returned by the swath (black) and cluster (orange) methods for the system shown in Figure 3.1. The peak values of these distributions are highlighted with a vertical line and the 1σ uncertainty for the swath method is shown in grey. The swath method tends to produce broad distributions for the optimum timing, which provides an estimation of how effective an observation will be at any given time along the orbital period. The cluster method typically results in a sharp peak, highly favoring a small window of time. Hence, the broad distribution of the swath method compliments the sharp peak of the cluster method and together they provide high confidence in an optimal observing time as well as information on the effectiveness of observing at other times.

3.2.2 The Cluster Method

Similar to the swath method, the cluster method operates under the hypothesis that the time where the sample orbits have maximum divergence over one best-fit orbital period is where the orbit is least constrained, and observations at that time will yield maximum constraint. Rather than comparing representative sample orbits, however, the cluster method focuses on the spread of all orbit samples throughout their orbits.

Over the same 1,000 randomly selected sample orbits, the X-Y distance between each orbit and the best-fit orbit were computed at 100 equally divided points in mean anomaly along the best-fit orbit. The sum of these distances were then taken to be the total spread of the cluster of orbits. The time at which this cluster spread was the largest, within one best-fit orbit period, was then taken to be the optimal revisit time for that iteration (see Figure 3.5). As with the swath method, this process was repeated 100 times and the peak of the resulting revisit time distribution (along with 1σ uncertainties) was the returned best observing time for the method (see Figure 3.4).

3.2.3 The Comprehensive Method

Our final method for calculating optimum revisit observation time was not based on a hypothesis derived from divergence of sample orbits. Instead, the comprehensive method approach to the problem involved assuming the best-fit orbit to be the true orbit, making future synthetic observations, and determining a new best-fit orbit based on these observations.

First, we divided the best-fit orbit into 100 points equally spaced in mean anomaly. For each point, we made a synthetic observation as if the best-fit orbit were the true orbit of the system. The synthetic observations were generated using `REBOUND` in the same manner as the original synthetic planet observations. Then, a separate MCMC run was carried out for each new synthetic point along with the original synthetic data. These MCMC runs were carried out in parallel on the Northern Arizona University high performance computing cluster, *Monsoon*.[4]

Uncertainties on each of the orbital elements were carefully tracked for each set of simulated observations. The resulting distribution of best-fit parameters was then processed to remove outliers, by replacing them with the median of nearby points. The distribution was then smoothed using a rolling histogram approach with a bin radius of seven timesteps. This outlier removal process reduced deviations resulting from the probabilistic nature of MCMC sampling and was more computationally feasible (as this method requires on the order of 100 times more compute time than the swath and cluster methods) than repeating the process many times to get an aggregate distribution as in the previous methods. Once the data were cleaned

[4] *Monsoon* has over 4000 cores, peak CPU performance above 220 teraflops, more than 2 PB of storage space, and is free for research and academic use for NAU faculty, staff, and students. https://in.nau.edu/hpc/details/

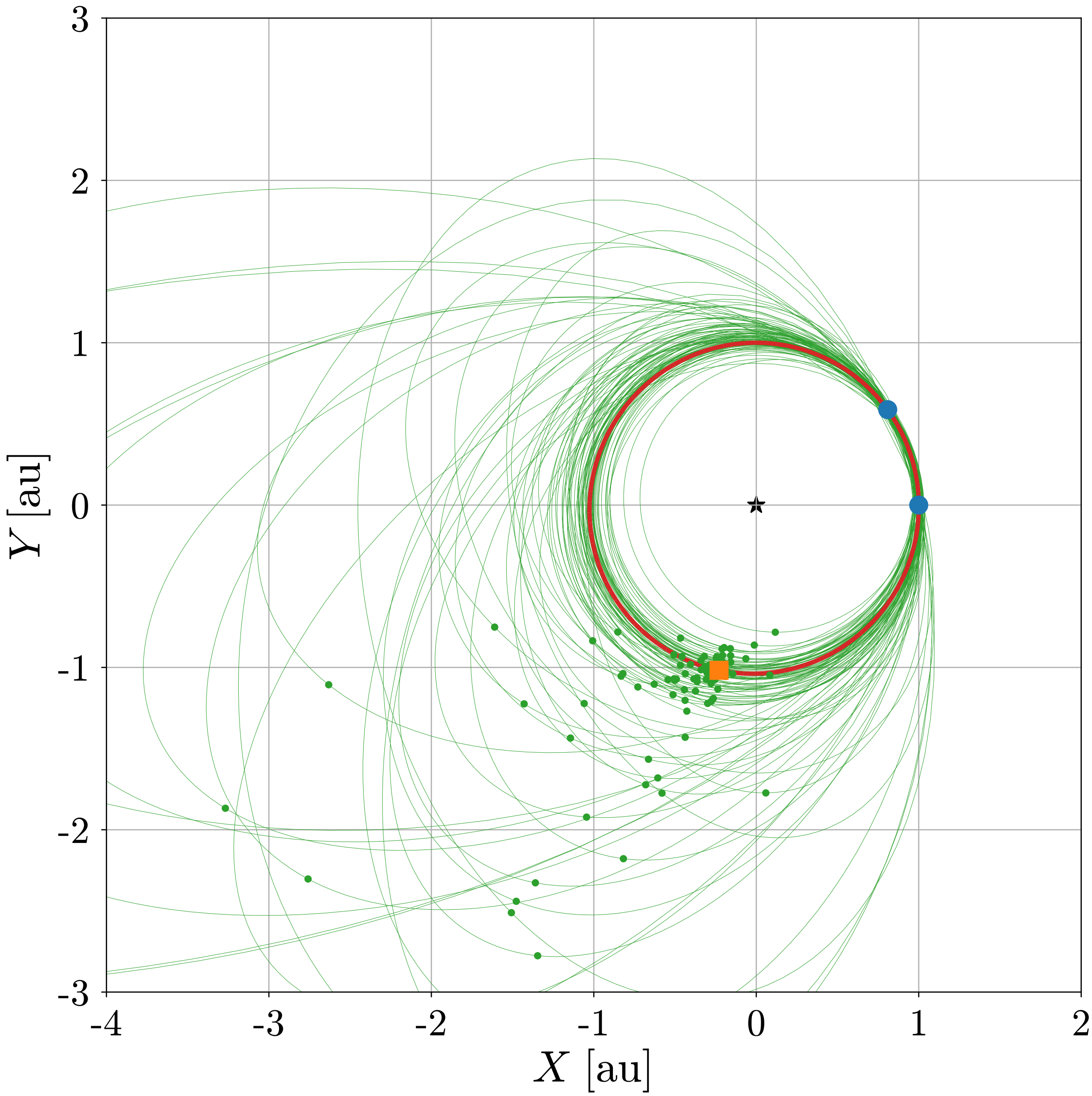


Figure 3.5 The cluster method. Best-fit orbit, data points, and sample orbits are shown as in Figure 3.1. For this group of sample orbits, locations along the orbits corresponding to the mean anomaly at which the cluster of orbits reaches maximum divergence, within one period of the best-fit orbit, are shown as green points. The peak of the distribution for all runs of the cluster method on this system (see Figure 3.4) is shown as an orange square. The uncertainties on this value are smaller than the square.

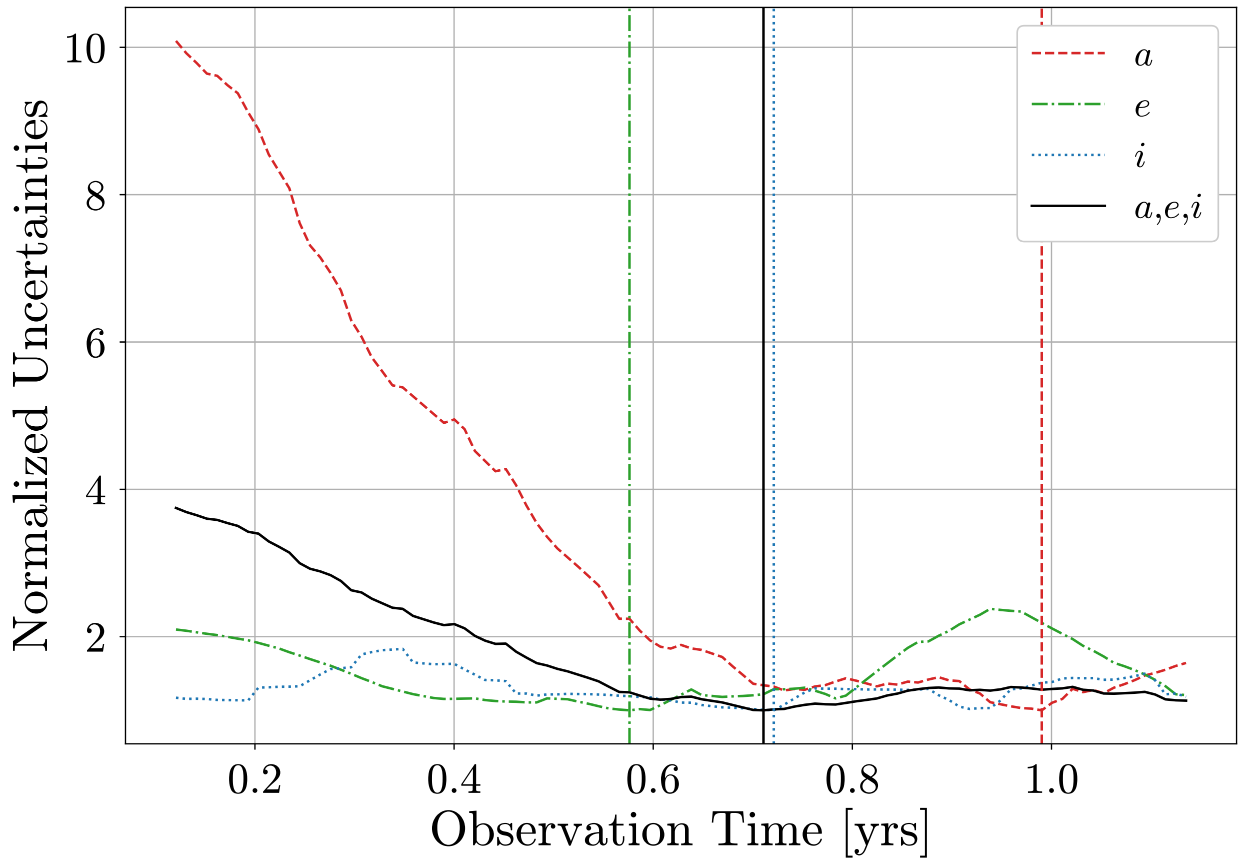


Figure 3.6 Normalized uncertainty distributions from the comprehensive method for a (dashed red), e (dashed dotted green), i (dotted blue), and the average of these 3 parameters (solid black) for the system shown in Figure 3.1. Note the good agreement between the optimum timing derived from the three-parameter average and the times obtained from the other two methods shown in Figure 3.4. It is typical when using this method for the semi-major axis uncertainty to decrease with time up to a certain point at which it roughly levels off and similar behavior is seen in the three-parameter average.

and smoothed, the uncertainties on each element were then normalized to their lowest value, for ease of comparison. A cumulative orbital element uncertainty was also calculated, taking into account a, e, and i uncertainties. Simplified results with the same planet configuration as in Figure 3.1 are shown in Figure 3.6, where the black curve represents the combined orbital uncertainty and the black vertical line designates to optimal observing time.

3.3 Results

We find that our methods outperform the standard fixed cadence approach (in some cases dramatically so) as shown in Figure 3.7. The top panel highlights that all of our methods are excellent at greatly reducing the uncertainty on the semi-major axis by the 3rd observation. In the bottom panel, the swath, cluster, and comprehensive three-parameter average methods all perform comparably, with the comprehensive three-parameter average reducing uncertainty in slightly less time than the other methods. This accelerated uncertainty reduction is the result of reaching a threshold time beyond which semi-major axis uncertainty is near its minimum value over the best-fit orbit period. Additionally, the revisit time must be before the time when uncertainty on the eccentricity can increase as it approaches one orbital period from previous observations. Near one orbital period after the previous observation, additional observations will provide relatively little information about the orbital eccentricity (see Figure 3.6). It is perhaps unsurprising then that both the swath and cluster methods predict optimum revisit times near the same time as the comprehensive 3 parameter average as all three methods seek to minimize overall orbit uncertainty.

The comprehensive method focused on minimizing the semi-major axis uncertainty is a particularly interesting case. Initially, it outperforms the other methods by selecting to observe at almost exactly one orbital period after the first observation. Because period and semi-major axis are intrinsically related through Kepler's 3rd law $P^2 \propto a^3$ (Kepler, 1619), precisely determining the orbital period in this manner would naturally provide excellent constraint on the semi-major axis. However, after this 3rd observation, the rate at which further observations using this method improve the uncertainty in a is reduced. Because the only consideration for this method is a single orbital element, the other elements incur sub-optimal constraint which, in turn, reduces the capacity of future observations to continue to improve orbital uncertainties (see Figure 3.7). So, while it may initially appear logical to focus optimization solely on the semi-major axis, this results in loss of overall orbit constraining capability.

3.4 Discussion and Future Work

From our simple example, we are able to constrain the semi-major axis to within acceptable parameters ($<$ 10% uncertainty) with just three observations. Though some fixed cadences may result in a similar result (albeit with larger final uncertainty) the best fixed cadence timing is not known *a priori*, hence, many cadences, including the one used in this work, require more than three observations in order to sufficiently constrain the orbit of the planet.

Revisit observations need not necessarily yield detections. Complications such as the planet being located within the inner working angle of the telescope, or becoming too faint due to phase function effects may result in a lack of detections for a given observing epoch. There is orbital information embedded in these non-detections and future models should take these factors into account when determining optimum observing time, however, the use of these constraints is outside the scope of this work.

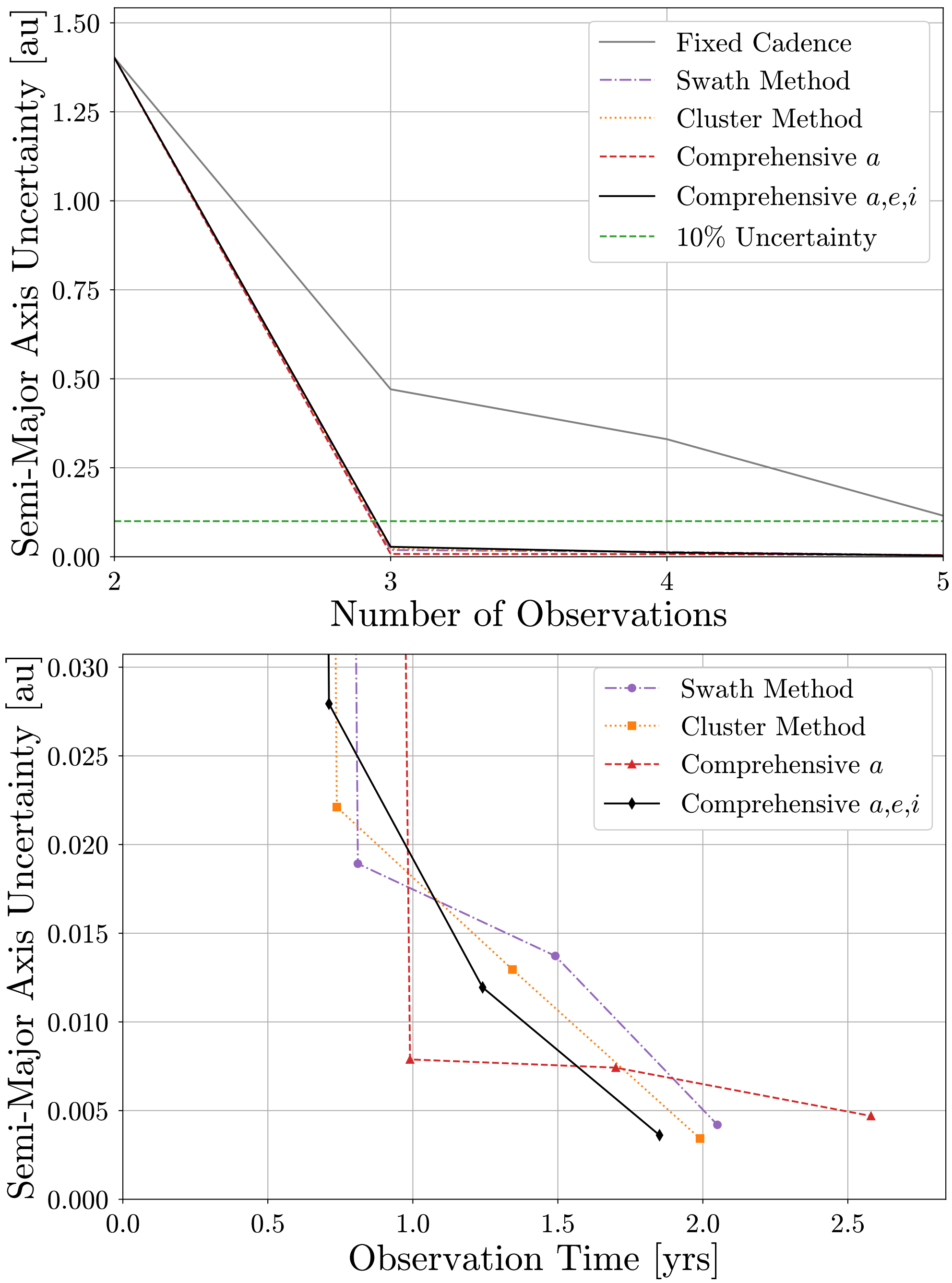


Figure 3.7 **Top:** semi-major axis uncertainty as a function of the number of observations using a fixed cadence of 0.1 yr (solid gray), the swath method (dot dashed purple), the cluster method (dotted orange), the comprehensive method with each observation taken at the time of minimum semi-major axis error (dashed red), and the comprehensive method with observations taken at the minimum of the three-parameter average (over a, e, and i; solid black). All three of our proposed methods greatly outperform the fixed cadence observations. The horizontal dashed green line indicates a semi-major axis uncertainty of 0.1 au (10% of the true a value). **Bottom:** semi-major axis uncertainty as a function of observation time for an Earth-mass planet orbiting a solar mass star at 1 au (1 year orbital period) in a face-on orbital configuration. Only observations three through five are shown. This is effectively a magnified view of the tightly clustered model results from the top panel, but with the added dimension of the observation time. Note the similarities in effectiveness between the models and their ability to constrain the orbit over a fixed cadence.

The results above provide a promising avenue forward for dramatically reducing the time required in a direct imaging survey for exoplanet orbit characterization. For example, if a mission design assumed 5 total visits for accurate constraints on exoplanet orbits at fixed cadence, whereas the optimization procedure given above would require only 3–4 total visits, then ∼20–40% of survey time could be saved.

Observation timing optimization approaches that include the option to solve for multiple planets within a system and/or account for mission constraints — such as scheduling, inner working angles, and limiting magnitudes — in addition to the methods presented here, would further improve the efficiency of current and future surveys. Additionally, broader suites of simulations analyzing the effectiveness and limitations of these methods could provide information on trends for suitable observing times, other than the most optimum time, that give sufficient orbital constraint for a given system.

The majority of this work focuses on determining the optimal revisit timing from the third observation on. The second observation is more difficult to constrain, since only one data point (from the initial observation) provides limited orbital information. One method for determining when to perform an initial revisit is to assume that the distance between the planet and the star on the plane of the sky is the semi-major axis of the planet, convert this to an orbital period, and make the observation at some fraction of this period (e.g., 25%). Although this method relies on a multitude of assumptions, it can provide order of magnitude estimations for whether a planet is within the habitable zone of its host star.

3.5 Summary

In this work, we explored three methods for determining the optimum observing time to maximize orbital constraint of a directly imaged Earth-like exoplanet. We find that (1) our methods all achieve similar results, pointing to a true optimum observing time for a subsequent observation to constrain the orbit; (2) all of our methods dramatically out perform a fixed cadence approach; and (3) we slightly prefer the comprehensive method over a, e, and i, outlined in Section 3.2.3, for maximum orbital constraint in the fewest number of observations (see Figure 3.7).

We began by fitting a suite of orbits to two X-Y observation points of an Earth-like exoplanet orbiting a Sun-like star in a face-on orientation (see Figure 3.1). We then applied our three methods to determine the best time to observe the system in order to maximize orbital constraint. For our first two methods, the swath method and the cluster method, we begin with the assumption that the time at which the suite of fitted orbits has maximum separation is the time we have the least information about the true orbit. Hence, we should observe at this time of maximum orbital divergence.

With the swath method, we select interior and exterior 1-σ representative sample orbits from the suite of orbit fits, find the time within an orbital period where they have maximum separation and designate this time as the best time to observe the system (see Figure 3.3). Similarly, for the cluster method, we calculate the distance of each orbit fit from the best-fit orbit at each time to determine the time when the suite of orbits has the greatest separation and designate this time as the correct time to observe (see Figure 3.5). As both of these methods tend to produce varied results with each iteration, we repeated the process 100 times for each method to build a distribution of best fits and designate the peak of the distribution as the optimum timing given by that method (see Figure 3.4).

For our final method, the comprehensive method, rather than look for divergence of orbit fits, we assume that the best-fit orbit is close to the actual orbit and make synthetic observations along the best-fit orbit. By pairing these synthetic observations with the original data points, we determine the orbital constraint provided by observing at each time within the calculated orbital period of the planet. This neatly provides a distribution of orbital constraint as a function of observation time (see Figure 3.6) and, when optimizing constraint over the semi-major axis, eccentricity, and inclination of the planet, provides excellent limits on its orbit (see Figure 3.7).

3.6 Acknowledgments

We thank Michael Line (ASU) for his suggestions which contributed to the development of the cluster method, Chad Trujillo (NAU) for helpful feedback on the implications of this work, and Patrick Tribbett (NAU) for recommendations regarding MCMC fit probabilities. This work was funded through NASA Planetary Astronomy grant 80NSSC18K1006. TDR gratefully acknowledges support from NASA's Exoplanets Research Program (No. 80NSSC18K0349), Habitable Worlds Program (No. 80NSSC20K0226), and Exobiology Program (No. 80NSSC19K0473), the Nexus for Exoplanet System Science Virtual Planetary Laboratory (No. 80NSSC18K0829), and the Cottrell Scholar Program administered by the Research Corporation for Science Advancement.

This work was made possible in part through the State of Arizona Technology & Research Initiative Program. Some of the computational analyses were run on Northern Arizona University's Monsoon computing cluster, funded by Arizona's Technology and Research Initiative Fund. This research has made use of NASA's Astrophysics Data System. This research has made use of the NASA Exoplanet Archive, which is operated by the California Institute of Technology, under contract with the National Aeronautics and Space Administration under the Exoplanet Exploration Program.

Software: corner (Foreman-Mackey, 2016), emcee (Foreman-Mackey et al., 2013), matplotlib (Hunter, 2007), numpy (Harris et al., 2020), pandas (McKinney, 2010), REBOUND (Rein & Liu, 2012; Rein et al., 2019), scipy (Virtanen et al., 2020), termcolor[5], tqdm (da Costa-Luis et al., 2022).

[5]https://pypi.org/project/termcolor/

Chapter 4

Active Asteroid Dynamics

This chapter is derived from work that has been submitted to Astrophysical Journal Letters for which I am second author. Here, we have expanded on the dynamical analysis portions of this work and condensed much of the discussion regarding other aspects of the object of study.

ABSTRACT

We have detected a second epoch of activity on comet/active asteroid 282P/(323137) 2003 BM_{80}. This detection is one of the first science results of the recently launched *Active Asteroids* citizen science project, in which volunteers identify comet-like activity in archival telescope survey images. In connection with this finding, we have evaluated the orbital stability of 282P using a cluster of orbital clones in order to determine its orbital behavior. We find that the orbit of 282P is unstable on the order of hundreds of years, both in the future and the past, due to a series of close encounters with Jupiter and, in some cases, Saturn. Moreover, these encounters cause 282P to migrate between the dynamical regions of the Jupiter Family Comets, the Quasi-Hildas, and the outer main asteroid belt on the order of tens of times within our 1,000 year simulations. This implies that a dynamical pathway may exist between Jupiter Family Comets and active asteroids and that the Quasi-Hildas may be an intermediate population between them.

4.1 Introduction

Volatile ices, such as water and carbon dioxide, are essential for life in the solar system. By studying bodies that show comet-like activity, we can learn more about the origin and distribution of these important substances within the solar system (Hsieh & Jewitt, 2006). One poorly understood class of objects that displays comet-like activity (i.e., tails and/or comae) are the active asteroids (see Jewitt, 2012, for a review of this elusive population). Although less than 30 active asteroids have been discovered, approximately 40% of these have shown activity at multiple epochs (i.e., over multiple orbital periods) (Chandler et al., 2018), a key indicator that the activity may be caused by the sublimation of volatiles (Snodgrass et al., 2017).

To expand the known population of active asteroids we utilize the citizen science project *Active Asteroids*[1] run through the popular *Zooniverse* platform (PI: Colin Chandler). In the *Active Asteroids* project, Citizen Scientists[2] inspect telescope images for objects showing signs of cometary activity (such as tails or comae). Since its launch in August 2021, over 6,600 volunteers have made more that 2.8 million classifications of images of active and inactive objects. The process of finding and preparing images of potentially active objects is described in detail by Chandler et al. (2018, 2019, 2020, 2021).

Once images have been examined, typically by ∼15 Citizen Scientists, we receive a score indicating how many of these volunteers have classified the image as containing activity. Images with a high fraction of reported positive scores are then examined by our team and, if there is indeed apparent activity, other images of the same object are pulled from the database for comparison. We also determine if follow-up observations are feasible based on available telescope time and current brightness and location of the object of interest. When possible, we make additional observations of the object and examine all available images of it for signs of activity. By pairing these visual comparisons with dynamical analysis of the object's orbit, we gain insights into the probable mechanisms that cause the object to be active. For example, if an object is only active once, but there are several images of it across multiple orbital periods that would likely show activity if it were present, then the activity may be a single impact or outburst event; however, if there are repeated signs of activity that correspond to when the object is near perihelion, the activity mechanism is most probably sublimation.

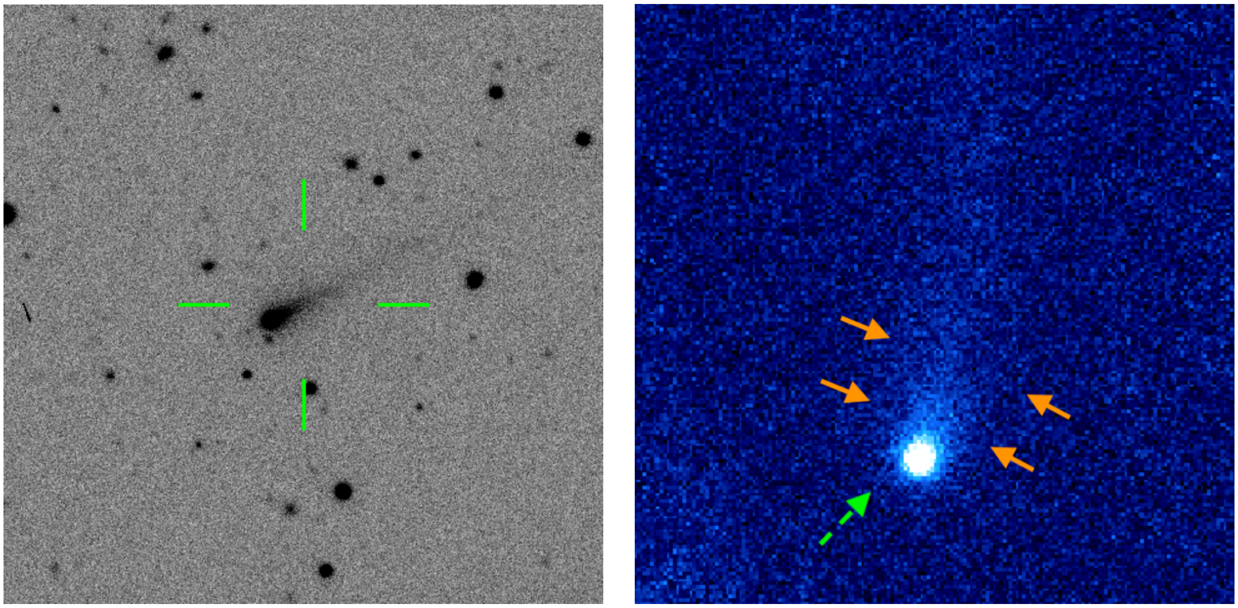
One of the most promising *Active Asteroids* discoveries thus far is an additional epoch of activity on comet/active asteroid 282P/(323137) 2003 BM_{80}, which we will refer to as 282P (see Figure 4.1)[3]. Activity for 282P was detected in 2012-2013 and 2021-2022 and all detections were within a year of its time of perihelion, where sublimation-driven activity is most expected for a volatile rich active asteroid. Hence, this discovery strongly indicates that 282P is recurrently active near each perihelion passage and, therefore, the activity is most likely driven by sublimation.

It is widely accepted that Centaurs originated in the trans-Neptunian region and are scattered to the giant planet region through encounters with Neptune and, subsequently, the other giant planets as well (see, for example, review by Morbidelli & Nesvorný, 2020, and references therein). The Jupiter Family Comets

[1] http://activeasteroids.net

[2] Anyone who wants to can go to the link and participate.

[3] In preparation for presenting my dissertation, I needed a nickname for 282P since our work had not yet been published. My son, Garrett, who was 5 years old at the time, came up with the name "Lightning Ball" based on the cool blue look of Figure 4.1b and we have been calling it that in our research group ever since.



(a) (b)

Figure 4.1 (a) Image of 282P taken with the Dark Energy Camera on UT 2021 March 14, Prop. ID 2019A-0305 (P.I. Drlica-Wagner). Here a tail is clearly visible extending from the object toward the upper right of the image. (b) Composite image of 282P from six co-added images taken with the Gemini Multi-Object Spectrograph imager on UT 2022 June 7, Prop. ID GS-2022A-DD-103 (P.I. Chandler). Activity (indicated by orange arrows) can be seen extending from 282P towards the top of the image. Both of these images highlight the second detected epoch of activity for 282P.

(JFCs) are also sourced from this population and recent dynamics work has shown a specific region in orbital element space, called the "JFC gateway," where Centaurs are actively transitioning from Centaurs to JFCs (Sarid et al., 2019). Both of these populations contain objects that show signs of cometary activity and are the result of dynamical migration inwards from the outer solar system. Further dynamical analysis of objects migrating between the JFC and asteroidal regions, such as 282P, may provide additional insight into a continuation of this pathway for volatile transport from the outer solar system to the inner solar system. In the rest of this chapter, we present the methods used to carry out our dynamical modeling of 282P and discuss our results and their implications.

4.2 Dynamical Methods

After confirming the multi-epoch nature of the activity associated with 282P, we analyzed its dynamical stability — how long an object will stay on a given orbit — to determine its orbital class and to search for evidence of the underlying mechanism of the observed activity. For these analyses, we simulated a cloud of 500 orbital clones of 282P. Orbital clones are copies of the original object with a small offset in its orbital parameters determined randomly within the uncertainty of the orbit. In this analysis, clones were drawn

Table 4.1. 282P Orbital Parameters

Parameter	Value	Uncertainty
Orbital Period	$P = 8.732$ yr	2.174×10^{-7} yr
Semi-major Axis	$a = 4.240$ au	7.039×10^{-8} au
Perihelion	$q = 3.441$ au	3.468×10^{-7} au
Aphelion	$Q = 5.039$ au	8.366×10^{-8} au
Eccentricity	$e = 0.188$	7.790×10^{-8}
Inclination	$i = 5.812°$	$1.166° \times 10^{-5}$
Argument of Perihelion	$\omega = 217.626°$	$7.816° \times 10^{-5}$
Longitude of Ascending Node	$\Omega = 9.297°$	$5.974° \times 10^{-5}$
Mean Anomaly	$M = 9.979°$	$3.815° \times 10^{-5}$

Note. — Orbital solution date: 8 October 2021, JPL Horizons (Giorgini et al., 1996).

from Gaussian distributions centered on the current fitted parameters of 282P with widths corresponding to the uncertainties on those fits (see Table 4.1).

We modeled the gravitational interactions of the orbital clones with the Sun and the planets, excepting Mercury, — which has negligible effect on the orbit of 282P — using the highly-accurate (typically to machine precision) IAS15 N-body integrator from the `REBOUND` `python` package (Rein & Spiegel, 2015). Integrations were run forward and backward in time for 1,000 years. Longer integrations were unnecessary as dynamical chaos (where small differences in initial position result in large differences in final orbit) ensues for 282P on the order of hundreds of years in both the past and the future.

As part of our orbital analysis of 282P we utilized two dynamical quantities that help in determining the effect that Jupiter and other giant planets (e.g., Saturn) have on the orbit of an object. The first of these is the Tisserand parameter with respect to Jupiter, defined as

$$T_{\mathrm{J}} = \frac{a_{\mathrm{J}}}{a} + 2\cos(i)\sqrt{\frac{a}{a_{\mathrm{J}}}\left(1 - e^2\right)}, \tag{4.1}$$

where a_{J} is the semi-major axis of Jupiter and a, e and i are the semi-major axis, eccentricity and inclination of the body, respectively[4]. One of the main uses of T_{J} is to quantify how much of an effect Jupiter has on the orbit of a given object. For values of $T_{\mathrm{J}} > 3$, the object will not cross Jupiter's orbit, but for values less than 3, their orbits will cross (but not necessarily intersect), leading to a higher probability of close encounters. Because of this, T_{J} is often used to differentiate between different classes of objects near Jupiter,

[4] T_{J} is a commonly used form of the Tisserand relation, obtained by multiplying the relation by 2 and replacing the semi-major axis in the relation with the ratio of the semi-major axes of the body to Jupiter a/a_{J}; see Section 3.4 of Murray & Dermott (1999) for a discussion of the Tisserand relation.

for example, JFCs are defined as having $2 < T_{\mathrm{J}} < 3$ (Levison, 1996), and asteroids have $T_{\mathrm{J}} > 3$ (Jupiter has a T_{J} approximately equal to 3).

The second dynamical quantity we will use throughout our analyses is the Hill radius of Jupiter and Saturn. The Hill radius r_H is a metric of orbital stability and indicates the region where a secondary body (e.g. a planet) has dominant gravitational influence on a tertiary body (e.g. a moon) compared to the primary body (e.g. the sun). At perihelion, the Hill radius of the secondary body can be approximated as

$$r_H \approx a(1-e)(m/3M)^{1/3}, \tag{4.2}$$

where a, e, and m are the semi-major axis, eccentricity and mass of the secondary (Jupiter or Saturn in our case) respectively and M is the mass of the primary (i.e. the Sun; see Section 2.1.2 of Hamilton & Burns (1992) and Hill (1878)). Close passages of a small body within a few Hill radii of a planet are generally considered to be significant perturbations and may drastically alter the orbit of the small body.

Using these two metrics, combined with the orbital elements of 282P, we analyzed its orbital evolution and probable orbital history (forward and backward integration). Non-gravitational effects on the orbit, such as those resulting from sublimation, were not accounted for in our analyses. These effects are typically negligible in comparison with gravitational forces and, although they may have a minor effect on close encounter distances, should fall well within the observational uncertainties for 282P. We present our results through a series of figures and their accompanying discussion in the following section.

4.3 Results

One of the first things to notice in our dynamical analysis of 282P is the proximity of its orbit (blue) to Jupiter's orbit (orange) as seen in Figure 4.2. Because of this proximity, we can expect that close encounters between Jupiter and 282P are probable[5].

We can clearly see close encounters happening between 282P and Jupiter by plotting the distance between them as as function of time in Figure 4.3. Here, the distance between the bodies is shown in log space and horizontal lines show the values of 1, 3, and 5 Jupiter Hill radii. For our purposes, we will define that anytime 282P is within approximately 5 Jupiter Hill radii ($\sim$2 au) that it qualifies as a close encounter. We also plot the semi-major axes of two of Jupiter's moons for reference. Callisto is the outermost of the large Galilean moons of Jupiter and Sinope is a distant irregular Jovian moon that was likely captured by Jupiter via a close encounter rather than forming in-situ (Grav et al., 2003). Time $t = 0$ corresponds to JD2459600.5 (Jan

[5]Unless 282P and Jupiter were in a protective Mean Motion Resonance, which they are not (see Chapter 1.2.4).

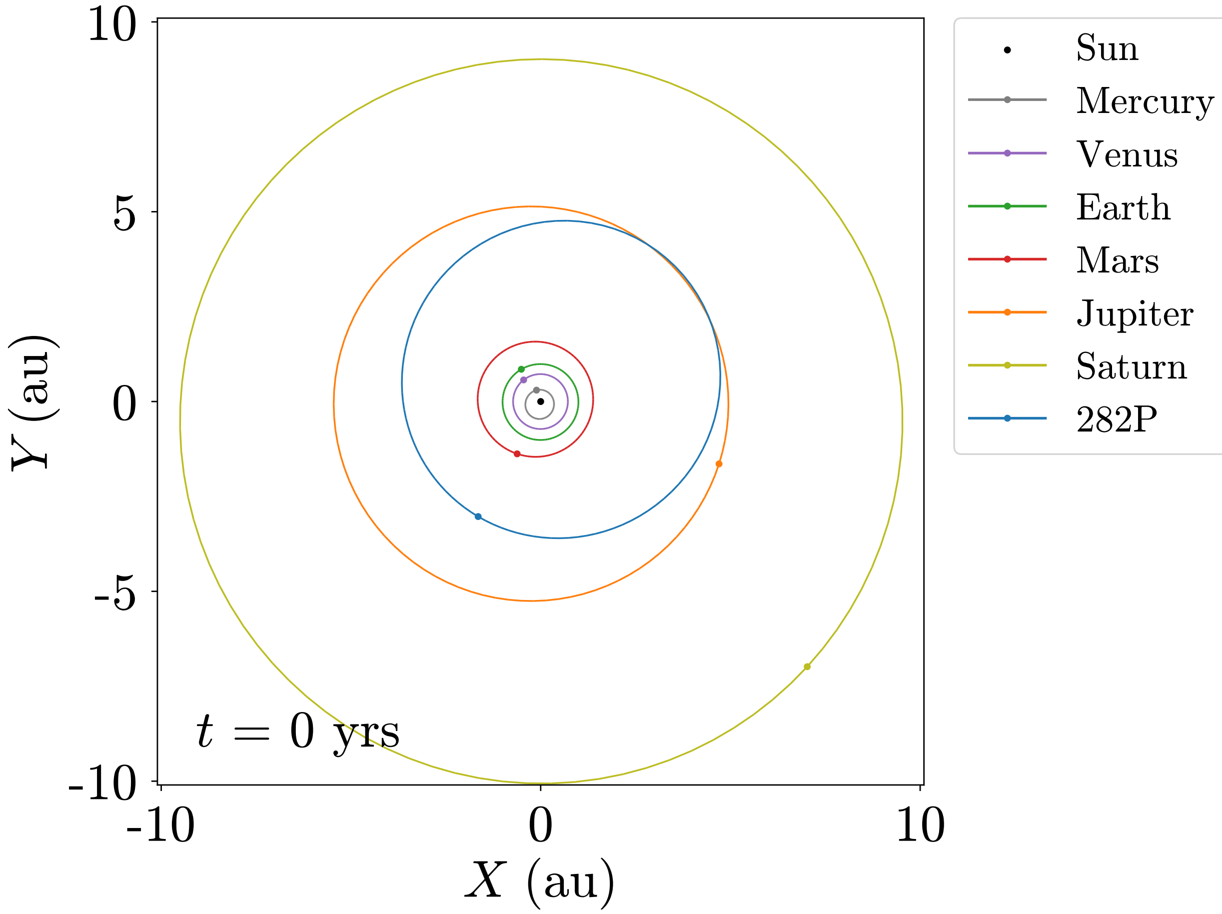


Figure 4.2 Orbital plan view of 282P (blue) and the planets interior to Saturn. $t = 0$ corresponds to Jan 21, 2022. Note the extended region of overlap between the orbits of Jupiter and 282P.

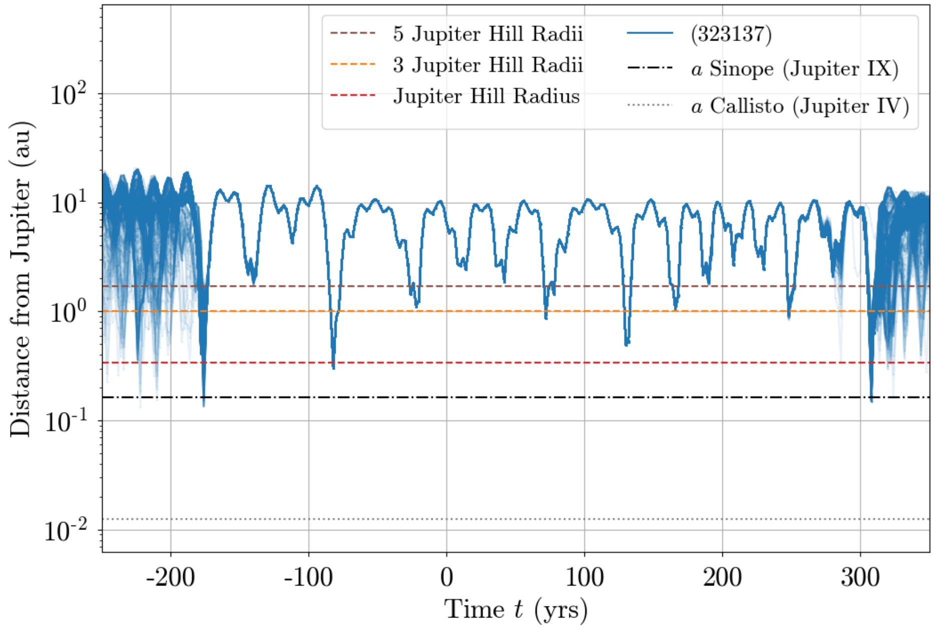


Figure 4.3 Log distance of 282P from Jupiter as a function of time (blue). Dashed horizontal lines representing 1 (brown), 3 (orange), and 5 (red) Hill radii are shown. Close encounters within 5 Hill radii of Jupiter are considered significant and can greatly alter 282P's orbit. Dotted gray and dot-dashed black horizontal lines show the semi-major axes of two of Jupiter's moons, Callisto and Sinope, respectively. Callisto is the outermost Galilean satellite and Sinope is a distant irregular moon likely captured by Jupiter (Grav et al., 2003). Deep encounters (less than one Hill radius) between 282P and Jupiter result in dynamical chaos for 282P's orbit beyond the bounds of the plot.

21, 2022) and the span from $t = -250$ to $t = 350$ (1772-2372 AD) is shown in this and subsequent figures. Here we see on the order of 10 instances where 282P and Jupiter have close encounters, including at least three that occur within one hill radius of Jupiter.

The effect of the close encounters between 282P and Jupiter can be readily seen by examining the evolution of the orbital elements of 282P. In Figure 4.4, the semi-major axis of 282P as a function of time is shown along with the semi-major axes of Jupiter and Saturn. Here, close encounters appear as jumps in the semi-major axis of 282P. Similar jumps in eccentricity and inclination caused by the same close encounters can be seen in Figures 4.5 and 4.6, respectively.

One of the most noticeable features in these figures is the deep encounter, within one Hill radius (∼0.3 au), with Jupiter around the year 1940 ($t = -82$ yrs). This encounter dramatically alters 282P's orbit, shifting it from an orbit between Saturn and Jupiter to an orbit interior to Jupiter, similar to its current orbit. This can be seen in Figure 4.4 and also in Figure 4.7, where the heliocentric distance of 282P, Saturn, Jupiter, and Mars are all shown as a function of time. This figure highlights that, not only is 282P's orbit between Jupiter and Saturn, but it also crosses the orbits of both of these giant planets. The 1940 encounter

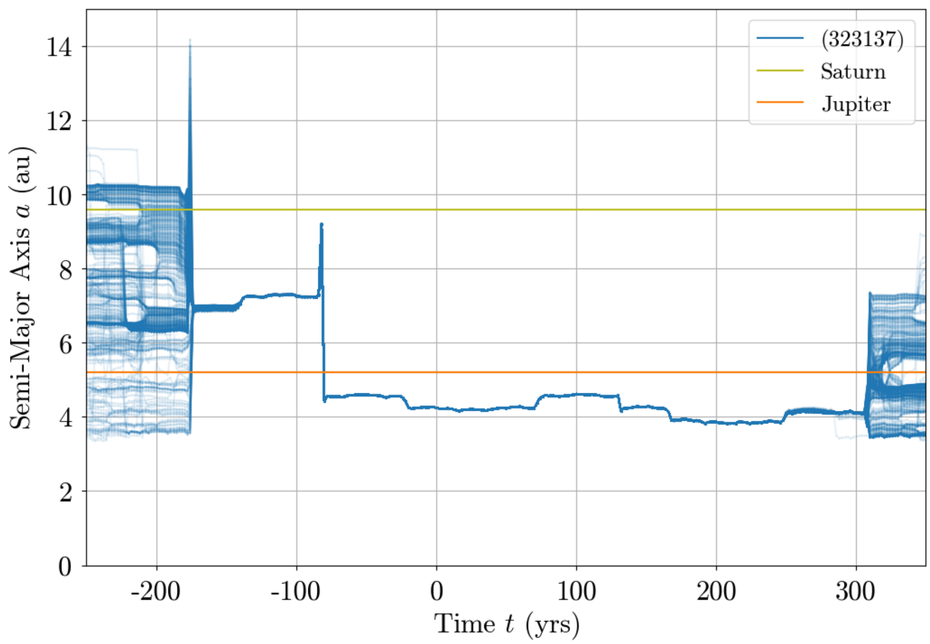


Figure 4.4 Semi-major axis evolution of 282P (blue). Horizontal olive and orange lines are the semi-major axes of Saturn and Jupiter respectively. Note the encounter at $t = -82$ years which causes 282P to transition from a semi-major axis between Jupiter's and Saturn's to an a interior to that of Jupiter. Dynamical chaos, as noted in Figure 4.3 is also evident here and in subsequent plots in this section.

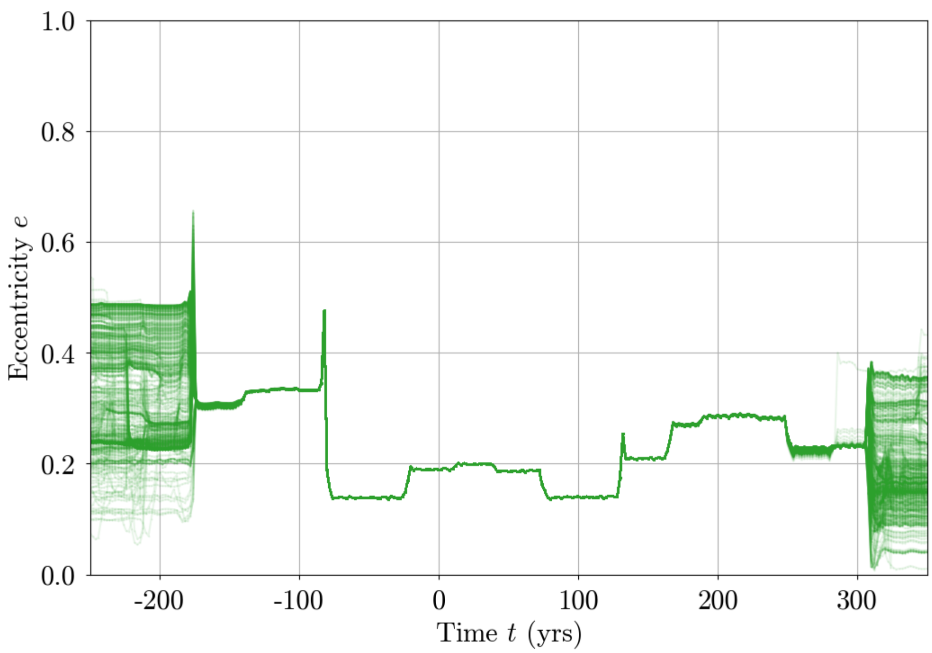


Figure 4.5 Eccentricity evolution of 282P. Changes in e are caused by close encounters with Jupiter and/or Saturn. Dynamical chaos is evident before ~ -180 years and after ~ 300 years.

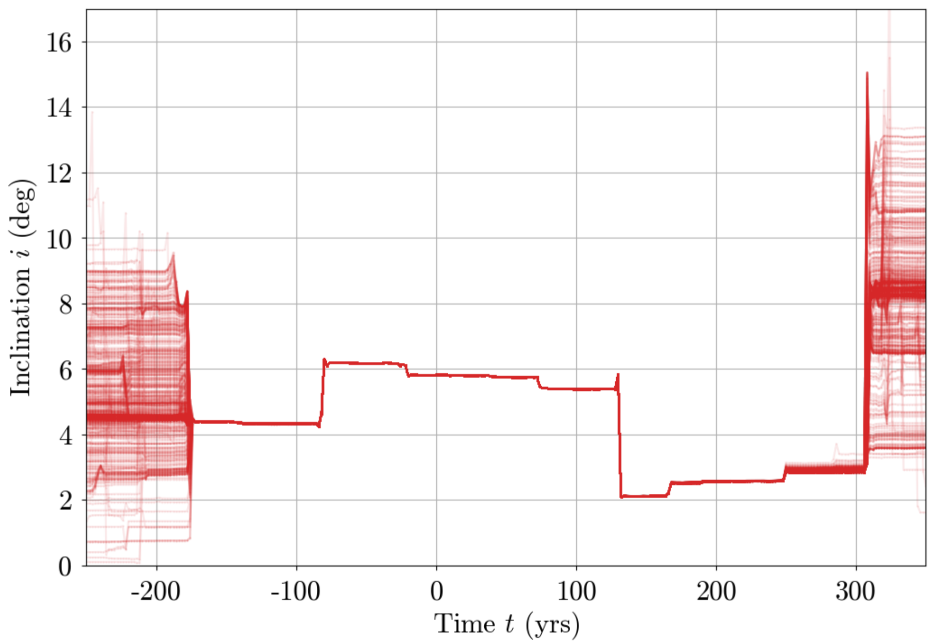


Figure 4.6 Inclination evolution of 282P. Changes in i are caused by close encounters with Jupiter and/or Saturn. Dynamical chaos is evident before ~ -180 years and after ~ 300 years.

with Jupiter also brings the T_J of 282P to within ∼0.01 of the $T_\mathrm{J} = 3$ boundary between cometary and asteroidal orbits, as seen in Figure 4.8.

Prior to its 1940 encounter with Jupiter, 282P was on an orbit that crossed the orbits of both Jupiter and Saturn. It was placed on this orbit by a strong close encounter ($< 1r_H$) with Saturn in approximately 1838 ($t \approx -184$ yrs), as seen in Figure 4.9, followed by another strong encounter with Jupiter (see Figure 4.3) in 1846 ($t = -176$ yrs). These highly perturbative passages placed 282P on the path towards its present day transitionary orbit; however, due to the closeness of these encounters, compounded by the orbital uncertainties of 282P, it is not possible to predict its past orbit beyond these encounters. This dynamical chaos is clearly visible in Figures 4.4 – 4.9 as orbital clones take a multitude of paths within the parameter space and result in a broad range of possible orbits caused only by slight variations in the initial orbital parameters of 282P.

Dynamical chaos is also evident after the major encounter with Jupiter set to occur around the year 2330 ($t \approx 308$ yrs; see Figure 4.3) and, after that event, the orbit of 282P is not resolved around a single solution. A slight diffusion following the previous several Jupiter passages is also visible in Figure 4.3 and adds uncertainty surrounding encounters between roughly 2301 and 2306 ($t \approx 280$ to 285 yrs). Although dynamical chaos dominates past and future orbits of 282P on the order of hundreds of years (beyond approximately $-180 \lesssim t \lesssim 300$ yrs given present orbital uncertainties), we can examine the fraction of orbital clones that

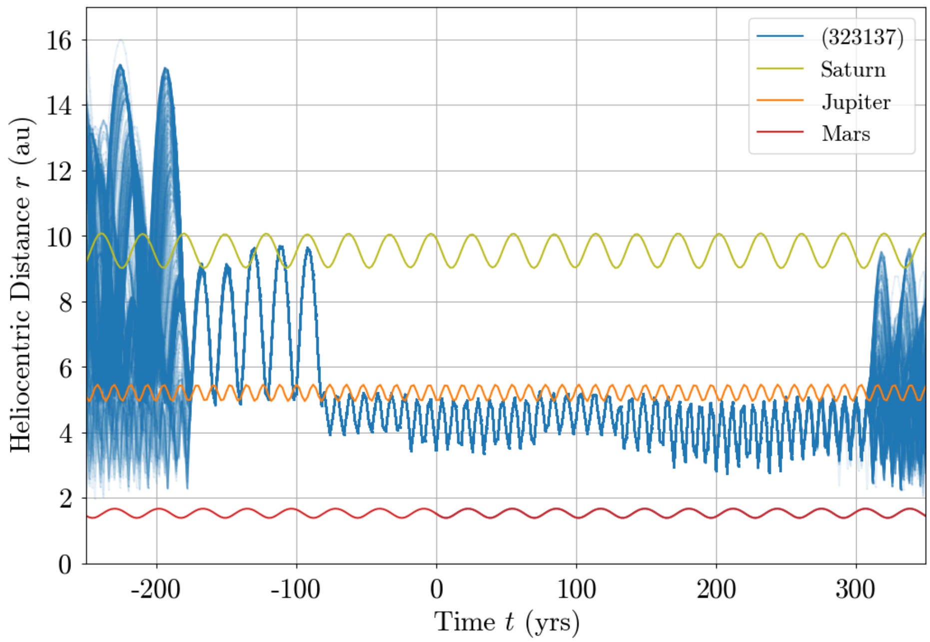


Figure 4.7 Heliocentric distance of 282P (blue), Saturn (olive), Jupiter (orange), and Mars (red). Note that between $\sim t = -176$ years and $t = -82$ years, the orbit of 282P crosses the orbits of both Saturn and Jupiter.

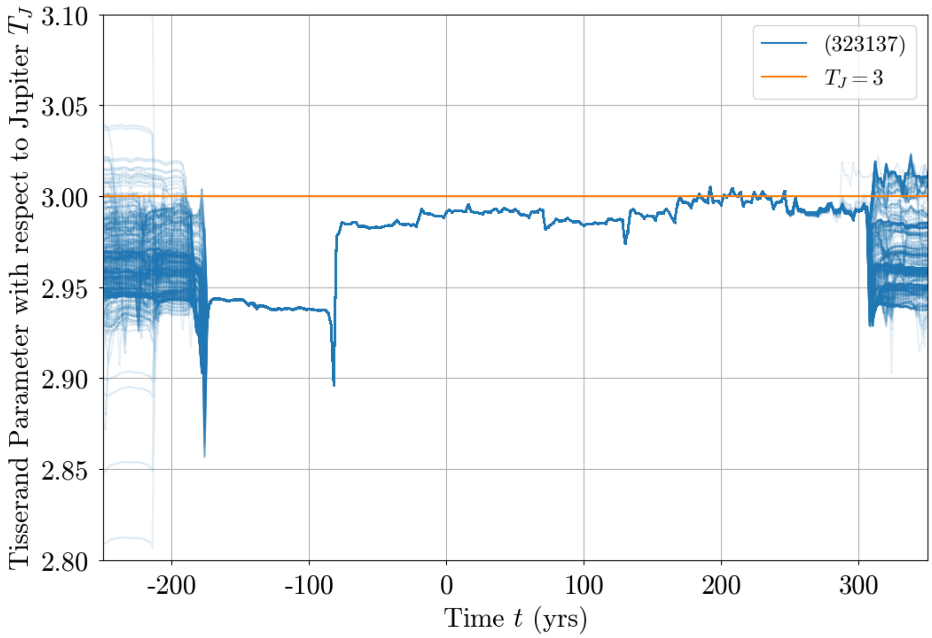


Figure 4.8 The Tisserand parameter with respect to Jupiter of 282P as a function of time (blue). The horizontal orange line highlights the traditional boundary at $T_{\mathrm{J}} = 3$ between cometary and asteroidal orbits. Note how 282P crosses this boundary on the order of ten times in the time surrounding $t = 200$ years.

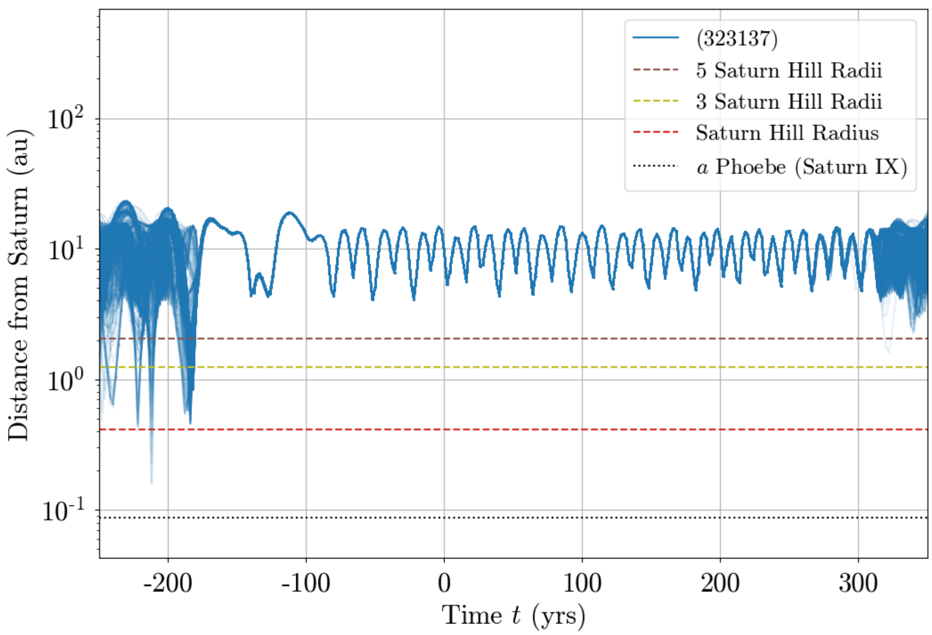


Figure 4.9 Log distance of 282P from Saturn as a function of time (blue). Dashed horizontal lines representing 1 (brown), 3 (olive), and 5 (red) Hill radii are shown. Similar to Figure 4.3, only the gray dotted line is the semi-major axis of Saturn's irregular moon, Phoebe, which is though to be captured by a close encounter with Saturn (Johnson & Lunine, 2005; Jewitt & Haghighipour, 2007).

finish the simulation (forwards and backwards) within the bounds of various orbital classes to obtain a statistical representation of 282P's probable orbital class.

By examining the orbital elements of the dynamical clones we determined the probability of 282P belonging to various dynamical classes and present the results in Table 4.2. We find that 282P falls within four different classes during the time span $-250 < t < 350$ years; (1) Jupiter Family Comets with $2 < T_J < 3$ and perihelia q interior to the semi-major axis of Jupiter a_J; (2) Quasi-Hildas (QHs), a small (∼300 known) population of objects occupying a similar orbital region as the Hilda asteroid group[6] (Toth, 2006; Gil-Hutton & García-Migani, 2016) with inexact boundaries of $3.5 \lesssim a \lesssim 4.4$ and $2.9 \lesssim T_J \lesssim 3.08$; (3) Centaurs with $q > a_J$ and, therefore, $T_J > 3$; and (4) asteroids with $3.2 < a < 4.6$ and $T_J > 3$. There is overlap between the QHs and the JFCs as well as the QHs and the asteroids. Hence, an object can be both a QH and a JFC or both a QH and an asteroid, but not a JFC and asteroid simultaneously.

We calculate the Tisserand parameter with respect to Jupiter of 282P at $t = 0$ to be $T_J = 2.991$. As this is less than the traditional $T_J = 3$ cutoff for asteroids, so, 282P is not presently an asteroid. It is, however, both a Jupiter Family Comet and a Quasi-Hilda. Prior to arriving on its current orbit, 282P likely migrated from either the Centaur or JFC regions (roughly equal probability). In the next 350 years, we predict that

[6]The Hildas are in the 3:2 interior mean motion resonance with Jupiter and the Quasi-Hildas are near but not in this resonance.

Table 4.2. Orbital Clone Outcomes

Orbit Class	$t = -250$ yrs	$t = 0$ yrs	$t = 350$ yrs
JFC	260 (52%)	500 (100%)	403 (81%)
QH	26 (5%)	500 (100%)	90 (18%)
Centaur	239 (48%)	0 (0%)	28 (5.6%)
Asteroid	1 (0.2%)	0 (0%)	69 (14%)

Note. — The number of orbital clones in a given class at three separate times are shown. Because some of these classes are not mutually exclusive, the total number in each column may exceed the 500 total orbital clones used; particularly, the Quasi-Hildas can also simultaneously be either JFCs or asteroids.

282P will likely return to a primarily JFC orbit; however, migration to asteroid or centaur regions are not unlikely. Both in the past and in the future, the probability of 282P remaining a member of the QH region is greatly diminished. In previous work, this has been shown to be common among the QHs, which have short dynamical lifetimes and a propensity to migrate to and/or from the JFC and Centaur regions (Jewitt & Kim, 2020).

In addition to the probability that 282P resides on an asteroidal orbit in 350 years, over that time span all 500 orbital clones cross $T_{\mathrm{J}} = 3$, the boundary between JFC and asteroid-like orbits, on the order of ten times (see Figure 4.8). This migration implies a potential pathway for JFCs to become active asteroids, with the QHs being an intermediate population between these two classes of objects. The migratory nature of this intermediate population is highlighted by the repeated oscillation of 282P across this $T_{\mathrm{J}} = 3$ boundary. This causes the orbital class of 282P, as traditionally defined, to rapidly change between JFC and asteroid with almost no change to the actual orbit of 282P. Hence, although the boundary has, in the past, been drawn as a sharp line (e.g., Levison, 1996), in practice, the surrounding region is highly transitory and prone to dramatic orbit alterations from close encounters with Jupiter.

In this work, we adopt a broader range of parameters for the QH region than is used in other works. We find this to more accurately represent the Quasi-Hilda population. In support of this decision, we provide Figure 4.10, which shows 200 year integrations of minor planets from various orbital classes in the reference frame corotating with Jupiter (where the angular position of Jupiter is fixed and relative motion is shown for the minor planet). Plots of this kind are often used to highlight mean motion resonances, like the one between (153) Hilda and Jupiter (panel d) which results in the characteristic trilobal shape. Here we see marked similarities between 246P, previously designated as a QH (Toth, 2006, panel e) and 282P (panel f) supporting the classification of 282P as a Quasi-Hilda.

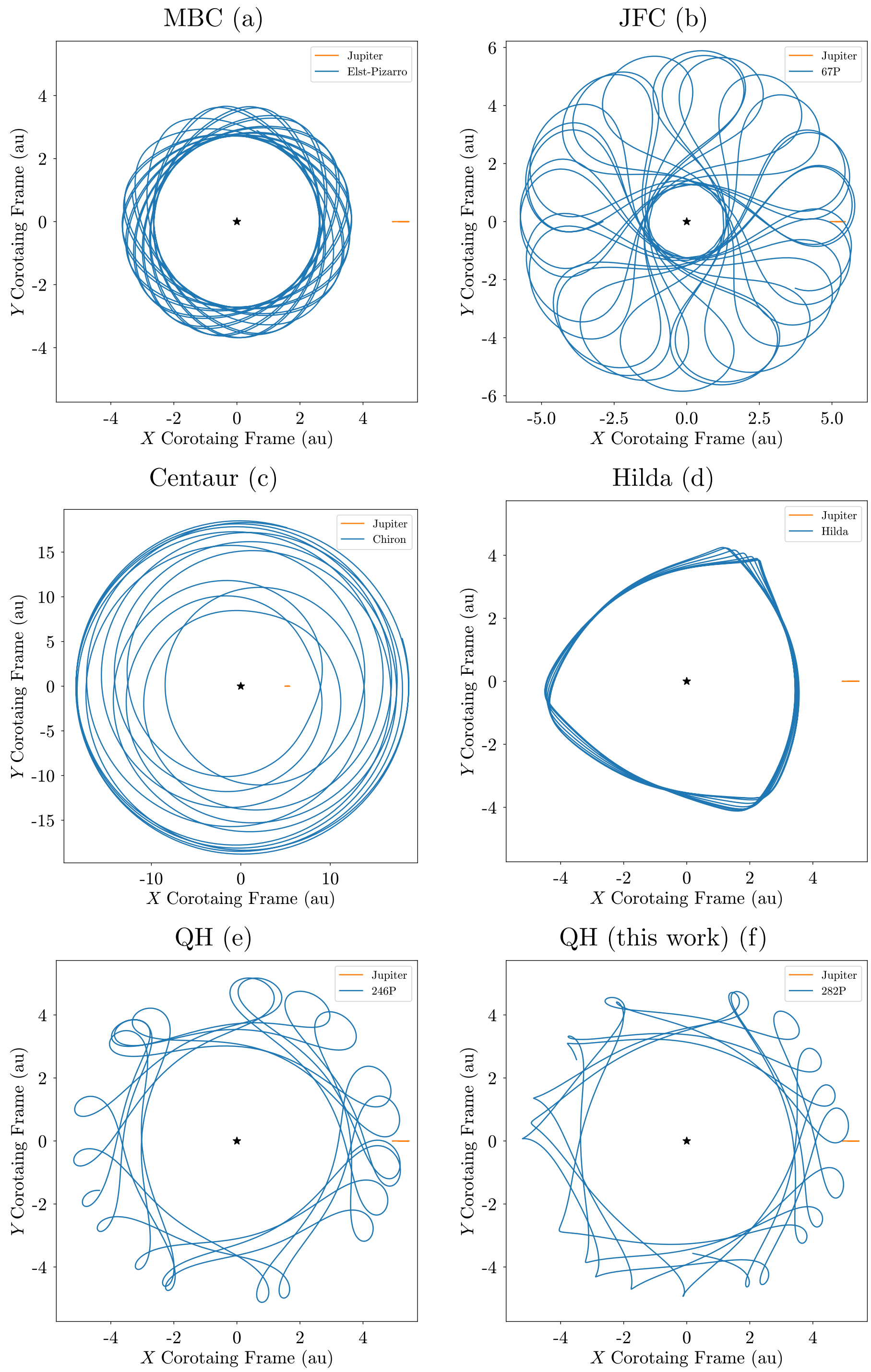


Figure 4.10 The orbital motion of minor planets (blue lines) as seen in the reference frame corotating with Jupiter (orange lines at right edge of plots). (a) Main Belt Comet (7968) Elst-Pizarro (133P). Main Belt Comets are typically defined as having $T_J > 3.08$ and orbits within the main asteroid belt (Jewitt et al., 2015; Hsieh et al., 2015). (b) JFC 67P/Churyumov-Gerasimenko (previously visited by the European Space Agency Rosetta Spacecraft). (c) Centaur (2060) Chiron (95P). (d) (153) Hilda, the namesake of the Hilda dynamical class, in the 3:2 interior mean-motion resonance with Jupiter. (e) Quasi-Hilda 246P/NEAT, also designated 2010 V_2 and 2004 F_3 (Toth, 2006). (f) Our object of study, 282P, in its Quasi-Hilda orbit.

4.4 Discussion and Summary

The fact that 282P is actively undergoing dramatic changes in its orbit due to close encounters with the giant planets further supports the idea that the activity observed on 282P is likely sublimation-driven since volatile ices will have increased potential to sublimate as 282P migrates closer to the Sun. Combined with the observation of two separate epochs of activity, this makes other explanations of activity, such as an impact event or activation caused by gravitational disruption after a close encounter with a giant planet, very unlikely. The fact that all activity on 282P has been discovered while it was near perihelion, where it is the warmest, further augments this argument (Hsieh et al., 2012).

Through our dynamical investigations of 282P, we clearly see that it migrates many times across the $T_J = 3$ boundary over the course of our simulations. This behavior suggests that 282P is actively transitioning between an orbit of a Jupiter Family Comet and an asteroidal orbit. This finding has key implications for objects in this region, such as the Quasi-Hildas. First, it indicates that there is a pathway for Jupiter Family Comets, which are primarily active bodies, to become active asteroids. Additionally, the reverse pathway will also be present, allowing active asteroids to become Jupiter Family Comets. It may be that Jupiter Family Comets that are expected to be active but are quiescent spent enough time as active asteroids to shed most of their volatiles prior to migrating back out to the JFC region. Similar dynamical pathways have been explored for other populations in the solar system and it has been shown that there is a dynamical "gateway" between the Centaurs and the Jupiter Family Comets that allows for efficient transfer from the more distant Centaur population (which are, in turn, thought to originate in the Trans-Neptunian region (Duncan et al., 2004; Volk & Malhotra, 2008)) to the active Jupiter Family Comet region (Sarid et al., 2019). Although much is still unknown about these pathways, these transitioning objects may be important in unveiling the underlying dynamical mechanisms behind the migration patterns of active small bodies in the solar system.

Secondly, the transitions of 282P around the $T_J = 3$ boundary indicate that the difference between Jupiter Family Comets and active asteroids is unlikely to be accurately defined by a sharp line at $T_J = 3$, as has historically been done, and is rather encompassed by a transition region, which potentially coincides with the Quasi-Hilda population. Within this region, objects have chaotic orbits that will eventually lead to a more stable orbit within either the JFC or active asteroid populations, or in having their orbits completely disrupted through close encounters with the giant planets.

In summary, we have detected a second epoch of activity on the object 282P/(323137) 2003 BM_{80} through the *Active Asteroids* citizen science project. By gravitationally modeling orbital clones of 282P, we find that it has undergone and will undergo several close encounters with Jupiter and likely Saturn as well in the past 250 years and in the next 350 years. These close encounters dramatically alter the orbit of 282P on the order of hundreds of years, both in the past and in the future. During this time, 282P migrates across the orbital boundary between the Jupiter Family Comets and the active asteroids many times and through the Quasi-Hilda region. This migration hints at a potential pathway for Jupiter Family Comets to become active asteroids and vice versa. Additionally, our findings support that the transition boundary between these two populations may be less-defined than the traditional line drawn at a Tisserand parameter with respect to Jupiter of three, and instead encompass a broader region of this parameter space. Further studies of objects near the transition boundary may yield clues to understanding the active asteroids as a population and, consequently, the role they play in the distribution of volatiles throughout the solar system.

4.5 Acknowledgments

We thank Colin Chandler (NAU) and Chad Trujillo (NAU) for their work on this project as well as Henry Hsieh (PSI) and David Jewitt (UCLA) for their helpful feedback. We also express gratitude to Mark Magbanua (UCSF), Mike Gowanlock (NAU), Jay Kueny (UA/Lowell), and the Trilling Research Group (NAU) for their insights. The unparalleled support provided by Monsoon cluster administrator Christopher Coffey (NAU) and the High Performance Computing Support team facilitated the scientific process.

We thank Gemini Observatory Director Jennifer Lotz for granting our DDT request for observations, German Gimeno for providing science support, and Pablo Prado for observing. Proposal ID GS-2022A-DD-103, PI Chandler.

We are grateful to the Vatican Observatory for the generous time allocations at the Vatican Advanced Technology Telescope (Proposal ID S165, PI Chandler). We especially thank Vatican Observatory Director Br. Guy Consolmagno, Rev. Pavel Gabor, S.J., Rev. Richard P. Boyle, S.J., Chris Johnson, Michael Franz, Summer Franks, and Gary Gray.

This work was funded through NASA Planetary Astronomy grant 80NSSC18K1006, NASA Solar System Observation grant 80NSSC19K0869, and National Science Foundation Graduate Research Fellowship grant No. 2018258765.

Some of the computational analyses were run on Northern Arizona University's Monsoon computing cluster, funded by Arizona's Technology and Research Initiative Fund. This work was made possible in part through the State of Arizona Technology and Research Initiative Program.

World Coordinate System corrections facilitated by the *Astrometry.net* software suite (Lang et al., 2010). This research has made use of data and/or services provided by the International Astronomical Union's Minor Planet Center. This research has made use of NASA's Astrophysics Data System. This research has made use of The Institut de Mécanique Céleste et de Calcul des Éphémérides (IMCCE) SkyBoT Virtual

Observatory tool (Berthier et al., 2006). This work made use of the `FTOOLS` software package hosted by the NASA Goddard Flight Center High Energy Astrophysics Science Archive Research Center. This research has made use of SAOImageDS9, developed by Smithsonian Astrophysical Observatory (Joye, 2006). World Coordinate System validation was facilitated with Vizier catalog queries (Ochsenbein et al., 2000) of Gaia Data Release 2 (Gaia Collaboration, 2018) and the Sloan Digital Sky Survey Data Release 9 (Ahn et al., 2012) catalogs. This work made use of the Lowell Observatory Asteroid Orbit Database *astorbDB* (Bowell et al., 1994; Moskovitz et al., 2021).

Based on observations at Cerro Tololo Inter-American Observatory, the National Science Foundation's National Optical and Infrared Laboratory (NOIRLab Prop. ID 2019A-0305; PI: A. Drlica-Wagner, NOIRLab Prop. ID 2013A-0327; PI: A. Rest), which is managed by the Association of Universities for Research in Astronomy under a cooperative agreement with the National Science Foundation. This project used data obtained with the Dark Energy Camera (DECam), which was constructed by the Dark Energy Survey (DES) collaboration. Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the Ministério da Ciência, Tecnologia e Inovação, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey. The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Enérgeticas, Medioambientales y Tecnológicas–Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenössische Technische Hochschule (ETH) Zürich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciències de l'Espai (IEEC/CSIC), the Institut de Física d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität München and the associated Excellence Cluster Universe, the University of Michigan, the National Science Foundation's Nationoal Optical and Infrared Laboratory, the University of Nottingham, the Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, and Texas A&M University.

These results made use of the Lowell Discovery Telescope at Lowell Observatory. Lowell is a private, non-profit institution dedicated to astrophysical research and public appreciation of astronomy and operates the Lowell Discovery Telescope in partnership with Boston University, the University of Maryland, the University of Toledo, Northern Arizona University, and Yale University. The Large Monolithic Imager was built by Lowell Observatory using funds provided by the National Science Foundation (AST-1005313).

Very Large Telescope OMEGACam (Arnaboldi et al., 1998; Kuijken et al., 2002; Kuijken, 2011) data were originally acquired as part of the Kilo-Degree Survey (de Jong et al., 2015).

The Panoramic Survey Telescope and Rapid Response System (Pan-STARRS) 1 Surveys (PS1) and the PS1 public science archive have been made possible through contributions by the Institute for Astronomy, the University of Hawaii, the Pan-STARRS Project Office, the Max-Planck Society and its participating institutes, the Max Planck Institute for Astronomy, Heidelberg and the Max Planck Institute for Extraterrestrial Physics, Garching, The Johns Hopkins University, Durham University, the University of Edinburgh, the Queen's University Belfast, the Harvard-Smithsonian Center for Astrophysics, the Las Cumbres Observatory Global Telescope Network Incorporated, the National Central University of Taiwan, the Space Telescope Science Institute, NASA under Grant No. NNX08AR22G issued through the Planetary Science Division of the NASA Science Mission Directorate, National Science Foundation Grant No. AST-1238877, the University of Maryland, Eotvos Lorand University, the Los Alamos National Laboratory, and the Gordon and Betty Moore Foundation.

Based on observations obtained with MegaPrime/MegaCam, a joint project of the Canada-France-Hawaii Telescope and Commissariat a l'Energes Atomique/Département d'Astrophysique, de physique des Particules, de physique Nucléaire et de l'Instrumentation Associée, at the Canada-France-Hawaii Telescope which is operated by the National Research Council of Canada, the Institut National des Science de l'Univers of the Centre National de la Recherche Scientifique of France, and the University of Hawaii. The observations at the Canada-France-Hawaii Telescope were performed with care and respect from the summit of Maunakea which is a significant cultural and historic site.

Magellan observations made use of the Inamori-Magellan Areal Camera and Spectrograph instrument (Dressler et al., 2011).

This research has made use of the NASA/Infrared Processing and Analysis Center Infrared Science Archive, which is funded by NASA and operated by the California Institute of Technology.

Facilities: Astro Data Archive, Blanco (DECam), CFHT (MegaCam), Gaia, Gemini-South (GMOS-S), IRSA, LDT (LMI), Magellan: Baade (TSIP), PO:1.2m (PTF, ZTF), PS1, Sloan, VATT (VATT4K), VST (OmegaCAM)

Software: astropy (Price-Whelan et al., 2018), astrometry.net (Lang et al., 2010), FTOOLS[7], JPL Horizons (Giorgini et al., 1996), matplotlib (Hunter, 2007), numpy (Harris et al., 2020), pandas (McKinney, 2010), REBOUND (Rein & Liu, 2012; Rein & Spiegel, 2015), SAOImageDS9 (Joye, 2006), scipy (Virtanen et al., 2020), Siril[8], SkyBot (Berthier et al., 2006), termcolor[9], tqdm (da Costa-Luis et al., 2022), Vizier (Ochsenbein et al., 2000).

[7]https://heasarc.gsfc.nasa.gov/ftools/
[8]https://siril.org
[9]https://pypi.org/project/termcolor/

Chapter 5

Discussion and Summary

Each of the three projects presented here focus on more precisely determining the location and orbit of different kinds of planets. By doing so, we learn about the interactions between bodies in the system and develop a greater understanding of the architecture of our solar system, and also other star systems. Gravitational modeling techniques make these analyses possible and open up a myriad of methods for exploring the composition, formation, and evolution of these systems. Here we provide summaries of each of the three main projects involved in this work, as well as a discussion on the importance of these and other scientific studies in astronomy.

5.1 The Perihelion Gap and Planet X

The principle project of this dissertation was published under the title Outer Solar System Perihelion Gap Formation Through Interactions with a Hypothetical Distant Giant Planet (Oldroyd & Trujillo, 2021). This work establishes the outer solar system perihelion gap as a real feature of our solar system, shows how a hypothetical distant giant planet, referred to as Planet X in this work, can produce both the gap and the extreme populations that surround it, and describes the underlying mechanism for how Planet X can form the gap. Topics involving Planet X and the perihelion gap are relatively new and have been frequently debated at conferences and in the literature (see, for example, reviews by Trujillo (2020) and Kavelaars et al. (2020) and references therein). The work presented in Chapter 2 provides useful insight into this discussion as well as additional evidence for the existence of a Planet X.

In Chapter 2, we begin by laying out some of the previous research that has been done surrounding Planet X and the gap. One of the main catalysts of the search for the planet has been the discoveries of the three known Inner Oort Cloud objects (IOCs) Sedna, 2012 VP_{113}, and Leleākūhonua. These distant dwarf planets are particularly interesting as their orbits are not explained by Neptune migration and scattering, the main formational process responsible for the Kuiper belt and scattered disk. Several formation mechanisms have been suggested (e.g., Zderic & Madigan, 2020; Kaib et al., 2011; Morbidelli & Levison, 2004), however, they only partially reproduce the observed structure, namely a gap in perihelion between the IOCs and the Extreme Trans-Neptunian Objects (ETNOs; see Figure 2.1).

The first goal of this research was, then, to determine whether the gap was real, or simply the result of observational biases in the discoveries of the ETNOs and IOCs. We start by setting up a suite of observational simulations designed to mimic many of the characteristics of recent and ongoing large outer solar system surveys. By placing realistic detection limits on synthetic objects in our simulations, we were able to directly compare the simulation results with the real ETNO and IOC populations. So, we developed several different synthetic populations of ETNOs and IOCs and applied our detection limits on each of these to determine the observed distributions for the synthetic populations. We then compared these observed distributions with the real ETNO/IOC population (for example, see Figure 2.2) and found that all of these continuous synthetic populations were poor fits.

Seeing that none of the smooth distributions we tried looked anything like the real distribution[1], we performed a Markov Chain Monte Carlo statistical fit, using two Gaussian curves as the model, in order to find the best possible match between the observed ETNOs/IOCs and a synthetic population. This two-Gaussian fit, shown in Figure 2.3, was approximately 10 million times more probable than any of the other distributions we tested. This strongly supports the hypothesis that the perihelion gap is real, since it fits the data so much better than distributions without a gap.

Once we had determined that the gap was real, we designed an experiment to test if Planet X could be the cause of the gap. In this experiment, we ran hundreds of orbital dynamics simulations to determine the influence of Planet X and the known giant planets on a Kuiper-belt-like distribution of test particles. The first simulations we tried did not include Planet X. We found that just interactions with Neptune[2] do not produce the gap or the ETNOs and IOCs (see Figure 2.4). Next, we included several different versions of Planet X, each with different combinations of mass and orbital elements. We found that some versions of Planet X can cause perihelion gaps near the real gap (see Figure 2.7) and that Planet X can turn objects scattered to high eccentricity by Neptune into ETNOs and IOCs (see Figure 2.5).

After finding that Planet X can cause the gap, the next question we ask relates to how Planet X can form the gap. We noticed in our simulations that through resonances, Planet X clearly causes ETNOs and IOCs to pass through the gap region. However, carefully comparing the amount of time that ETNOs and IOCs

[1]Observing a gap if the gap did not exist was extremely statistically improbable (see Table 2.4).

[2]Neptune is the most important of the giant planets for this population because it orbits so much closer to these objects than do the other planets.

spend in the gap with the time they spend on either side of it shows that these objects spend far less time in the gap than they do in the ETNO or IOC regions (see Figure 2.8). The resonances between Planet X and these objects are such that objects do not remain in the gap for long periods of time. Interestingly, these objects instead tend to end their resonance cycle on either side of the gap. Hence, the gap is a transition region with relatively few objects compared to the surrounding region. We predict that the gap is, therefore, not empty, but that we expect to find roughly 5 times as many objects directly beyond the gap in the IOC region (we have discovered 3 so far) than we will find in the gap[3].

These results provide additional evidence that the existence of Planet X is not only compatible with the perihelion gap, but that Planet X is a plausible explanation for the presence of the perihelion gap in our solar system. This evidence is independent of the observed alignments (Trujillo & Sheppard, 2014; Batygin & Brown, 2016a) for which observational bias has been cited as the main counter-argument (e.g., Kavelaars et al., 2020). For the perihelion gap, observational biases would predict that objects within the gap would be much easier to detect than objects beyond the gap, hence, the existence of the gap strengthens the case for Planet X in the presence of observational biases. Although other hypotheses (e.g., Zderic & Madigan, 2020) have been suggested for gap formation and placing the ETNOs and IOCs on their current orbits, none of them have adequately explained both phenomenon[4]. Hence, this work provides valuable support for the developing hypothesis that Planet X is somewhere on the outskirts of the known solar system, waiting to be discovered.

One question that remains to be answered concerning the gap and its relation to Planet X, is whether simulations will be able to sufficiently constrain the parameters of Planet X by matching the resulting distributions of the simulations with the real distribution of ETNOs and IOCs. The primary difficulty of this method is the high computation cost. Our preliminary tests have shown that simulations on the scale used in this study (many thousands of particles) contain too few particles to produce a statistically significant result. Hence, constraining the parameters of Planet X using this method would require orders of magnitude more particles and, therefore, correspondingly longer compute times. Additionally, the probability of successfully constraining the parameters of Planet X using this method are unknown. It could be quite effective, but it could also provide only broad constraints that are already known using other methods. This uncertainty of success combined with the high computing cost and effort required to carry out such an experiment makes

[3] "A scientific theory without a testable prediction is useless." - Chad Trujillo

[4] Zderic & Madigan (2020) have suggested an alternate hypothesis of a large self-gravitating disk of planetessimals which undergoes an instability in inclination causing features that resemble both the gap and the alignment in the argument of perihelion (both features attributed to Planet X). However, despite some potential, recent work by Das & Batygin (2022) has largely shown their methods to be insufficient and subsequent study has been unable to reproduce these conditions in a realistic solar system scenario.

this method for placing constraints on Planet X less appealing than other methods for placing constraints on its orbit.

Despite the difficulty in constraining the parameters and location of Planet X, there are multiple ongoing surveys searching for the planet. The two surveys with the greatest probability of finding Planet X (in my opinion) are run by (1) Scott Sheppard and Chad Trujillo (Sheppard et al., 2016, 2019) and (2) Mike Brown and Konstantine Batygin (Brown & Batygin, 2021, 2022). These surveys are the only ones I am aware of that consistently use wide-field ($> 1°$) 8 meter class telescopes to survey the sky with proper cadences to find extremely faint (∼25th magnitude) slow moving (∼1”/day) objects in the solar system. Other groups have searched for the planet, but these two surveys are the only ones I am aware of that seem to have a good chance of finding Planet X.

One frequent question regarding the search for Planet X is whether the James Webb Space Telescope (JWST) would be able to find the planet. JWST would be excellent for observing Planet X once it is found, however, JWST’s field of view is not large enough to be effective at surveying the sky for Planet X. The upcoming Legacy Survey of Space and Time (LSST), however, will have a more reliable opportunity to search for Planet X. LSST will survey most of the available sky every three nights. This is an excellent cadence for searching for Planet X. However, it may be the case that Planet X is either too faint to be observed in single exposures from LSST, or the planet may lie outside the designed survey area. If either of these is the case, it will be unlikely that LSST will detect Planet X. Hence, dedicated surveys, such as those mentioned previously, in combination with efficient dynamical constraints, are the most promising options in the search for Planet X.

5.2 Exoastrometry Optimization

Focusing on the dynamics of worlds outside of our solar system, this second project touches more on theoretical optimization methods than on a specific object. The work presented in Chapter 3 is in its final stages and we are preparing the corresponding manuscript for publication. Because next generation space telescopes, such as HabEx and LUVOIR, will be capable of observing Earth-like exoplanets directly as they orbit nearby stars, a high priority for the astronomy community is to increase the efficiency of these and similar missions in order to observe in detail as many potentially habitable worlds (suitable for life) as possible. A significant portion of the observing time for these missions will be allocated to the task of orbit determination. Since exoplanets are much more likely to be habitable if they orbit within the temperate habitable zone around their host star, determining whether an exoplanet orbits in the habitable zone or not in as few observations as possible is essential for efficiency.

Previous studies have suggested that up to five observations are needed in order to determine if a directly imaged exoplanet orbits in its host star's habitable zone (Gaudi et al., 2020). If we can reduce that number to only three or four observations, or reduce the number of system architectures and orientations that require five observations, then we will have saved a large fraction of observing time (∼20% of detection and orbit characterization time). This saved time can then be devoted to further studies of the surfaces, atmospheres, and other physical properties of the most interesting Earth-like exoplanets. Hence, our methods have the potential to significantly increasing the science yield of these missions.

In order to examine the problem, we created a pair of synthetic observations of a single Earth-like exoplanet orbiting a Sun-like star. Then, we generated a suite of potential orbit fits for these observations, including a best-fit orbit that is the current best approximation of the actual orbit based on the observations so far (see Figure 3.1). Our goal was to find the most beneficial time to make a third observation in order to maximize the constraint on the orbit (minimize its uncertainties). So, we developed and tested three methods for determining this optimum observing time.

The first of these methods begins by separating the suite of orbital fits into two groups: orbits that are interior to the best-fit orbit, and orbits that are exterior to the best-fit orbit. We then take representative sample orbits from each group (1σ away from the best-fit orbit within the distribution of sample orbits using Gaussian statistics) and compare the inner and outer sample orbits at equal intervals. The location where these two orbits are farthest apart from each other (when measured at the same mean anomaly) marks the best place to observe (since we know the least about the orbit there) and we then simply calculate the corresponding time to find when it is best to make subsequent observations (see Figure 3.3).

Our second method is quite similar to the first, only we do not separate the orbits or select sample orbits. Instead we take the on-sky separation of all orbits in our sample suite of orbit fits at equal intervals. The place where this cluster of points is the widest is also the least constrained location, and is, therefore, the best location to observe in order to minimize the uncertainties on the orbit (see Figure 3.5). After calculating the time corresponding to this location, we arrive at a similar conclusion to our first method (see Figure 3.4).

In our third method, rather than focusing on sample orbits to find the best time to observe, we instead make synthetic observation at 100 equally spaced times all along the orbit. Then, we run each through a statistical simulation to determine which has the smallest uncertainty. This method appears to be the most straightforward, but it also takes roughly 100 times more computing power than the other methods and produces similar results.

All three methods are dramatically more efficient that the previously suggested equal timing of observations and reduce the number of observations required from five to three in the example shown in Figure 3.7.

By using these improved techniques, current and future direct imaging exoplanet endeavors can significantly improve their science yield by saving time on orbit determination. These methods may also be useful in other scenarios where orbit determination is necessary based only a few closely spaced data points, such as newly discovered binary asteroids, planetary moons, and other solar system objects. This work is nearly completed and will be submitted to the Astronomical Journal in the coming months.

5.3 Active Asteroid Dynamics

In the final project in this dissertation we study the Quasi-Hilda 282P, a minor planet that orbits near the Hilda asteroid group. Hildas orbit in the 3:2 interior mean motion resonance with Jupiter and Quasi-Hildas orbit near this resonance, but are not in it. Activity for 282P has been observed in the past, and our study was motivated by the detection of a secondary epoch of activity through the NASA Partner *Active Asteroids* Citizen Science project.

In order to explore the cause of activity on 282P, we developed a dynamical simulation with 500 orbital clones of 282P. Orbital clones are identical to each other and given slight offsets in their orbital elements within the bounds of the measured orbital uncertainties for 282P. The main reason for making orbital clones of a body is that, when there are close encounters between that body and a planet, tiny differences in the small body's orbital elements can make drastic differences in the outcome of the encounter. For example, one clone may have its orbit completely changed, another may collide with Jupiter, and a third could be ejected from the solar system all from the same encounter with just slightly different orbital elements. This chaotic response necessitates looking at the probability of all the different outcomes of an encounter in order to say anything useful about the resulting orbit of the object.

After simulating the orbits of the 282P clones, both in the future and the past, for 1,000 years, we find that it has undergone roughly ten close encounters with Jupiter in the timespan between 250 years ago and 350 years in the future. These encounters make dramatic differences in 282P's orbit, including shifting it from an orbit between the orbits of Jupiter and Saturn to it current orbit, which is smaller and interior to Jupiter's orbit. Additionally, 282P's Tisserand parameter with respect to Jupiter, an indicator of whether an object is an asteroid or a comet, oscillates between cometary (Jupiter Family Comets) and asteroidal values approximately 10 times in the next 250 years and 282P will alternate between being an asteroid and a comet during this timeframe.

Finding this object in transition with cometary features is a promising clue towards solving the mystery of the origins of the active asteroids. It has been shown that the Jupiter Family Comets come from the more distant Centaurs, which, in turn, come from the Kuiper belt beyond Neptune. Hence, the Quasi-

Hildas may play an important role as a missing link between the outer solar system and the active asteroids, potentially even providing a mechanism for sourcing Earth's water from beyond Neptune in the early stages of the growth of our solar system. In the event that the Quasi-Hildas do play an important transitional role between the Jupiter Family Comets and the active asteroids, they may represent part of a new dynamical class of objects that fill an important void in our understanding of how material is transported throughout the solar system. This work has been included in a publication submitted to Astrophysical Journal Letters. In our submitted work, we deduce that the activity observed on 282P is driven by sublimation.

5.4 Broader Impacts

As we end our discussion on this work, one of the most pertinent questions to ask is *Why do we care?* At times it seems as though academic research of this kind does not have a clear purpose or benefit to society. As a rebuttal to this line of thought, I present three reasons to both care about and support research in dynamical astronomy.

First, new methods in mathematics and new lines of thought in physics are often developed as a result of trying to understand how the universe works. Newton's theory of gravity, Gauss' methods of successive approximation, and Einstein's theory of relativity all came from a desire to understand astronomy at a higher level. These and countless other developments arising from the study of dynamical astronomy shape our view of the world around us and help us to understand the basic principles that govern it. Perhaps the transition of active asteroids from the outer solar system to the inner will provide new insights into how the natural world works.

The second reason is the development of new technology. Advances such as digital cameras and portable computers were heavily influence by advances in astronomy and space exploration and there are many others directly and indirectly resulting from astronomy research. These technologies dramatically improve the quality of life for countless people and future advances will continue to do so. It may be that our methods for optimizing exoplanet orbit determination could be valuable in other fields as well and the techniques we use or derivatives of them could lead to new advancement in technology yet to be imagined.

A third and final reason to care about and support astronomy is more abstract than the first two, but also even more important; new inspiration. When Pluto was discovered in 1930 it united people in wonder and appreciation for something so far beyond themselves as individuals. The discovery of something new in the universe is amazing and inspiring and it has the ability to give purpose and passion to our lives. There is a reason that over 90 years after its discovery and more than 15 years after it was no longer classified as a planet, when asked what their favorite planet is, many people will still say it is Pluto. That is the power of discovery. That is why science fiction is so prevalent and well-loved in our society. It is also why nations invest in space programs and research. Because the wonder of discovery makes us better people both individually and as a society. When Planet X is discovered, it will be as exciting, if not more so, as when Pluto was discovered. It will bring people together across boundaries and will make the world a better place by inspiring countless people to find greater perspective in their lives. It is difficult to predict whether a particular research project may result in a new technology or a Nobel prize-winning scientific advance, but as we continue to seek and discover new things about our solar system and the universe beyond it, we will doubtless inspire others to reach for the stars.

References

Adams, J. 1846, On the Perturbations of Uranus (London: W. Clowes), 265

Agnor, C. B., & Hamilton, D. P. 2006, Nature, 441, 192

Ahn, C. P., Alexandroff, R., Allende Prieto, C., et al. 2012, ApJ Supplement Series, 203, 21

Apian, P. 1524, Cosmographia

Arnaboldi, M., Capaccioli, M., Mancini, D., et al. 1998, The Messenger, 93, 30

Bailer-Jones, C. A. L., Farnocchia, D., Ye, Q., Meech, K. J., & Micheli, M. 2020, A&A, 634, A14

Bailes, M., Bates, S. D., Bhalerao, V., et al. 2011, Science, 333, 1717

Bailey, E., Batygin, K., & Brown, M. E. 2016, AJ, 152, 126

Bailey, E., Brown, M. E., & Batygin, K. 2018, AJ, 156, 74

Bailey, V., Meshkat, T., Reiter, M., et al. 2014, ApJL, 780, L4

Barucci, M. A., & Merlin, F. 2020, in The Trans-Neptunian Solar System, ed. D. Prialnik, M. A. Barucci, & L. Young, 109–126

Batygin, K., Adams, F. C., Brown, M. E., & Becker, J. C. 2019, PhR, 805, 1

Batygin, K., & Brown, M. E. 2016a, AJ, 151, 22

—. 2016b, ApJL, 833, L3

Batygin, K., & Morbidelli, A. 2017, AJ, 154, 229

Becker, J. C., Adams, F. C., Khain, T., Hamilton, S. J., & Gerdes, D. 2017, AJ, 154, 61

Berthier, J., Vachier, F., Thuillot, W., et al. 2006, SkyBoT, a New VO Service to Identify Solar System Objects

Blunt, S., Nielsen, E. L., De Rosa, R. J., et al. 2017, AJ, 153, 229

Blunt, S., Wang, J. J., Angelo, I., et al. 2020, AJ, 159, 89

Bode, J. E. 1772, Anleitung zur Kenntnis des gestirten Himmels

Bonamente, M. 2017, Statistics and Analysis of Scientific Data, 2nd edn. (New York: Springer)

Bouvard, A. 1821, Tables astronomiques publiees par le Bureau des Longitudes de France

Bowell, E., Muinonen, K., & Wasserman, L. H. 1994, Symposium-International ...

Brasser, R., Duncan, M. J., Levison, H. F., Schwamb, M. E., & Brown, M. E. 2012, Icarus, 217, 1

Brown, M., & Batygin, K. 2019, in EPSC-DPS Joint Meeting 2019 (Geneva: Europlanet Society), EPSC–DPS2019–1994

Brown, M. E., & Batygin, K. 2016, ApJL, 824, L23

—. 2021, AJ, 162, 219

—. 2022, AJ, 163, 102

Brown, M. E., Trujillo, C., & Rabinowitz, D. 2004, ApJ, 617, 645

Brown, M. E., Trujillo, C. A., & Rabinowitz, D. 2005, IAU Circular, 8577, 1

Brunini, A. 1992, Celestial Mechanics and Dynamical Astronomy, 53, 129

Cameron, A. G. W., & Truran, J. W. 1977, Icarus, 30, 447

Canup, R. M., & Asphaug, E. 2001, Nature, 412, 708

Carrera, D., Gorti, U., Johansen, A., & Davies, M. B. 2017, ApJ, 839, 16

Carter, A. L., Hinkley, S., Bonavita, M., et al. 2021, MNRAS, 501, 1999

Chambers, J. E. 2016, ApJ, 825, 63

Chandler, C. O., Curtis, A. M., Mommert, M., Sheppard, S. S., & Trujillo, C. A. 2018, Publications of the Astronomical Society of the Pacific, 130, 114502

Chandler, C. O., Kueny, J., Gustafsson, A., et al. 2019, The Astrophysical Journal Letters, 877, L12

Chandler, C. O., Kueny, J. K., Trujillo, C. A., Trilling, D. E., & Oldroyd, W. J. 2020, The Astrophysical Journal Letters, 892, L38

Chandler, C. O., Trujillo, C. A., & Hsieh, H. H. 2021, The Astrophysical Journal, 922, L8

Christenson, A. J. 2003, Popol Vuh: The Sacred Book of the Maya, Winchester, Inglaterra: O Books.

Christy, J. W., & Harrington, R. S. 1978, AJ, 83, 1005

Clement, M. S., & Kaib, N. A. 2020, AJ, 159, 285

Copernicus, N. 1543, De Revolutionibus Orbium Coelestium (Johannes Petreius, Nuremberg)

da Costa-Luis, C., Larroque, S. K., Altendorf, K., et al. 2020, tqdm: A fast, Extensible Progress Bar for Python and CLI, v4.54.0, Zenodo

—. 2022, tqdm: A fast, Extensible Progress Bar for Python and CLI, v4.64.0, Zenodo

Das, A., & Batygin, K. 2022, in AAS/Division of Dynamical Astronomy Meeting, Vol. 54, AAS/Division of Dynamical Astronomy Meeting, 202.08

de Jong, J. T. A., Verdoes Kleijn, G. A., Boxhoorn, D. R., et al. 2015, Astronomy and Astrophysics, 582, A62

de la Fuente Marcos, C., & de la Fuente Marcos, R. 2019, A&A, 621, A137

Dermott, S. F., & Murray, C. D. 1983, Nature, 301, 201

Doppler, C. 1903, Ueber das farbige Licht der Doppelsterne und einiger anderer Gestirne des Himmels: Versuch einer das Bradley'sche Aberrations-Theorem als integrirenden Theil in sich schliessenden allgemeineren Theorie (K. Böhm Gesellschaft der Wissenschaften)

Dressler, A., Bigelow, B., Hare, T., et al. 2011, Publications of the Astronomical Society of the Pacific, 123, 288

Duncan, M., Levison, H., & Dones, L. 2004, in Comets II, ed. M. C. Festou, H. U. Keller, & H. A. Weaver, 193

Duncan, M., Quinn, T., & Tremaine, S. 1987, AJ, 94, 1330

Edgeworth, K. E. 1949, MNRAS, 109, 600

Einstein, A. 1920, Relativity: the Special and General Theory (Methuen & Co. Ltd.)

Ferguson, J., & Bell, A. 1771, Astronomy in Encyclopaedia Britannica 1st Edition (Encyclopaedia Britannica Inc.)

Fienga, A., Di Ruscio, A., Bernus, L., et al. 2020, A&A, 640, A6

Foderà Serio, G., Manara, A., & Sicoli, P. 2002, Giuseppe Piazzi and the Discovery of Ceres in Asteroids III, 17–24

Foreman-Mackey, D. 2016, JOSS, 1, 24

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306

Fortney, J. J., Dawson, R. I., & Komacek, T. D. 2021, Journal of Geophysical Research (Planets), 126, e06629

Gaia Collaboration. 2018, VizieR Online Data Catalog, I/345

Galilei, G. 1610, Sidereus nuncius magna, longeque admirabilia spectacula pandens lunae facie, fixis innumeris, lacteo circulo, stellis nebulosis, ... Galileo Galileo : nuper a se reperti beneficio sunt observata in apprime vero in quatuor planetis circa Iovis stellam disparibus intervallis, atque periodis, celeritate mirabili circumvolutis ... atque Medicea sidera nuncupandos decrevit

Gaudi, B. S., Seager, S., Mennesson, B., et al. 2018, Nature Astronomy, 2, 600

—. 2020, arXiv e-prints, arXiv:2001.06683

Gauss, K. F. 1809, Theoria motvs corporvm coelestivm in sectionibvs conicis solem ambientivm.

Ghez, A. M., Salim, S., Weinberg, N. N., et al. 2008, ApJ, 689, 1044

Gil-Hutton, R., & García-Migani, E. 2016, A&A, 590, A111

Giorgini, J. D., Yeomans, D. K., Chamberlin, A. B., et al. 1996, BAAS, 28, 1158

Gladman, B., & Chan, C. 2006, ApJL, 643, L135

Gladman, B., Marsden, B. G., & Vanlaerhoven, C. 2008, in The Solar System Beyond Neptune, ed. M. A. Barucci, H. Boehnhardt, D. P. Cruikshank, A. Morbidelli, & R. Dotson, 43

Gomes, R., Levison, H. F., Tsiganis, K., & Morbidelli, A. 2005, Nature, 435, 466

Gomes, R. S., Fernández, J. A., Gallardo, T., & Brunini, A. 2008, in The Solar System Beyond Neptune, ed. M. A. Barucci, H. Boehnhardt, D. P. Cruikshank, A. Morbidelli, & R. Dotson, 259

Gould, B. A. 1852, AJ, 2, 80

Grav, T., Holman, M. J., Gladman, B. J., & Aksnes, K. 2003, Icarus, 166, 33

Grundy, W. M., Noll, K. S., Virtanen, J., et al. 2008, Icarus, 197, 260

Grundy, W. M., Noll, K. S., Roe, H. G., et al. 2019, Icarus, 334, 62

Guimond, C. M., & Cowan, N. B. 2019, AJ, 157, 188

Gulbis, A. A. S., Elliot, J. L., Adams, E. R., et al. 2010, AJ, 140, 350

Guyon, O., Pluzhnik, E. A., Galicher, R., et al. 2005, ApJ, 622, 744

Hahn, J. M., & Malhotra, R. 1999, AJ, 117, 3041

Hamilton, D. P., & Burns, J. A. 1992, Icarus, 96, 43

Harris, A. W., & Harris, A. W. 1997, Icarus, 126, 450

Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Nature, 585, 357

Hatzes, A. P. 2016, Space Science Reviews, 205, 267

Hernandez, D. M., Zeebe, R. E., & Hadden, S. 2022, MNRAS, 510, 4302

Herschel, W. 1802, Philosophical Transactions of the Royal Society of London Series I, 92, 213

Hesselbrock, A. J., & Minton, D. A. 2017, Nature Geoscience, 10, 266

Hill, G. W. 1878, American Journal of Mathematics, 1, 5

Holman, M. J. 1994, PhD thesis, Massachusetts Institute of Technology

Holman, M. J., & Payne, M. J. 2016, AJ, 152, 80

Homer. c. 630 BC, The Iliad of Homer

Hsieh, H. H., & Jewitt, D. 2006, Science, 312, 561

Hsieh, H. H., Yang, B., & Haghighipour, N. 2012, ApJ, 744, 9

Hsieh, H. H., Denneau, L., Wainscoat, R. J., et al. 2015, Icarus, 248, 289

Hunter, J. D. 2007, CSE, 9, 90

Jenkins, J. M. 2002, Galactic Alignment: The Transformation of Consciousness According to Mayan, Egyptian and Vedic Traditions (Bear & Company)

Jewitt, D. 2009, AJ, 137, 4296

Jewitt, D. 2012, The Astronomical Journal, 143, 66

Jewitt, D., & Haghighipour, N. 2007, ARA&A, 45, 261

Jewitt, D., Hsieh, H., & Agarwal, J. 2015, in Asteroids IV, 221–241

Jewitt, D., & Kim, Y. 2020, PSJ, 1, 77

Jewitt, D., & Luu, J. 1993, Nature, 362, 730

Johnson, H. L., Mitchell, R. I., Iriarte, B., & Wisniewski, W. Z. 1966, Communications of the Lunar and Planetary Laboratory, 4, 99

Johnson, T. V., & Lunine, J. I. 2005, Nature, 435, 69

Joye, W. A. 2006, in Astronomical Data Analysis Software and Systems XV ASP Conference Series, Vol. 351, 574–

Juric, M., Eggl, S., Moeyens, J., et al. 2020, in American Astronomical Society Meeting Abstracts, 220.06

Kaib, N. A., Roškar, R., & Quinn, T. 2011, Icarus, 215, 491

Kaib, N. A., & Volk, K. 2022, arXiv e-prints, arXiv:2206.00010

Karttunen, H., Kröger, P., Oja, H., Poutanen, M., & Donner, K. J., eds. 2007, Fundamental Astronomy, 5th edn. (Berlin: Springer)

Kasting, J. F., Whitmire, D. P., & Reynolds, R. T. 1993, Icarus, 101, 108

Kavelaars, J. J., Lawler, S. M., Bannister, M. T., & Shankman, C. 2020, in The Trans-Neptunian Solar System, ed. D. Prialnik, M. A. Barucci, & L. Young (Amsterdam: Elsevier), 61

Kenyon, S. J., & Bromley, B. C. 2004, Nature, 432, 598

Kenyon, S. J., & Luu, J. X. 1999, ApJ, 526, 465

Kepler, J. 1619, Ioannis Keppleri harmonices mundi libri V, ed. J. Planck (Godofredi Tampachii Linz, Austria), 189

Kepler, J., Copernicus, N., Mästlin, M., & Schöner, J. 1596, Prodromus dissertationum cosmographicarum, continens mysterium causis coelorum numeri, magnitudinis, motuumque periodicorum genuinis cosmographicum, de admirabili proportione orbium coelestium, deque & propriis, demonstratum, per quinque regularia corpora geometrica

Khain, T., Becker, J. C., & Adams, F. C. 2020, PASP, 132, 124401

Kirkwood, D. 1867, Meteoric astronomy: A treatise on shooting-stars, fire-balls, and aerolites

Kollmeier, J. A., & Raymond, S. N. 2019, MNRAS, 483, L80

Konopacky, Q. M., Marois, C., Macintosh, B. A., et al. 2016, AJ, 152, 28

Kopparapu, R. K., Ramirez, R., Kasting, J. F., et al. 2013, ApJ, 765, 131

Kraft, R. P., Burrows, D. N., & Nousek, J. A. 1991, ApJ, 374, 344

Kuijken, K. 2011, The Messenger, 146, 8

Kuijken, K., Bender, R., Cappellaro, E., et al. 2002, The Messenger, 110, 15

Kuiper, G. P. 1951, Proceedings of the National Academy of Science, 37, 1

Lang, D., Hogg, D. W., Mierle, K., Blanton, M., & Roweis, S. 2010, Astronomical Journal, 139, 1782

Le Verrier, U. J. 1846, Astronomische Nachrichten, 25, 65

Leonard, F. C. 1930, Leaflet of the Astronomical Society of the Pacific, 1, 121

Levison, H. F. 1996, in Astronomical Society of the Pacific Conference Series, Vol. 107, Completing the Inventory of the Solar System, ed. T. Rettig & J. M. Hahn, 173–191

Levison, H. F., Duncan, M. J., Brasser, R., & Kaufmann, D. E. 2010, Science, 329, 187

Levison, H. F., Morbidelli, A., Tsiganis, K., Nesvorný, D., & Gomes, R. 2011, AJ, 142, 152

Levison, H. F., Morbidelli, A., Van Laerhoven, C., Gomes, R., & Tsiganis, K. 2008, Icarus, 196, 258

Li, G., Hadden, S., Payne, M., & Holman, M. J. 2018, AJ, 156, 263

Lowell, P. 1915, Memoirs of the Lowell Observatory. John Wiley & Sons, 1, 1

Luu, J. X., & Jewitt, D. 1988, AJ, 95, 1256

MacKay, D. J. C. 2003, Information Theory, Inference, and Learning Algorithms (Cambridge University Press)

Malhotra, R., Volk, K., & Wang, X. 2016, ApJL, 824, L22

Margot, J.-L. 2015, AJ, 150, 185

Marois, C., Macintosh, B., Barman, T., et al. 2008, Science, 322, 1348

Mayor, M., & Queloz, D. 1995, Nature, 378, 355

McKinney, W. 2010, in Proc. 9th Python in Science Conf., ed. S. van der Walt & J. Millman (Austin, TX: SciPy), 56

Metropolis, N. 1987, Los Alamos Science Special Issue, 15, 125

Millholland, S., & Laughlin, G. 2017, AJ, 153, 91

Montmerle, T., Augereau, J.-C., Chaussidon, M., et al. 2006, Earth Moon and Planets, 98, 39

Morbidelli, A., Chambers, J., Lunine, J. I., et al. 2000, Meteoritics & Planetary Science, 35, 1309

Morbidelli, A., & Levison, H. F. 2004, AJ, 128, 2564

Morbidelli, A., & Nesvorný, D. 2020, in The Trans-Neptunian Solar System, ed. D. Prialnik, M. A. Barucci, & L. Young (Amsterdam: Elsevier), 25

Morgan, R., Savransky, D., Turmon, M., et al. 2021, Journal of Astronomical Telescopes, Instruments, and Systems, 7, 021220

Moskovitz, N., Burt, B., Schottland, R., et al. 2021, AAS Division of Planetary Science meeting #53, id. 101.04, 53, 101.04

Murison, M. 2006, A Practical Method for Solving the Kepler Equation, Technical Report

Murray, C. D., & Dermott, S. F. 1999, Solar System Dynamics (Cambridge: Cambridge Univ. Press)

Napier, K. J., Gerdes, D. W., Lin, H. W., et al. 2021, PSJ, 2, 59

National Academies of Sciences, Engineering, a. 2021, Pathways to Discovery in Astronomy and Astrophysics for the 2020s

Nesvorný, D. 2015, AJ, 150, 68

—. 2018, ARA&A, 56, 137

Nesvorný, D., Vokrouhlický, D., Dones, L., et al. 2017, ApJ, 845, 27

Newton, I. 1687, Philosophiae Naturalis Principia Mathematica. Auctore Js. Newton

Ochsenbein, F., Bauer, P., & Marcout, J. 2000, Astronomy and Astrophysics Supplement, 143, 23

Oldroyd, W. J., Ragozzine, D., & Porter, S. 2018, in AAS/Division of Dynamical Astronomy Meeting, 402.03

Oldroyd, W. J., & Trujillo, C. A. 2021, AJ, 162, 39

Oort, J. H. 1950, Bulletin of the Astronomical Institutes of the Netherlands, 11, 91

Petit, J.-M., & Morbidelli, A. 2005, in Lectures in Astrobiology, Volume 1, ed. M. Gargaud, B. Barbier, H. Martin, & J. Reisse, Vol. 1, 61

Piazzi, G. 1801, Risultati delle osservazioni della nuova stella scoperta il di' 1. Gennaio all'Osservatorio Reale di Palermo

Piazzi, G., & von Zach, F. X. 1801, Monatliche Correspondenz, Sept 1801, 280

Pravec, P., Thomas, C. A., Rivkin, A. S., et al. 2022, PSJ, 3, 175

Press, W., Teukolsky, S., Vetterling, W., & Flannery, B. 2007, Numerical Recipes 3rd Edition: The Art of Scientific Computing (Cambridge: Cambridge Univ. Press)

Price-Whelan, A. M., Sipőcz, B. M., Günther, H. M., et al. 2018, AJ, 156, 123

Ptolemy, C., & Toomer, G. 1984, Ptolemy's Almagest, 2nd centruy AD, translated and annotated by G.J. Toomer (Duckworth)

Rayleigh, L. 1879, Philosophical Magazine, 5, 8, 261

Rein, H., & Liu, S. F. 2012, A&A, 537, A128

Rein, H., & Spiegel, D. S. 2015, MNRAS, 446, 1424

Rein, H., & Tamayo, D. 2015, MNRAS, 452, 376

Rein, H., Hernandez, D. M., Tamayo, D., et al. 2019, MNRAS, 485, 5490

Ribeiro de Sousa, R., Morbidelli, A., Gomes, R., et al. 2022, Icarus, 379, 114933

Rivkin, A. S., Chabot, N. L., Stickle, A. M., et al. 2021, PSJ, 2, 173

Roberge, A., & Moustakas, L. A. 2018, Nature Astronomy, 2, 605

Sarid, G., Volk, K., Steckloff, J. K., et al. 2019, ApJL, 883, L25

Schaffer, S. 1981, Journal for the History of Astronomy, 12, 11

Schwamb, M. E., Brown, M. E., Rabinowitz, D. L., & Ragozzine, D. 2010, ApJ, 720, 1691

Sheppard, S. S., Tholen, D. J., & Trujillo, C. A. 2021, Minor Planet Electronic Circulars, 2021-C187

Sheppard, S. S., & Trujillo, C. 2015, AJ, 149, 44

—. 2016, AJ, 152, 221

Sheppard, S. S., Trujillo, C., & Tholen, D. J. 2016, ApJL, 825, L13

Sheppard, S. S., Trujillo, C. A., Oldroyd, W. J., Tholen, D. J., & Williams, G. V. 2018, Minor Planet Electronic Circulars, 2018-Y14

Sheppard, S. S., Trujillo, C. A., Tholen, D. J., & Kaib, N. 2019, AJ, 157, 139

Siraj, A., & Loeb, A. 2020, ApJL, 899, L24

Snodgrass, C., Agarwal, J., Combi, M., et al. 2017, The Astronomy and Astrophysics Review, 25, 5

Sori, M. M., Sizemore, H. G., Byrne, S., et al. 2018, Nature Astronomy, 2, 946

Soto, G. J., Savransky, D., Garrett, D., & Delacroix, C. 2019, Journal of Guidance Control Dynamics, 42, 2671

Spahn, F., Schmidt, J., Albers, N., et al. 2006, Science, 311, 1416

Stukeley, W. 1752, Memoirs of Sir Isaac Newton's Life, Royal Society Library, London, UK 2004

The LUVOIR Team. 2019, arXiv e-prints, arXiv:1912.06219

Thomas, C., Rivkin, A., Pravec, P., et al. 2018, in AGU Fall Meeting Abstracts, Vol. 2018, P51A–02

Tombaugh, C. W. 1961, in Planets and Satellites, ed. G. P. Kuiper & B. M. Middlehurst, 12

Toth, I. 2006, A&A, 448, 1191

Trujillo, C. 2020, in The Trans-Neptunian Solar System, ed. D. Prialnik, M. A. Barucci, & L. Young (Amsterdam: Elsevier), 79

Trujillo, C. A., & Brown, M. E. 2001, ApJL, 554, L95

Trujillo, C. A., Brown, M. E., Rabinowitz, D. L., & Geballe, T. R. 2005, ApJ, 627, 1057

Trujillo, C. A., & Sheppard, S. S. 2014, Nature, 507, 471

Tsiganis, K., Gomes, R., Morbidelli, A., & Levison, H. F. 2005, Nature, 435, 459

Turtle, E. P., Perry, J. E., Hayes, A. G., et al. 2011, Science, 331, 1414

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261

Volk, K., & Malhotra, R. 2008, ApJ, 687, 714

—. 2017, AJ, 154, 62

Voltaire. 1727, An Essay upon the Civil Wars of France, extracted from curious Manuscripts and also upon the Epick Poetry of the European Nations, from Homer down to Milton (Samuel Jallasson: London, England)

von Zach, F. X. 1801, Monatliche Correspondenz zur Beforderung der Erd- und Himmels-kunde, 3, 592

Whitmell, C. T. 1907, Obs., 30, 96

Winn, J. N., & Fabrycky, D. C. 2015, ARA&A, 53, 409

Wisdom, J. 2015, AJ, 150, 127

Wright, J. T., & Gaudi, B. S. 2013, in Planets, Stars and Stellar Systems. Volume 3: Solar and Stellar Planetary Systems, ed. T. D. Oswalt, L. M. French, & P. Kalas, 489

Zderic, A., & Madigan, A.-M. 2020, AJ, 160, 50