The Kuiper Belt

Harold F. Levison

Department of Space Sciences

Southwest Research Institute, Boulder, CO 80302

hal@gort.boulder.swri.edu

and

Paul R. Weissman

Earth and Space Sciences Division

Jet Propulsion Laboratory, Pasadena, CA 91109

pweissman@issac.jpl.nasa.gov

Table of Contents

Glossary

Albedo: The fraction of incident light reflected from a surface. An albedo of 1.0 indicates a surface that is totally reflecting; an albedo of zero indicates a surface that is totally absorbing.

Astronomical unit, AU: Formally, the distance from the Sun to a massless object in an unperturbed circular orbit which has an orbital period of 365.2568983 days, equal to $1.4959787066 \times 10^{11}$ meters. The AU is commonly taken to be the average distance of the Earth from the Sun, and is the unit of distance used by planetary astronomers.

Chaotic motion: A dynamical situation where the error in the prediction of the long-term motion of a body grows exponentially over time. This exponential growth leads to an inability to predict the long-term behavior of the body. Chaotic motion can be confined to fairly narrow regions of space so that the orbit of the body will not change much over time. However, this is unusual for objects in the solar system. Most often chaotic motion will lead to sudden and drastic changes in the orbit of a body.

Centaur: A body in a heliocentric orbit with an average distance from the Sun the lies between the orbits of Jupiter and Neptune. Typically the orbits of these objects also cross one or more of the orbits of the giant planets Saturn, Uranus, and Neptune. Most Centaurs have likely been perturbed out of the Kuiper belt and/or the scattered disk, and some fraction of them will evolve to and appear as Jupiter-family comets. Centaurs are part of the ecliptic comets population.

Eccentricity: The measure of the departure of an elliptical orbit from circularity. The eccentricity, e, is equal to $(1 - b^2/a^2)^{1/2}$, where a and b are the major and minor axes of the elliptical orbit. Circular orbits have $e = 0$, elliptical orbits have $0 < e < 1$, radial and parabolic orbits have $e = 1$, and hyperbolic orbits have $e > 1$.

Ecliptic: The plane of the Earth's orbit around the Sun. The planets, most asteroids, most of the short-period comets, and the Kuiper belt are all in orbits with small or moderate inclinations (or tilts) relative to the ecliptic.

Ecliptic comet: An object which has escaped the Kuiper belt or the scattered disk and which is being scattered inward to an orbit crossing one or more of the planets. The designation of such objects as comets assumes an icy composition because of their likely formation far from the Sun in the colder regions of the solar nebula. Ecliptic comets include Centaur objects and Jupiter-family comets.

Inclination: A measure of the orientation of an orbit. The inclination is the angle between the plane of the orbit and some reference plane, usually the ecliptic.

Invariable plane: The plane passing through the barycenter (center-of-mass) of the solar system and perpendicular to the total angular momentum vector of the solar system.

Jupiter-family comet: An active comet in a low to moderate inclination orbit with a semi-major axis less than that of Jupiter's orbit. Most Jupiter-family comets are in orbits which cross or can closely approach Jupiter's orbit. The Jupiter-family comets are a subset of the ecliptic comets.

Kuiper belt: A collection of a billion or more bodies in low eccentricity, low inclination orbits beyond Neptune, extending out possibly as far as 1,000 astronomical units.

Longitude of the ascending node: A measure of the orientation of an orbit. The *nodes* of an orbit are the points where the orbit crosses some reference plane, usually the ecliptic. The *ascending node* is where the nodal crossing is in the upward direction (where "upward" is defined by some reference direction, typically the angular momentum vector of the Earth's orbit). The *longitude of the ascending node* is the angle between the location of the ascending node and some standard direction in the reference frame, usually the direction of the vernal equinox.

Mean motion resonance: A dynamical situation where the ratio of the orbital periods of two orbiting objects can be expressed as the ratio of two small integers. Mean motion resonances can lead to strong changes in the orbit of one or more of the bodies, or can actually enhance orbital stability depending on the precise nature of the resonance.

Precession: The slow, smooth increase or decrease of an angle.

Perihelion: The location in an orbit of closest approach to the Sun.

Planetesimal: A small body formed in the early solar system by accretion of dust and ice (if present) near the central plane of the solar nebula.

Scattered disk: A collection of icy planetesimals in high eccentricity orbits in the ecliptic plane beyond Neptune. Objects in the scattered disk may be escapees from the Kuiper belt and/or may be scattered Uranus-Neptune planetesimals.

Secular resonance: A dynamical situation where there is a commensurability among the frequencies associated with the rates of precession of the argument of perihelion or of the regression of the nodal line of two bodies. Secular resonances can lead to very large changes in the orbit of one or more of the bodies.

Semi-major axis: One half of the major axis of an elliptical orbit. The semi-major axis is commonly thought of as the average distance between an object and the body it is in orbit about.

Trojan asteroids: Asteroids which are locked in a 1:1 mean motion resonance with Jupiter and which librate about the L_4 and L_5 Lagrange points, 60° ahead and behind Jupiter in its orbit. Trojan-type orbits are also possible for objects in 1:1 resonances with other planets. A Trojan-type librating asteroid has been found for Mars.

Vernal equinox: The direction of the Sun as viewed from the Earth as it crosses the celestial equator moving northward. Denotes the beginning of spring in the northern hemisphere.

I. Historical Perspective

Since its discovery in 1930, Pluto has traditionally been viewed as the last vestige of the planetary system — a lonely outpost at the edge of the solar system, orbiting beyond Neptune with a 248 year period. Pluto receives very little light from the Sun (being almost 40 times farther from the Sun on average than the Earth) and thus it is very cold. The view was that it was a distant, isolated, and unfriendly place, with nothing of substance beyond it.

Pluto itself has always appeared to be an oddity among the planets. Traditionally, the planets are divided into two main groups. The first group, the terrestrial planets, formed in the inner regions of the solar system where the material from which the planets were made was too warm for water and other volatile gases to be condensed as ices. These planets, which include the Earth, are small and rocky. Farther out from the Sun, the cores of the planets grew from a combination of rock and condensed ices, and captured significant amounts of nebula gas. These are the Jovian planets, the giants of the solar system; they most likely do not have solid surfaces. But, then there is Pluto, unique, small (its radius is only $\sim$1180 km) and made of a mix of rock and frozen ices.

The planets formed in a disk of material that originally surrounded the Sun. As the Sun formed from the collapse of its parent molecular cloud, it faced a problem. The cloud had a slight spin and as it collapsed, the spin rate had to increase in order to conserve angular momentum. The cloud could not form a single star with the amount of angular momentum it possessed, so it shed a disk of material that contained very little mass (as compared with the mass of the Sun), but most of the angular momentum of the system. As such, the planets formed in a narrow disk structure; the plane of that disk is known as the *invariable plane*. But, then there is Pluto, unique, having an orbital inclination of 15.6° with respect to the invariable plane.

The orbits of the planets are approximately ellipses with the Sun at one focus. As the planets formed in the original circumsolar disk, they tended to evolve onto orbits that were well separated from one another. This was required so that their mutual gravitational attraction would not disrupt the whole system. (Or to put is another way, if our system had not formed that way we would not be here to talk about it!) But, then there is Pluto, unique, having an orbit that crosses the orbit of its nearest neighbor, Neptune.

So, the historical view was that Pluto was an oddity in the solar system. Unique for its physical makeup and size as well as its dynamical niche. But, this view changed in September, 1992 with the announcement of the discovery of the first of a population of small (compared to planetary bodies) objects orbiting beyond the orbit of Neptune, in the same region as Pluto. Since that time over 60 objects with radii between about 50 and 500 km have been discovered. Given the very small area of the sky which has been searched so far, these 60+ objects imply that there are approximately $70,000$ similar objects occupying this region of space, between 30 and 50 AU from the Sun. There are almost certainly many more smaller ones. As discussed in more detail in the following sections, these objects likely have a similar physical makeup to that of Pluto and many have similar orbital characteristics. Thus, in the last few years, Pluto has been transformed from a oddity, to the founding member of what is perhaps the most populous class of objects in the planetary system.

The discovery of the *Kuiper belt*, as it has come to be known, represents a revolution in our thinking about the solar system. First predicted on theoretical grounds and later confirmed by observations, the Kuiper belt is the first totally new class of bodies to be discovered in the solar system since the first asteroid was found on New Year's day, 1801. Its discovery is on a par with the discovery of the solar wind and the planetary magnetospheres in the 1950's and 60's, and it has radically changed our view of the outer solar system.

Speculation on the existence of a trans-Neptunian disk of icy objects dates back over 50 years. In the 1940's and early 1950's, Kenneth Edgeworth and Gerard

Kuiper independently considered the structure of our planetary system. They noticed that if one where to grind up the giant planets and smooth out their masses to form a disk, that this disk would have a very smooth distribution, with a density that decreased as the distance from the Sun. That is until Neptune, at which point there is an apparent edge, beyond which there was thought to be nothing except tiny Pluto. Edgeworth and Kuiper suggested that perhaps this edge was not real. Perhaps the disk of planetesimals that formed the planets extended past Neptune, but that the density was too low or the formation times to long to form large planets. If so, they argued, these planetesimals should still be there in nearly circular orbits beyond Neptune. Unfortunately, Edgeworth's contribution was overlooked until recently, and thus this disk has come to be known as the Kuiper belt.

The idea of a trans-Neptunian disk received little attention for many years. The objects in the hypothetical disk were too faint to be seen with the telescopes of the time, so there was no way to prove or disprove their existence. Comet dynamicists showed that the lack of detectable perturbations on the orbit of Halley's Comet limited the mass of such a disk to no more than 1.3 Earth masses ($M_{\oplus}$) if it was at 50 AU from the Sun.

However, the idea was resurrected in 1980 when Julio Fernandez proposed that a cometary disk beyond Neptune could be a possible source reservoir for the short-period comets (those with orbital periods $<$ 200 years). Subsequent dynamical simulations showed that a comet belt beyond Neptune is the most plausible source for the low inclination, Jupiter-family comets. This work led observers to search for Kuiper belt objects, the first one being discovered in August, 1992. Since then over 60 Kuiper belt objects with estimated radii $\gtrsim 50$ km (assuming a typical cometary nucleus albedo of 0.04) have been found by ground-based searches, and $\sim$30 comet-sized (radii $\lesssim 10$ km) object have been detected using the Hubble Space Telescope. (Since the Kuiper belt is believed to be the source of the Jupiter-family short-period comets, typical cometary albedos of 4% are assumed for the surfaces of Kuiper belt objects).

There are two populations of objects in the solar system that are related to the Kuiper belt. The first consists of a large number ($\sim 10^6$) of (presumably) icy objects on planet-crossing orbits interior to Neptune's orbit. These objects are known as *ecliptic comets* and they most likely originated in the Kuiper belt. Ecliptic comets include two distinct sub-populations: the Centaurs and the Jupiter-family comets. The known Centaurs are 20–100 km-radius objects with semi-major axes beyond the orbit of Jupiter, most of which appear to be inactive (the one known exception is the first and largest Centaur discovered, 2060 Chiron). Jupiter-family comets are much smaller objects a few kilometers in diameter, on Jupiter-crossing orbits but with semi-major axes interior to Jupiter. Most of known Jupiter-family comets are active comets since active comets are much brighter than inactive ones.

Both the Centaurs and the Jupiter-family comets are just the brightest members of the ecliptic comets, i.e., the most easily discovered. The active Jupiter-family comets are bright because they are relatively close to the Earth and because they produce active cometary comae. The Centaurs are bright because they are relatively big.

The second related population of objects is the known as the *scattered comet disk*. It was suggested by a number of theoretical researchers in the 1980's, predicted by numerical models in the 1990's, and then the first member was discovered by observers in 1996. The scattered disk occupies the same volume of physical space as the Kuiper belt, but as discussed below, it has a different dynamical character and a different origin.

In this chapter we discuss the available information on the Kuiper belt, scattered disk, and Centaurs (for a discussion of the Jupiter-family comets see: COMETARY DYNAMICS, PHYSICS AND CHEMISTRY OF COMETS). The observational information comes in the form of: *i*) the orbits determined for the discovered objects, *ii*) the sizes of the objects estimated from their brightnesses, and *iii*) the physical observations (for example, broad-band colors) of the objects. In addition, there is a vast body of theoretical studies of the trans-Neptunian region, based primarily on computer-based simulations. These simulations have provided considerable insight into the motion, spatial distributions, size distribution, and evolution of bodies in this region, as well as their relationship to other solar system populations.

But, why study the Kuiper belt and the scattered disk? It is rare in the history of astronomy that a whole new region of the solar system is discovered that needs to be understood. For that reason alone, the Kuiper belt deserves study. Indeed, as the following sections illustrate, the structure of the orbits of objects in the trans-Neptunian region is much more complex, and therefore more interesting, than would have been dreamed of only a few years ago.

It has also become clear that not only is this region interesting because of its novelty and complexity, but also because it has supplied (and most likely will continue to supply) us with important clues about the formation of the solar system. The discovery of dust disks around nearby stars by the IRAS satellite in 1983-84 suggests that Kuiper belts may be a natural consequence of star and planet formation, and thus provides us with a tool for studying solar system formation around other stars.

II. Basic Orbital Dynamics

Much of the story of the Kuiper belt to date involves the distribution of the orbits of its members. In this section, we present a brief overview of the important aspects of the orbits of small bodies in the solar system. [For a more detailed discussion see SOLAR SYSTEM DYNAMICS]

The most basic problem of orbital dynamics is the two-body problem: a planet, say, orbiting a star. In this case, the orbit of the planet is constrained to lie in a single plane. The orbit's trajectory is an ellipse with the Sun at one of the foci. Energy, angular momentum, and the orientation of the ellipse are conserved quantities. The *semi-major axis*, a, of the ellipse is a function of the orbital energy. The *eccentricity*, e, of the ellipse is a function of the energy and the angular momentum. For a particular semi-major axis, the angular momentum is a maximum for a circular orbit, $e = 0$. These two-body orbits are known as *Kepler orbits.*

In the real solar system there are nine planets and many smaller bodies, each acting to gravitationally perturb the orbits of the others. However, the Sun is much more massive than any of the planets. As a result, the orbits of objects in the solar system can be viewed as slightly perturbed Keplerian orbits about the Sun (unless the object in question gets particularly close to another perturber, such as a planet). A Keplerian orbit is characterized by its instantaneous semi-major axis and eccentricity, as well as by three angles which describe the orientation of the orbital ellipse in space. The first, known as the *inclination,* i, is the angle between the angular momentum vector of the orbit and some reference direction for the system. In our solar system, the reference direction is usually taken as the angular momentum vector of the Earth's orbit (which defines the *ecliptic plane*, the reference plane), but is sometimes taken to be the angular momentum vector of all the planetary orbits combined (which defines the *invariable plane*).

The point where the orbit passes through the reference plane in an 'upward' direction is called the *ascending node.* Here the reference plane is the plane that is perpendicular to the reference angular momentum vector and 'upward' is in the direction along the reference angular momentum vector. The second orientation angle of the orbit is the angle between the ascending node and some reference direction in the reference plane, as seen from the Sun. In our solar system, the reference direction is usually taken to be the direction toward the vernal equinox. This angle is known as the *longitude of the ascending node*, Ω.

The third and final orientation angle is the angle between the ascending node and the point where the orbit is closest to the Sun (known as *perihelion*), as seen from the Sun. It is called the *argument of perihelion*, ω. Another useful angle, known as the *longitude of perihelion*, $\tilde{\omega}$, is defined to be $\omega + \Omega$.

The first-order gravitational effect of the planets on one another is that each applies a torque on the other's orbit, as if the planets were replaced by rings of material smoothly distributed along their orbits. This torque causes both the longitude of perihelion, $\tilde{\omega}$, and the longitude of the ascending node, Ω to precess. For a given planet the precession of $\tilde{\omega}$ is typically dominated by one frequency. The same is true for Ω, although the dominant frequency is different. The periods associated with these frequencies range from 4.6×10^4 to 2×10^6 years in the outer planetary system. This is much longer than the orbital periods of the planets (which are all less than 250 years).

The orbit of a small object in the solar system, when it is not being strongly perturbed by a close encounter with a planet or is not located near a resonance (see below), is usually characterized by a slow oscillation in e and i and a circulation (i.e., continuous change) in $\tilde{\omega}$ and Ω. The variation in the eccentricity is coupled with the $\tilde{\omega}$ variation and the variation in the inclination is coupled with the Ω variation. Figure 1 shows this behavior for the first discovered Kuiper belt object, 1992 QB_1.

The behavior of objects that are in a resonance can be very dramatic. There are two types of resonances that are known to be important in the Kuiper belt. The most basic is known as a *mean motion* resonance. A mean motion resonance is a commensurability between the orbital period of two objects. That is, the ratio of the orbital periods of the two bodies in question is a ratio of two (usually small) integers. Perhaps the most well-known and important example of a mean motion resonance in the solar system is the one between Pluto and Neptune.

As noted above, one of the unique aspects of Pluto's orbit is that when Pluto is at perihelion, it is closer to the Sun than Neptune. Normally, this configuration would soon lead to close encounters between the two planets that would eventually scatter Pluto (i.e., perturb it to a very different orbit) and eventually, most likely out of the solar system (Because Pluto is much less massive than Neptune, conservation of energy and angular momentum would lead to Pluto's orbit undergoing radical changes, and not Neptune's). However, close encounters do not occur because Pluto is locked in a mean motion resonance where it goes around the Sun twice every time Neptune goes around three times. This resonance is known as the 2:3 mean motion resonance.

Figure 2a illustrates how the 2:3 resonance works. The trajectory of Pluto is shown in the figure in a frame that rotates with Neptune. The figure is constructed so that in a non-rotating frame, all the planets would move in a counter-clockwise direction. However, in the rotating frame shown, Neptune (N) and the Sun (S) are fixed; their locations are shown as dots. Pluto moves in a clockwise direction when further from the Sun than Neptune and moves in a counter-clockwise direction when closer to the Sun. In the figure, an object with Pluto's eccentricity and exactly at Neptune 2:3 mean motion resonance would have a trajectory that is a double-lobed structure oriented so that the loops (which are the object's perihelion passages) are always 90° from Neptune. The configuration shown in the figure will remain fixed only if the object is *exactly* at the location of the resonance. This is not the case for Pluto. In Pluto's case, the double-lobed structure will slowly precess in the rotating frame, possibly leading to close encounters with Neptune. What makes resonances like this one important is that there are stabilizing forces that lock nearby orbits into the resonances. In other words, the resonances are stable.

How is the resonance stabilized in Pluto's case? Suppose the semi-major axis of Pluto was slightly larger than the exact location of the resonance. Pluto would be moving slightly too slowly to maintain the configuration in Figure 2a. In other words, it would lag behind an object that was exactly at the resonance and Pluto's double-lobed trajectory would precess clockwise, as shown in Figure 2b. Because of the symmetry of the exact resonance (Figure 2a) the forces on Pluto by Neptune average out over long periods of time. This is not true for the new configuration. Indeed, Pluto would receive its largest acceleration (a_m) from Neptune when in or near the upper lobe in Figure 2b. At this point, Pluto is leading Neptune in their orbits and thus Pluto will be slowed down, decreasing its semi-major axis. Remember, this scenario started out with Pluto's semi-major axis being too large. As Pluto's semi-major axis decreases, its precession rate in the figure decreases, goes through zero (when its semi-major axis is equal to that of the resonance), and starts to move the other way (when its semi-major axis is less than that of the resonance).

Now Pluto's semi-major axis is smaller than the resonance and the double-lobed trajectory is precessing in the counter-clockwise direction. The trajectory moves through the point where the lobes are at 90° to the Sun-Neptune line and evolves into a configuration similar to the one shown in Figure 2c. In this case, Pluto gets its largest acceleration when it is near perihelion and is trailing Neptune in their orbits (near the lower lobe of the trajectory). Thus, Pluto's velocity is increased, increasing its semi-major axis. As Pluto's semi-major axis increases, its precession rate in the figure decreases, goes through zero (when its semi-major axis is equal to that of the resonance), and starts to move the other way (when its semi-major axis is greater than that of the resonance). The end result is that the double lobe structure oscillates or 'librates' back and forth in a frame that rotates with Neptune, as shown in Figure 2d. Therefore, this orbit is stable.

Mean motion resonances of small bodies are identified by the planet involved and the two integers that define that ratio of orbital periods. The two integers are separated by a colon, the one associated with the small body appears first. For example, Pluto orbits the Sun twice (2) every time Neptune orbits the Sun three (3) times. So, Pluto is said to be in Neptune's 2:3 mean motion resonance.

The other type of resonance that is important in the Kuiper belt is called a *secular resonance.* A secular resonance occurs when there is a commensurability between the frequencies associated with the precession rates of either $\tilde{\omega}$ or Ω of two objects. A small body in a secular resonance can undergo drastic changes in its eccentricity and inclination. Indeed, it has been shown that secular resonances can drive objects into the Sun (i.e. $e \rightarrow 1$).

To understand how such extreme behavior can come about, imagine a system with two planets and one small body. For simplicity, assume that the orbits of these three objects are in the same plane. The principal effect of the two planets on one another is to cause their orbits to precess. The orbit of the small body will also precess. Figure 3a shows the long-term trajectories of these objects in a fixed frame. Figure 3b shows the same system in a frame that rotates with the precession rate of the small body. Note that the orbit of the secular small body (the outermost orbit) is now an ellipse. The trajectories of the two planets are still, on average, axisymmetric and thus the small body experiences no long-term torques.

However, if the small body is in a secular resonance with one of the planets then it precesses at the same rate as the planet. For example, Figure 3c shows a system where the small body (outermost orbit) is in a secular resonance with the inner planet. The system is shown in a frame that rotates with the small body's precession rate. In this case, the long-term trajectory of the planet is a fixed ellipse and is no longer axisymmetric. Thus the small body feels a significant long-term torque, which can lead to a significant change in the eccentricity (which is related to the angular momentum) of the small body.

There are actually two types of secular resonances. The first, which was discussed above, is a resonance between the precession rates of the longitudes of perihelion. As discussed, this can lead to changes in eccentricity. These resonances are identified by the Greek letter ν with a numbered subscript that indicates the resonant planet (1 for Mercury through 9 for Pluto). In the Kuiper belt, the perihelion secular resonance with Neptune, or ν_8, is most important. The other type of secular resonance occurs when the small body's nodal precession rate is the same as for a planet. This type of resonance can cause significant changes in the inclination of the orbit of the small body. These resonances are identified by ν_{1x}, where x is the number of the resonant planet. For example, the nodal resonance with Neptune is the ν_{18}

The dynamical structure of the Kuiper belt and the scattered disk have been sculpted by a combination of mean motion and secular resonances and by the evolution of these resonances during the formation of Uranus and Neptune. In order to illustrate this point, we must first discuss what is known of the objects in the Kuiper belt and scattered disk. We review the observational data in the next section.

III. Observations of the Kuiper belt and Related Structures

The greatest areal search for trans-Neptunian objects is that by Clyde Tombaugh which covered the entire northern sky photographically to a visual magnitude (V, yellow-green in color) of 15. Tombaugh succeeded in discovering Pluto in 1930. Charles Kowal searched ~16% of the sky photographically, generally near the ecliptic, to a visual magnitude of ~20, and discovered the first outer solar system, planet-crossing object (other than Pluto and recognized comets), 2060

Chiron, in 1977. Chiron's orbit crosses Saturn's with a perihelion of 8.45 AU and an aphelion of 19.0 AU, just inside the orbit of Uranus. As of January 1998, six additional outer solar system, large, planet-crossing objects have been discovered. These objects are known as Centaurs.

Orbital data and estimated radii for the seven Centaurs are given in Table 1. The fact that these objects are on orbits that cross the orbit of one or more planets and the fact that they are not trapped in resonances implies that they will, in time, suffer a close encounter with one of these planets. Such encounters will lead to drastic changes in the orbits of these objects. Indeed, these objects will bounce around the solar system, gravitationally scattering off one planet after another until they finally reach their ultimate fate of *i*) being thrown out of the planetary system altogether, *ii*) impacting one of the planets, or *iii*) hitting the Sun. On average, it will take between $\sim 10^5$ and 10^8 years for these objects to reach their final states with a median lifetime of 5×10^7 years. The maximum inclination among the seven known Centaurs is 25° for 5145 Pholus, suggesting that their source reservoir is likely in the ecliptic plane. As will be discussed in more detail blow, the most likely source is the Kuiper belt.

The first successful detection of an object beyond the orbit of Neptune (other than Pluto and its satellite Charon) was by David Jewitt and Jane Luu. Using a CCD camera on the 2.2 meter University of Hawaii telescope, they searched ~ 1 square degree of sky to a visual magnitude of 24 and found object 1992 QB_1 in August, 1992. The discovery image is shown in Figure 4. At the time 1992 QB_1 was at a heliocentric distance of 41.2 AU. The object had an R (red) magnitude of 23.5, was reddish in color (see next section), and stellar in appearance with no evidence of cometary coma. The radius of this object can be estimated from the following formula:

$$R^2 = 4.53 \times 10^5 \, r^2 \Delta^2 p^{-1} 10^{-0.4\,V},$$

where R is the radius in kilometers, r is its heliocentric distance in AU, Δ is its distance from the Earth in AU, p is the fraction of incoming light that the object reflects (known as the geometric albedo), and V the object's visual magnitude. If 1992 QB_1 has a typical cometary albedo of 0.04 (the albedo of 1992 QB_1 and other Kuiper belt objects have not yet been measured), then it has a radius of $\sim$120 km. Subsequent observations allowed Brian Marsden to determine an orbit for 1992 QB_1 with semi-major axis of 43.8 AU, eccentricity of 0.088, inclination of 2.2°, and orbital period of 290 years. The perihelion distance of 40.0 AU is well beyond the orbit of Neptune. Long-term dynamical simulations suggest that orbits like that of 1992 QB_1 are stable over the age of the solar system.

The second trans-Neptunian object, designated 1993 FW, was discovered in March, 1993 at a heliocentric distance of 42.1 AU. 1993 FW is similar in size to 1992 QB_1 (possibly slightly larger), but less red in color, and again, stellar in appearance. A subsequent orbit solution found $a = 43.9$ AU, $e = 0.041$, $i = 7.7°$, and $P = 291$ years. Again, this orbit is expected to be stable over the age of the solar system. Thus, it appeared that the Kuiper belt had a dynamical structure similar to what most researchers expected at the time. That is, Kuiper belt objects were in random, low-inclination, low-eccentricity orbits about the Sun.

A real surprise came with the discovery of the next four objects. These objects were significantly different in that their heliocentric distances at discovery were substantially closer to the Sun, in a region where the orbits cannot be stable unless protected by some dynamical mechanism. The four objects: 1993 RO, 1993 RP, 1993 SB, and 1993 SC were found at heliocentric distances ranging from 32.3 to 35.4 AU. Interestingly, all four objects were approximately 60° from Neptune in the sky, suggesting a possible Trojan-type (1:1 mean motion resonance) dynamical relationship. However, further observations allowed Brian Marsden to show that all but 1993 RP (for which there are insufficient observations and which is now considered lost) are trapped in the 2:3 mean motion resonance with Neptune, similar to the orbit of Pluto.

Continued ground-based searches have now discovered a total of 60 trans-Neptunian objects, which are listed in Table 2, in order of discovery. The columns in the table are the semi-major axis, eccentricity, inclination, orbital period, and an estimated radius, based on an assumed cometary albedo of 0.04. Those objects marked with an asterisk have been discovered too recently to have accurate orbits determined. Those marked with a dagger do not have accurate orbits and are presumed lost. Many of these are assumed to be on circular orbits if they are beyond 40 AU, and on orbits locked in mean motion resonances inside of 40 AU.

The largest object so far discovered appears to be 1996 TO_{66} with a radius of $\sim$500 km. This is only slightly smaller than Pluto's satellite Charon, which has a radius of $\sim$600 km. For comparison, Pluto has radius of $\sim$1180 km. The next three largest trans-Neptunian objects are similar in size with radii of $\sim$300 km. The smallest known object is 1993 RP at $\sim$50 km.

In addition to large objects, approximately 30 objects have been detected in the Kuiper belt with V magnitudes between 28.6 and 27.8 (radii between 5 and 10 km assuming an albedo of 0.04) using the Hubble Space Telescope's Wide Field Planetary Camera 2. These objects were detected using statistical techniques and are not currently recoverable. Thus, only very limited information about them is available.

IV. Physical Observations of Trans-Neptunian Objects

Physical studies of Kuiper belt and Centaur objects have been limited because of the intrinsic faintness of these distant objects. Obtaining meaningful physical observations with good signal-to-noise is extremely difficult at magnitudes fainter than $V = 20$, and requires considerable dedication of telescope resources as well as very careful data reduction and calibration. Observations to date (January 1998) have been limited to broad-band spectral measurements of about one-third of the Centaur and Kuiper belt objects, infrared spectra of a few objects, and light curve measurements of a few objects. Coverage is better for the Centaur objects than for the Kuiper belt objects because the former are closer and hence, generally brighter.

Broad-band spectral observations in the visible region of the spectrum have demonstrated that both Centaurs and Kuiper belt objects apparently fall into two distinct groups characterized by their B–V (blue minus visible) and V–R (visible minus red) colors. One group consists of objects which are essentially grey, with colors similar to or only slightly redder than the sunlight which illuminates them. The other group consists of objects which display some of the reddest colors observed for any small bodies in the solar system.

This dichotomy in Centaur and Kuiper belt object colors is illustrated in Figure 5 which shows B–V and V–R colors for 13 objects, which include four Centaurs, eight Kuiper belt objects, and one scattered-disk object, 1996 TL_{66}. The grey objects in the figure are similar in color to C and D-type asteroids in the main asteroid belt (both C and D are carbonaceous asteroid types, see: ASTEROIDS) and include three Centaurs, two Kuiper belt objects, and one scattered-disk object. The red objects include one Centaur (5145 Pholus) and six Kuiper belt objects. For three other distant objects, one Centaur and two Kuiper belt objects, only V–R colors have been determined. Of these, both Kuiper belt objects appear to be members of the grey group, and the Centaur appears to be a member of the red group. None of them as yet appear to fall between the two populations. Also, the grey group includes 2060 Chiron, the first Centaur discovered, which is not among those shown in Figure 5.

Grey colors have generally been identified with the presence of carbon rich surface materials and are prevalent throughout the asteroid belt, though predominantly in the outer main belt. Red colors are more common in the outer planets region and have generally been taken to be indicative of the presence of complex hydrocarbons on the icy surfaces of the objects.

Primitive solar system materials in the outer planets region are expected to be initially grey or black, due to the high fraction of carbon and simple hydrocarbons mixed with relatively transparent ices. Laboratory experiments have shown that irradiation of such materials with UV photons from stars and cosmic rays will produce more complex hydrocarbons and turn the material reddish in color. However, continuing irradiation of the material will eventually produce elementary carbon and refractory organic residue that will once again turn grey.

Various hypotheses have been put forward to explain the different colors of the Centaur and Kuiper belt objects. Conceivably objects may be reddened by irradiation, but then resurfaced due to impacts, which excavate primitive grey material from the bottoms of craters and spreads it across the surface, "repainting" the objects grey. However, it is not clear that individual impacts on such large objects could excavate sufficient material to cover the entire surface area. Additionally, the fact that there appears to be two distinct color groups with little evidence of intermediate color objects, suggests that whatever might change the colors of Centaurs and Kuiper belt objects acts quickly compared to the length of time that they retain their colors, and acts over the complete surface. This might be consistent with a sudden resurfacing event due to a major impact, but is not consistent with the slow reddening of the new materials that would susequently take place.

Alternatively, since the first Centaur, 2060 Chiron, is known to display cometary-like outbursts and is grey in color, it has been suggested that the generation of a coma may spread grey material originating at depth over the surface of a formerly red object. Arguing against this idea is the fact that many of the Kuiper belt objects are far from the Sun in regions where surface temperatures would be too low to produce significant cometary activity, unless extremely volatile ices were involved. Also, the first two Centaurs found, 2060 Chiron and 5145 Pholus, have very similar perihelion distances, 8.45 and 8.68 AU respectively, yet one is grey and active, and the other is red and without apparent activity.

Of course, the dichotomy in colors may simply represent some intrinsic difference in the formation mechanism for the two groups of objects. However, the red and grey colors are found equally, to first order, for both Centaurs and Kuiper belt objects; statistics are not yet good enough to detect significant population differences. No correlations have been detected between colors and orbital elements (semi-major axis, eccentricity, inclination) or absolute magnitudes (i.e. physical radius).

The range of colors seen for the Centaurs and the Kuiper belt objects is also seen in long- and short-period comet nuclei. Comets do not display the same clear dichotomy in colors, but that may be a result of the contamination of the cometary observations by dust comae, as well as the poor signal-to-noise often present in cometary observations.

Broad-band spectral measurements have been extended into the infrared for about half a dozen Centaur and Kuiper belt objects. These all show essentially grey colors in the infrared, regardless of whether the objects appear grey or red in the visible region of the spectrum.

A moderate resolution spectrum in the near-infrared has been obtained for 1993 SC, one of the brightest of the Kuiper belt objects (Figure 6). The spectrum shows absorption features at 1.62, 1.79, 1.95, 2.20, and 2.32 microns (μ), and is similar in many ways to reflection spectra in the near-IR for both Triton and Pluto (which are also shown in the Figure). The features at 1.79, 2.20, and 2.32 microns have previously been interpreted as evidence for methane (CH_4) in solid solution with nitrogen (N_2) ice on Triton's surface. The identification of these bands with N_2 ice on 1993 SC is far less certain and they may represent pure methane ice alone. A small inflection in the spectrum of 1993 SC at 2.15 microns is also similar to one seen on Triton and is attributed to N_2 ice. However, the issue is not so clear for 1993 SC and the error-bars on the 1993 SC measurements are such that this feature may not be real. Other spectral features on 1993 SC are suggestive of simple hydrocarbon ices such as ethane (C_2H_6), ethylene (C_2H_4), and acetylene (C_2H_2).

If nitrogen ices are indeed present in sufficient quantity on the surfaces of the Kuiper belt objects, then that would have important implications for the surface temperatures, and hence the albedos of these objects. Up to now it has generally been assumed that the Kuiper belt objects have very low albedos, similar to that for carbonaceous asteroids and cometary nuclei. Such albedos imply that sub-solar temperatures on 1993 SC at its current heliocentric distance of 34.3 AU should be on the order of 65 K. At that temperature methane and other simple hydrocarbon ices should be stable, but N_2 should have a fairly high sublimation rate. Thus, one would not expect to see stable N_2 ice unless the albedo of 1993 SC was considerably higher than 0.04, which would then result in lower equilibrium surface temperatures and hence, lower sublimation rates. Note that both Triton and Pluto have fairly high mean albedos, $A = 0.85$ and 0.55, respectively. These make the presence of N_2 ice more reasonable on these bodies, though the surface temperature of Pluto is controversial and N_2 may be confined to the near-polar regions on that planet. Conceivably, any N_2 ice on 1993 SC, if it is there, may also be confined to the polar regions of that object. The few measured albedos for Centaur or Kuiper belt objects include 2060 Chiron, which has $A \approx 0.14$, Pholus with $A \approx 0.04$, Pluto with $A \approx 0.55$, and Charon with $A \approx 0.38$.

If there are indeed stable N_2 ices on the surfaces of Kuiper belt objects, then it may be that the albedos of the Kuiper belt objects are considerably higher than 0.04. This would mean that the objects are considerably smaller than current estimates. Estimated radii are proportional to $A^{-1/2}$. Obviously, this would have substantial implications for estimates of the total mass of material in the Kuiper belt, and for what the Kuiper belt is telling us about the origin of the planetary system. However, more observations are clearly needed in order to clarify whether or not N_2 ice is really there or not. If there is no N_2 ice, then there would be no problem with the currently assumed low albedos for the Kuiper belt objects.

The spectrum of 1993 SC strengthens the argument that there is a common link between Pluto, Triton, and the Kuiper belt objects. However, the similarities may only be skin deep, because these spectra only measure surface compositions and do not necessarily reflect the bulk composition of these objects. If the Kuiper belt objects did form under the same conditions as Triton and Pluto, then their bulk composition will be approximately two-thirds rock and complex organics, and about one-third volatile ices, the latter being predominantly water ice. The methane and nitrogen ices on Triton and Pluto, and perhaps 1993 SC, may only be thin veneers of surface frost.

Another recently obtained infrared spectrum shows evidence for water ice on the surface of the Centaur object 1997 CU_{26}. Water ice or frost on the surfaces of Centaurs and KBO's may also result in higher albedos, and hence may imply smaller diameters and masses for these bodies.

Several attempts have been made to obtain light curves for Centaurs and Kuiper belt objects, which would presumably be indicative of their rotation periods. Rotation periods have been determined for several Centaurs including 2060 Chiron, 5.92 hours, 5145 Pholus, 9.98 hours. and 1995 GO, 8.93 hours. Among Kuiper belt objects, attention has again focused on 1993 SC, one of the brightest Kuiper belt objects. However, no evidence has been found for a systematic variation in the brightness of 1993 SC, to a level of about 12%. This could suggest that the object is close to spherical with little in the way of albedo features on its surface, or that it is currently being viewed pole-on, and that the same hemisphere is being observed at all times. Further observations of 1993 SC and other Kuiper belt objects are required to resolve this question. The rotation periods of both Pluto and Charon are 6.387 days, but they have been mutually, tidally despun, so their original periods were likely shorter.

Physical observations of Centaur and Kuiper belt objects should improve as more of the new, large aperture telescopes come on line, and as the associated detectors and instrumentation improve over time. At present we have only the tantalizing clues described above as to the nature of these fascinating objects.

V. The Long-Term Stability of Orbits in the Kuiper Belt

One of the fundamental difficulties with studying astronomy is that the systems we study change on timescales that are very long compared to the length of time that astronomers have been observing them. The solar system is about 4.6 billion years old, the precession periods of the planets are thousands to millions of years, and yet modern science has only existed for a few hundred years. For example, we have only studied Pluto for only one-quarter of its orbital period! We have only seen it close to the perihelion of its fairly eccentric orbit.

And yet, one of the fundamental goals of astronomy is to determine where the solar system came from and where is it going. Perhaps the most important modern tool in accomplishing this goal is the computer. In the last few decades a new branch of astronomy has developed known as numerical experimentation, which consists of running experiments on the computer. The advantage of this method is that it is possible to run an experiment on the computer that would take much too long to perform in normal (as opposed to cyber) space.

Here, we are interested in the long-term dynamical behavior of objects in the Kuiper belt. We are interested in asking questions like: 'If there was an object in some specific orbit 4 billion years ago, would it still be there today?' or 'Will Pluto still be in a stable orbit 4 billion years from now?' Clearly, we cannot wait around for the answer. So, exacting computer algorithms have been developed that simulate the motion of objects along their orbits, taking into account the gravitational effects of the Sun, planets, and sometimes smaller objects. Using these algorithms, billions of years can pass in cyber space while only a few weeks or months pass in real space.

So let us consider the Kuiper belt. As briefly described in Section 1, modern work on the Kuiper belt was prompted by a mystery that existed in the early 1980's, concerning the origin of a certain type of comet, known as the Jupiter-family comets. These comets are short-lived as compared to the age of the solar system, with lifetimes of only $\sim 10^5$–10^6 years, Thus, there must be a 'reservoir' serving as a source for these comets, continually replenishing the transient population on planet-crossing orbits. Also, Jupiter-family comets have small orbital inclinations compared to most comets; the median inclination of the Jupiter family is $\sim 10°$. Dynamical simulations showed that the only possible way to get such a flat inclination distribution for these comets was for them to have originated in a flattened disk. At the time (about 1988), the only possible place to hide such a disk was beyond the orbit of Neptune. Thus, to many researchers in the field, the existence of the Jupiter-family comets was the first evidence for the existence of the Kuiper belt.

However, there was a major problem with the Kuiper belt as the source of the Jupiter-family comets. Recall that the dynamical lifetimes of these comets are short. So, if the Kuiper belt is the source of these comets, then whatever mechanism removes comets from the Kuiper belt and feeds them into the inner solar system needs to be active today. Initially, Julio Fernandez suggested that Mars-sized bodies in the Kuiper belt itself would scatter smaller bodies on to planet-crossing orbits. Later, Michael Torbett showed that such massive objects were not necessary because orbits in the Kuiper belt just outside the orbit of Neptune were formally *chaotic*.

Usually, if two objects are placed very close to one another with the same initial velocities, the relative separation of the two bodies will increase linearly with time. However, in certain cases, the two bodies will separate exponentially. This kind of behavior is one of the formal definitions of *chaos* and has two very important consequences. First, it leads to unpredictability. That is, it is impossible to predict the dynamical behavior of objects for long periods of time. Second, objects on chaotic orbits may undergo drastic and unexpected changes in their orbits [see: CHAOS].

The behavior of the eccentricity of an object in an initially nearly-circular (small e) orbit in the Kuiper belt is shown in Figure 7. The eccentricity remains small for a long period of time, in this case for over 200 million years. The orbit looks very well behaved for this length of time. Then, at about 2.5×10^7 years, the object evolves very quickly on to a high eccentricity orbit. The eccentricity is so large that the object can and does encounter Neptune, which gravitationally kicks it out of the Kuiper belt.

The most complete study of the long-term behavior of objects in the Kuiper belt has been performed by Martin Duncan, Harold Levison, and collaborators. They followed the orbital evolution of thousands of test particles for the age of the solar system, studying a wide range of initial semi-major axes, inclinations, and eccentricities. Each particle was followed until it suffered a close encounter with Neptune. Once comets can encounter Neptune, they will rapidly evolve ($\sim 10^7$ to $\sim 10^8$ years) into the inner planets region, or be ejected to the Oort cloud or to interstellar space. This issue is described in more detail in section 7.

The results of these simulations are shown in Figure 8. The colored strips indicate the length of time required for a particle to encounter Neptune as a function of its initial semi-major axis and eccentricity. Strips that are colored yellow represent objects that survive for the length of the simulation, 4×10^9 years, the approximate age of the solar system. As can be seen in the figure, the Kuiper belt can be expected to have a complex structure, although the general trends are readily explained. Objects with perihelion distances less than $\sim 35\,AU$ (shown as a red curve) are unstable, unless they are near, and presumably librating about, a mean motion resonance with Neptune. Indeed, the results in Figure 8 show that many of the Neptunian mean motion resonances (shown in blue) are stable for the age of the solar system. Objects with semi-major axes between ~ 40 and $42\,AU$ are unstable. This is presumably due to the presence of three overlapping secular resonances that occur in this region of the solar system: two with Neptune and one with Uranus.

Indeed, secular resonances appear to play a critical role in ejecting particles from the Kuiper belt. This can be seen in Figure 9, which shows some of the results of a one billion year integration. Here, the color strips show the length of time required for a particle to encounter Neptune as a function of its initial semi-major axis and inclination. These particles all had initial eccentricities of 0.01. Also shown are the locations of the Neptune longitude of perihelion secular resonance (in red) and the Neptune longitude of the ascending node secular resonance (in yellow). It is important to note that much of the clearing of the Kuiper belt occurs where these two resonances overlap. This includes the low inclination region between 40 and $42\,AU$. The Neptune mean motion resonances are also shown (in green).

As mentioned above, many of the Neptune mean motion resonances are stable at low inclinations. However, it can be seen in Figure 9 that these resonances are often unstable at high inclinations. This instability is again most likely due to secular resonances. Figure 9 shows that the ν_8 secular resonance converges on the 3:4 and 2:3 mean motion resonances at large inclinations. The numerical integrations show that the unstable orbits in these regions of phase space are chaotic and temporarily librate about both the local mean motion resonance and the nearby secular resonance, confirming that the resonance overlap is the cause of the instability.

It is interesting to compare the numerical results to the current best orbital elements of the known Kuiper belt objects. This comparison is made in Figure 10. The locations of all the Kuiper belt objects with well established orbits are shown as filled circles in the figure. Stable regions as indicated by the numerical integrations in Figures 8 and 9 are shown as grey areas. The main conclusion from this comparison is that the orbits of objects inside of $\sim$41 AU have sufficiently high eccentricities that they must be in Neptune mean motion resonances to be stable. The orbits of objects outside of this region tend to have lower eccentricities and are not in obvious mean motion resonances (although there does appear to be two objects in the 3:5 resonance). It is interesting to note that the transition from resonant to non-resonant orbits occurs near the location of the overlapping secular resonances at 40–42 AU.

Looking at Figure 10, it is surprising that there are, as yet, no known objects in non-resonant orbits with semi-major axes between 36 and 39 AU, despite the fact that the dynamical simulations indicate that most orbits in this region are stable provided their initial eccentricities are less than $\sim$0.05. In order to interpret this result, let us consider what the integrations are saying about what we should expect in the Kuiper belt. The integrations were performed by: *i*) initially taking the giant planets in their current orbits and with their current masses, and massless test particles in a smooth distribution of initial orbits, and *ii*) integrating the orbits of the planets and test particles under the gravitational effects of the planets. Thus, if the 36–39 AU region indeed proves to be unpopulated, then some mechanism other than the long-term gravitational effects of the planets in their current configuration is likely required to have cleared it. Three mechanisms that may have accomplished this come to mind:

- A hypothetical early outward migration of Neptune would cause its mean motion resonances to sweep through this region thereby sweeping most objects into the mean motion resonances. The migration of Neptune was first demonstrated by Julio Fernandez and Wing Ip, and capture mechanism was later proposed by Renu Malhotra to explain Pluto's current orbit. The orbital element distribution predicted by this model is shown in Figure 11a. Note that the predicted orbital element distribution matches that for real objects in the Kuiper belt fairly well. In general, this mechanism predicts that the mean motion resonances will be over-populated relative to a more uniform initial distribution. One interesting consequence of this model is that in order to get eccentricities of 0.32 in Neptune's 2:3 mean motion resonance requires that Neptune must have had an initial semi-major axis of 25 AU (compared to a current value of 30 AU). One possible problem with this model is that it predicts that we should see a significant number of objects in orbits in the 1:2 mean motion resonance. A significant fraction of these objects should have large enough orbital eccentricities to get close enough to the Sun to have been discovered. Unfortunately for this mechanism, none have yet been found.

- Harold Levison, Alan Stern, and Martin Duncan have shown that the location of the perihelion secular resonance with Neptune would have been much closer to the Sun if the Kuiper belt were massive in the past. Indeed, if the Kuiper belt initially contained about 10 Earth masses ($M_\oplus$) of material between 30 and 50 AU, then the ν_8 resonance would have been at 36 AU. This amount of mass is significantly larger than what is in the Kuiper belt today (see section 6 on the size distribution of objects), but it is not inconsistent with what may have been there at earlier times (see section 8 on the formation and evolution of the Kuiper belt). As the Kuiper belt eroded to its current mass through a combination of physical collisions and dynamical ejection, the ν_8 secular resonance swept outward exciting the eccentricities of the orbits of objects along the way. Only those objects in orbits between the initial and final positions of the secular resonance would be effected. In addition, only those objects that were not in orbits in mean motion resonances would have been removed from the Kuiper belt due to encounters with Neptune. Figure 11b shows the orbital element distribution predicted by this mechanism. It correctly predicts that objects in orbits inside of 41 AU will be locked in mean motion, while those outside will not be. However, it does not predict the large eccentricities seen for the orbits of objects in the 2:3 mean motion resonance.

- Some process may have pumped up the eccentricity and inclination of the orbits of objects in the region between 36 and 39 AU to values greater than $e \sim 0.05$ and/or $i \sim 10°$ where the dynamical lifetimes are short. One method for exciting random motions in a disk is by mutual gravitational encounters between objects in the disk. In this vein, it is interesting to note that the escape velocity of the largest known Kuiper belt object (assuming the sizes in section 3 and a density of 1 gm cm^{-3}) is approximately 400 m/s, which is about 10% of its heliocentric orbital velocity. Thus, if there initially were enough of these large objects for the Kuiper belt to be dynamically relaxed by mutual gravitational scattering, then the Kuiper belt objects would have typical orbital eccentricities of about 10%, just enough to depopulate the region between 36 and 39 AU, except in the resonances where high eccentricity orbits are stable. Unlike the previous hypothesis, this mechanism predicts that the mean motion resonances will not be over-populated relative to a more uniform initial distribution.

These three explanations are not exhaustive. Physical collisions and gas drag may also have played an important role. Nonetheless, these three models make very different predictions about the distribution of the orbits of objects in the Kuiper belt. Thus, it seems likely that further observations will help resolve whether any of the mechanisms described above have played an important role.

VI. The Sizes of Trans-Neptunian Objects and the Total Mass of the Kuiper Belt

As briefly described in section 1, the disk out of which the planetary system accreted was created as a result of the Sun shedding angular momentum as it formed. As the Sun condensed from a molecular cloud, it left behind a disk of material (mostly gas with a little bit of dust) that contained a small fraction of the total mass but most of the angular momentum of the system. It is believed that the initial solid objects in the proto-planetary disk were boulder-sized, on the order of meters to tens of meters in diameter. These objects formed larger objects through a process of accretion, whereby bodies collided and stuck to one another. The boulders accreted to form mountain-sized bodies, which in turn accreted to form asteroids and comets, which in turn accreted to form planets (or the cores of the giant planets which then accreted gas directly from the solar nebula). Understanding this processes is one of the main goals of astronomy today.

There are few clues in our planetary system about this process. We know that the planets formed and we know how big they are. Unfortunately, the planets have been so altered by internal and external processes that it is difficult to learn anything about their formation. Luckily, we also have the asteroid belt, the Kuiper belt, and the scattered disk. These structures contain the best clues to the planet formation process because they are regions where the process started, but for some reason, did not run to completion (i.e. a large planet). Thus, the size distribution of objects in these regions may show us how the processes progressed with time and (hopefully) what stopped them. Unfortunately, this type of information has been contaminated in the asteroid belt [see: ASTEROIDS] due to the collisional breakup of many large bodies. Thus, the Kuiper belt and the scattered disk are perhaps the best places to learn about the accretion process.

The best clue for understanding planet formation in the trans-Neptunian region is the size distribution of the objects found there. This distribution is usually given in the form of a power law of the form

$$n(R)\,dn \propto R^{-q}\,dR, \qquad\qquad q > 0$$

where n is the number of objects with radius, R, between R and $R + dR$. The slope of this distribution, q, contains important clues about the physical strengths, masses, and orbits of the objects involved in the accretion process.

For example, there are two extremes to the accretion process. If two large, strong objects collide at low velocities then the amount of kinetic energy in the collision is small compared to the amount of energy holding the objects together. In this case, the objects merge to form a larger object. If two small, weak objects collide at high velocities then the energy in the collision will overpower the gravitational and material binding energies. In this case the objects break apart, forming a large number of much smaller objects. In realistic models of the Kuiper belt with a range of sizes and velocities (where small objects usually move faster than large objects due to the equal partition of energy), we expect small objects to fragment and large objects to grow. This produces a size distribution with $q \sim 3$ at small sizes and a much steeper slope at large sizes where accretion is important.

By culling information from a variety of sources, including both observed Jupiter-family comets and Kuiper belt objects, it is possible to attempt to construct a preliminary estimate of the size distribution of icy planetesimals formed in the outer solar system. Much of what follows is crude, and the results are not intended to be a finished product. On the contrary, the results should be viewed as a forum for a comparison of the known theoretical and observational constraints on the sizes of observed long- and short-period comets and Kuiper belt objects.

In all, there are currently five major constraints on the size distribution of objects between 30 and 50 AU from the Sun (which is the dynamically active region of the Kuiper belt). They are:

1) As described above, the current interest in the Kuiper belt was prompted by the suggestion that Jupiter-family comets originated there. If so, it is possible to estimate that total number of comet-sized ($1 < R < 10$ km) objects in the Kuiper belt from the number of known Jupiter-family comets. The total number of Jupiter-family comets (both active and extinct) is expected to be

$$N_J = N_K \, p_e \, f_J L_J,$$

where N_K is the current number of comets in the Kuiper belt, p_e is the mean probability that any comet between 30 and 50 AU will escape the Kuiper belt per year, f_J is the fraction of those comets that become Jupiter-family comets once they leave the Kuiper belt, and L_J is the dynamical lifetime of a Jupiter-family comet. With the exception of N_K, all other variables have been estimated based on either observational or theoretical work. Thus, given the observed number of Jupiter-family comets, it is found that there are approximately 7×10^9 objects with radii between 1 and 10 km in the Kuiper belt between 30 and 50 AU from the Sun.

2) In 1994, Anita Cochran and collaborators used the Hubble Space telescope to detect Halley-sized objects in the Kuiper belt. They detected approximately 30 objects with V magnitudes between 28.6 and 27.8. Assuming an albedo of 0.04, this magnitude range corresponds to radii between 5 and 10 km. These observations covered only about 4 square arc minutes of sky. If one assumes a uniform inclination

distribution of objects between 0 and 12°, the HST observations imply that there are $\sim 2 \times 10^8$ comets in this size range in the Kuiper belt. This should be viewed as a lower limit to the number between 30 and 50 AU, because the objects detected were all closer than 40 AU.

3) It has been found recently that Pluto's moon Charon has a significant orbital eccentricity (near $e = 0.003$). This eccentricity is surprisingly large because the time it takes for tides to damp the eccentricity to zero, $\sim 9 \times 10^6$ years, is short compared to the age of the solar system [see: PLUTO]. It is possible to place crude estimates on the total number of large objects in the Kuiper belt because Kuiper belt objects passing between Pluto and Charon can excite the eccentricity of Charon's orbit. Under the assumption that this mechanism is the only one of importance, preliminary results suggest that there are between 3×10^6 and 3×10^8 Kuiper belt objects with radii between 20 and 330 km, between 30 and 50 AU. Since other unrecognized mechanisms may also be important, this estimate should be viewed as an upper limit.

4) The searches that discovered the Kuiper belt objects listed in Table 2 can also be used to estimate the total number of objects in the Kuiper belt. In order to accomplish this, a consistent set of observations is required. Such an analysis is possible on the 15 objects that were discovered in the Mauna Kea and Cerro-Tololo surveys, which covered 3.9 and 4.4 deg^2 of sky, respectively. The Mauna Kea survey found that there are $\sim$3.8 objects per square degree brighter than a limiting magnitude of $R \sim 24.2$. The Cerro-Tololo survey found $\sim$1.4 deg^{-2} brighter than $R \sim 23.2$. Combining these numbers, there must be $\sim 7 \times 10^4$ objects with $R > 50$ km and $\sim 2.6 \times 10^4$ objects with $R > 80$ km between 30 and 50 AU from the Sun. Since these searches were only along the ecliptic, and Kuiper belt objects have since been found in orbits inclined up to 30°, the estimates above are likely only lower limits.

5) Charles Kowal's search that discovered Chiron (the first Centaur) in 1977 covered 6,400 square degrees photographically to $V \sim 20$. This search should have found any object as large or larger than Pluto currently near the ecliptic and within 60 AU of the Sun, though Kowal's longitude coverage was not complete. Thus, there is probably only one such object within 50 AU, Pluto.

The constraints above are displayed in Figure 12, which plots the cumulative number of objects larger than radius R between 30 and 50 AU from the Sun, as a function of R. The constraints above are shown as filled circles in the plot. It should be noted that some of these numbers are very uncertain. For example, point 1 can be off by as much as an order of magnitude because of uncertainties in the physical lifetimes of comets and in the current orbital element distributions of objects in the Kuiper belt. In addition, the location of this point on the abscissa, which represents the radius of the smallest detectable Jupiter family comet in the inner solar system, is also unknown. Its value is thought to be near 1 km, but is uncertain by at least a factor of two.

However, the data do show that the size distribution is not a simple power law. The circles in Figure 12 show that $q = 3$ fits the constraints fairly well for objects with $R \lesssim 10\,\mathrm{km}$. At $R > 10\,\mathrm{km}$ the circles indicate a slope of about $q = 4.5$. Thus, the size distribution of objects in the Kuiper belt might best be described by a broken power law of the form

$$n(R)dn \propto \begin{cases} R^{-q_1}dR & \text{if } R \leq R_0 \\ R^{-q_2}dR & \text{if } R > R_0 \end{cases},$$

where R_0 is the radius where the power law changes slope. The solid black curve shows such a distribution with $R_0 = 10$ km, $q_1 = 3$, and $q_2 = 4.5$.

Interestingly, models of the accretion of objects in the outer solar system and of the size-distribution of long-period comets find similar broken power laws. In addition, an analysis of the size distribution of the Kuiper belt objects discovered to date by ground-based telescopes, those indicated by points 4 and 5 in Figure 12, found a slope of $q_2 = 4$, quite close to the value predicted in Figure 12. The broken power law shown in Figure 12 suggests that Kuiper belt objects with radii less than 10 km are collisionally evolved, but that larger objects are most likely the results of accretion.

Given the limited number of objects which have been discovered in the Kuiper belt to date, any conclusions about the size distribution in the belt must be regarded as speculative. Nevertheless, the Kuiper belt is clearly a rich area for future studies of the growth of planetesimals in the solar nebula. The discovery of additional objects, along with accurate photometric measurements from which their sizes can be inferred, will hopefully help to shed more light on this problem.

Finally, it is possible to integrate under the broken power law shown in Figure 12 in order to estimate the total mass in the Kuiper belt between 30 and 50 AU. Such an integration with limits between $R = 1$ km and 1200 km (the approximate radius of Pluto) and assuming a density of 1 gm cm^{-3}, shows that the total mass between 30 and 50 AU is 0.13 $M_\oplus$.

As with many scientific endeavors, the discovery of new information tends to raise more questions than it answers. Such is the case with the above mass estimate. Edgeworth's and Kuiper's original arguments for the existence of the Kuiper belt were based on the idea that it seemed unlikely that the disk of planetesimals that formed the planets would have abruptly ended at the current location of the outermost known planet. An extrapolation into the Kuiper belt (between 30 and 50 AU) of the current surface density of non-volatile material in the outer planets region predicts that there should originally have been about 30 $M_\oplus$ of material there. However, as stated above, our best estimate is over 200 times less than that figure!

Were Kuiper and Edgeworth wrong? Is there a sharp outer edge to the planetary system? The answer appears to be no to both questions. Over the last few years, evidence has mounted that supports the idea of a massive primordial Kuiper belt. Models of collisional processes have shown that the Kuiper belt is currently eroding away due to collisions. These models predict that the Kuiper belt will lose a significant fraction of its objects in the next few times 10^8 years. Since the Kuiper belt is currently losing mass due to collisional erosion, it clearly had to be more massive in the past.

These collisional evolution models have also shown that it is not possible for objects with radii greater than about 30 km to form in the current Kuiper belt, at least by two-body accretion, over the age of the solar system. The current surface density is too low to accrete bodies larger than this size. However, the models show that objects the size of 1992 QB_1 could have grown in a more massive Kuiper belt if the mean orbital eccentricities of the accreting objects were very small. A Kuiper belt of at least several Earth masses is required in order for 100 km sized objects to have formed. It is interesting to note that according to this model, the 100 km sized objects had to have formed before Uranus and Neptune grew to their current masses, because the presence of those planets would pump the eccentricities of the orbits of the initial objects too high for QB_1-sized objects to accrete.

The same applies even more strongly to the accretion of Pluto and Charon. For those two bodies to have grown to their current sizes in the trans-Neptunian region, there must have originally been a far more massive Kuiper belt. Otherwise, Pluto and perhaps even Charon would have had to have accreted in a denser region of the solar nebula (presumably closer to the Sun) and then been transported to their current orbit by gravitational scattering. There is no known mechanism that could accomplish this.

The idea of a historically massive Kuiper belt also solves two other mysteries. First, as discussed in section 4, objects in the Kuiper belt seem to fall into two distinct classes. Inside of 41 AU, they are all in mean motion resonances with Neptune and have eccentricities larger than 0.1. Beyond 41 AU they tend to be in non-resonant orbits and most have eccentricities less than 0.1. If the Kuiper belt were massive in the past, the ν_8 secular resonance would have been closer to the Sun than it is today. As the Kuiper belt's mass declined, the ν_8 resonance would have swept outward, thereby exciting the orbital eccentricity of objects that were between its initial and its current position. The orbits of these objects would be unstable unless they were in a mean motion resonance. This mechanism thus predicts that most objects inside of the current ν_8 resonance should be on high eccentricity orbits in Neptune mean motion resonances, while objects beyond this region should be in low eccentricity, non-resonant orbits. This is precisely what is observed. Thus, an early massive Kuiper belt could help explain the current orbital element distributions of the known Kuiper belt objects.

The second potentially solved mystery has to do with Neptune's eccentricity. The orbital eccentricity of the other three giant planets range from 0.047 to 0.054. Neptune's orbital eccentricity is 0.009, much smaller. Recent analytic models by William Ward and Joseph Hahn have shown that Neptune's eccentricity could have been decreased through gravitational interactions with a massive Kuiper belt.

With the arguments presented above in mind, it is possible to build a strawman model of the Kuiper belt, which is depicted in Figure 13. There are three distinct zones in the Kuiper belt in this model. Region A is a zone where the dynamical perturbations of the outer planets have played an important role in determining the Kuiper belt's heliocentric structure. In this region the planets have tended to pump up the eccentricities of the orbits of objects. About half of the objects in this region have been dynamically removed from the Kuiper belt by being thrown into Neptune-crossing orbits. The remaining objects have been perturbed to larger orbital eccentricities, resulting in encounter velocities between objects that are large enough that accretion has stopped and collisional erosion has become important. Thus, we expect that a significant fraction of the mass in region A has been removed by collisions. Indeed, the solid curve in this region of Figure 13 shows an estimate of the current Kuiper belt surface density, while the dotted curve shows an estimate of the initial surface density, extrapolated from that of the outer planets. If these curves are at all representative of reality, the current surface density in this region has been depleted by a factor of 10^3. The outer boundary of region A is somewhere between 50 and 65 AU.

Region B in Figure 13 is a zone where collisions are important in shaping the structure of the Kuiper belt, but where the gravitational effects of the planets are not important. Without the effects of the planets, the orbital eccentricities of objects in this region will most likely be small enough that collisions lead to accretion of bodies, rather than erosion. Large objects could have formed here, so that the size distribution of objects may be significantly different than in the region immediately interior to it. However, the surface density may not have changed very much in this region over the age of the solar system. The outer boundary of Region B is very uncertain, but will likely be beyond 100 AU. Region C is a zone where collision rates are low enough that the surface density of the Kuiper belt and the size distribution of objects in it have remained virtually unchanged over the history of the solar system.

It is only natural to draw parallels between the Kuiper belt and the dust disks discovered around main sequence stars by the IRAS satellite. These disks extend out to $\sim$900 AU (in the case of β Pictoris) or more from the central star. The most likely source of the dust in the IRAS disks is colliding and sublimating comets. Infrared excesses have been found to be common around solar-type stars and it would be unusual if the Sun did not have a similar dust disk.

The amount of dust in the Kuiper belt between 30 and 100 AU is constrained to be less than 0.3 $M_\oplus$, based on estimates of collisionally produced dust in the belt and upper limits on zodiacal dust infrared emission by the COBE satellite. This estimate is 10^{-2} and 10^{-4} times the luminosity of the two best extra-solar dust disks, Vega and β Pictoris, respectively. Both the Vega and β Pictoris disks are much younger than our solar system, since they are in orbit around stars that are more massive than the Sun and thus have shorter lifetimes. Thus, they have not had as much time to erode as the Kuiper belt about our solar system. Still, a more massive Kuiper belt, as discussed above, could be present at larger solar distances.

VII. Ecliptic Comets and the Scattered Disk

As described in section 1, the current renaissance in Kuiper belt research was prompted by the suggestion that the Jupiter-family comets originated there. Thus, as part of the research intended to understand the origin of these comets, a significant amount of research has gone into understanding the dynamical behavior of objects that are on orbits that can encounter Neptune. These studies show that once objects (either escapees from the Kuiper belt or leftover planetesimals from the Uranus-Neptune zone) encounter Neptune, gravitational encounters can spread them throughout the planetary system. The distribution of these objects as predicted by numerical integrations by Martin Duncan and Harold Levison is shown in Figure 14.

Figure 14 therefore shows the predicted distribution of the objects that make up the ecliptic comet distribution described above. Those that get close to the Sun are the Jupiter-family comets. It is somewhat surprising that about a third of the objects in the simulations spend at least some of their time as Jupiter-family comets. The Jupiter-family comets that we see today are all small, $R \lesssim 10$ km. However, if our understanding of the size-distribution of these objects is correct (see section 5), we should expect to see a 100 km-sized Jupiter-family comet about 0.4% of the time. What a show that would be!

Those ecliptic comets between Jupiter and Neptune are the Centaurs (only the largest of which are observable). The simulations predict that there are $\sim 10^6$ ecliptic comets currently in orbits between the giant planets.

Most ecliptic comets do not last forever. They either get scattered out of the planetary system by a close encounter with a planet, or hit something (a planet or the Sun), in a relatively short amount of time: a few tens of millions of years. However, Levison and Duncan's integrations showed that a small fraction, about 1%, are scattered into long-lived, eccentric orbits outside the orbit of Neptune, though mainly inside of one to two hundred AU from the Sun. The long-range gravitational effects of the planets, in particular resonances with the planets, can lift the orbits of the objects beyond the outer edge of the planetary system (i.e. Neptune's orbit), protecting them from close encounters with the planets.

An example of this type of behavior is shown in by the dynamical simulation in Figure 15. This hypothetical object initially underwent a random walk in semi-major axis due to encounters with Neptune. At about 7×10^7 years it was temporarily trapped in Neptune's 3:13 mean motion resonance for about 5×10^7 years. It then performed another random walk in semi-major axis until about 3×10^8 years, when it was trapped in the 4:7 mean motion resonance, where it remained for 3.4×10^9 years. Notice the increase in the perihelion distance near the time of capture into this resonance. While trapped, the particle's eccentricity became as small as 0.04. After leaving the 4:7 resonance, the object was trapped temporarily in Neptune's 3:5 mean motion resonance for $\sim 5 \times 10^8$ yr and then went through a random walk in semi-major axis for the remainder of the simulation.

Many of the objects that had their perihelion distances raised by resonances can survive in these eccentric orbits for the 4.6 billion-year age of the solar system.

Thus, Duncan & Levison's simulations predicted the existence of a scattered disk in the outer solar system.

The simulations described above indicate that objects in the scattered disk will tend to have more eccentric and/or inclined orbits than those in the Kuiper belt. Two recently discovered trans-Neptunian objects may be the first members of the scattered disk to be found. The first, 1996 TL_{66}, was discovered in October, 1996 by Jane Luu and colleagues and is estimated to have a semi-major axis of 84 AU, a perihelion of 35 AU, an eccentricity of 0.58, and an inclination of 24°. Such a high eccentricity orbit could only result from gravitational scattering by a giant planet, in this case Neptune.

A second possible scattered disk object, 1996 RQ_{20}, was discovered in September, 1996 by Eleanor Helin and colleagues. Recent observations indicate its semi-major axis is 44 AU, its eccentricity is 0.11, and its inclination is 32°. The high inclination of this object suggests that it may be a member of the scattered disk.

VIII. The Formation of the Kuiper Belt and the Scattered Disk

Given the recent advances in both our observational and theoretical understanding of the trans-Neptunian region, it is now possible to construct a scenario of how the Kuiper belt and the scattered disk formed. The following scenario is by no means unique and there are variations that could reproduce the observed structures just as well. In addition, many of the stages of this scenario have yet to be modeled in detail and thus are still very uncertain.

In the outer solar system, Jupiter and Saturn formed early-on when the gaseous component of the solar nebula, mainly hydrogen and helium, was still present. As the gas in the solar nebula dissipated, there was a large population of small, icy bodies in low-inclination, nearly-circular orbits about the Sun outside the orbit of Saturn and extending outward perhaps as far as several hundred astronomical units. These objects gradually grew into larger planetesimals by mutual collisions. As the planetesimals increased in mass, they began to gravitationally perturb one another into more eccentric and more inclined orbits about the Sun. The growth of planets was more rapid inside of $\sim$40 AU, leading to the formation of Uranus and Neptune. Outside of 40 AU (i.e., in the Kuiper belt) the bodies remained relatively small.

As the bodies which eventually became Uranus and Neptune grew in mass, they began to gravitationally scatter the remaining nearby planetesimals both outward into the trans-Neptunian region and inward toward Jupiter and Saturn. Usually, the objects scattered outward returned to the Uranus-Neptune region where they would again be scattered. However, Jupiter and Saturn were so massive that they could efficiently eject many of these planetesimals out of the planetary system, thereby acting as a sink for these objects. As a result, there was a net flux of objects toward Jupiter and Saturn. This inward transport of mass required an outward migration of the orbits of Uranus and Neptune in order to conserve the total angular momentum and energy of the system. As Neptune moved outward, it may have trapped planetesimals external to it (i.e., in the Kuiper belt) into mean motion resonances. Once the planetesimals were trapped in the resonances, Neptune's heliocentric migration (and other dynamical processes) would have pumped up the eccentricity of the orbits of these objects to values as large as 0.3, leading to much of the present Kuiper belt structure described in section 5.

As discussed above, the objects scattered outward by Uranus and Neptune can be relatively long-lived. After their first encounter with Neptune, about 1% of the objects survive for the age of the solar system, stored in an extended scattered disk beyond Neptune. How much mass could there be in the scattered disk? There are $\sim$32 Earth-masses of material in Uranus and Neptune. Since planet formation is not 100% efficient, it is not unreasonable to expect that a similar amount of mass was initially scattered outward by Neptune (the efficiency factor is not known to within a factor of a couple of orders of magnitude). So crudely, we may expect as much as a few tenths of an Earth-mass of material in the scattered disk. This value is consistent with current estimates of the number of scattered disk objects in the solar system.

IX. Concluding Remarks

Since the study of the trans-Neptunian region is still in its infancy, much of the above formation scenario is speculative, and much of it is likely to change rapidly in the coming years (or even months) as this field matures and more data become available. One reason for this is that only a very small fraction ($\lesssim 0.06\%$) of the Kuiper belt and scattered disk objects with radii > 50 km have been discovered so far. However, it has already been determined that the structure of the trans-Neptunian region is much more complex, and therefore more interesting, than would have been dreamed of only a few years ago. It has also become clear that not only is this region interesting because of its novelty and complexity, but also because it has supplied (and most likely will continue to supply) us with important clues to the formation of the solar system.

Pluto and its satellite Charon occupy an important position in this new scenario. They are no longer just a pair of odd objects on the fringes of the planetary system. We now recognize that they are the largest known members of a huge population of objects which grew in the outer reaches of the solar nebula beyond Neptune. Other Pluto-sized objects may still orbit in the more distant regions of the Kuiper belt.

The Kuiper belt also provides us with a link to dust disks detected around many main sequence stars. It suggests that those stars likely had accretion disks and may have even formed planetary systems like our own closer in to the stars.

Study of the Kuiper belt is one of the most active frontiers of solar system research at present, and it will be interesting to see what new discoveries and revelations are made in the coming years.

Figure Captions

Figure 1 — The temporal evolution of the orbit of the first Kuiper belt object found, 1992 QB_1. As described in the text, the eccentricity, e, and inclination, i, oscillate, while the longitude of the ascending node, Ω, and the longitude of perihelion $\tilde{\omega} = \omega + \Omega$ circulate.

Figure 2 — The mechanics of the 2:3 mean motion resonance between Neptune and Pluto. The double-lobed curve represents the orbit of Pluto as seen in a coordinate frame that rotates at the average speed of Neptune. Thus, Neptune is almost stationary in this figure. As drawn, the frame is rotating counterclockwise. The location of the Sun and Neptune are shown as dots and are labeled s and N respectively. A) The orbit of an object exactly at the resonance. The gravitational perturbations of Neptune cancel out due to the symmetry in the geometry. Thus, this orbit is stable. B) If the symmetry is broken then there is a net acceleration due to Neptune. Here, the strongest perturbation (a_m) is at the upper lobe. Pluto is leading Neptune here, so the net acceleration will decrease Pluto's semi-major axis and it will start to precess clockwise. C) The strongest perturbation is in the lower lobe. Pluto will start to precess counterclockwise. D) The orbit of an object that *librates* in the resonance.

Figure 3 — The mechanics of a secular resonance. Three orbits are shown in each panel. The inner two are planets, which are shown as solid lines. The outer orbit (dashed line) is for an object that is very small compared to either of the planets. The orbits of each object are ellipses, and the ellipses are precessing due to the mutual gravitational effects of the planets. A) The orbits of the objects over a period of time that is long compared to the precession time of the orbits. Here, we are looking in a fixed, non-rotating reference frame.

Each orbit sweeps out a torus of possible positions. B) The same as A, except that we are looking in a frame that rotates at the precession rate of the small outer body. Thus, its orbit is again an ellipse. This panel shows the geometry if no secular resonance exists. Note that the trajectories of the planets look axisymmetric. Therefore, there is no net torque on the outer small object. C) Same as B except that the outer object is in a secular resonance with the inner planet, i.e. both orbits precess at the same rate. As a result, the outer object no longer sees an axisymmetric gravitational perturbation from the inner planet. Indeed, it feels a significant torque.

Figure 4 — Discovery images of 1992 QB_1 with the University of Hawaii 2.2 m telescope on Mauna Kea taken by David Jewitt and Jane Luu. The Kuiper belt object is about R magnitude 23.5. The numbers at the upper right of each image refers to the time of day. The arrow indicates 1992 QB_1, which slowly moves through the field of view.

Figure 5 — The broad-band colors of Kuiper belt objects and Centaurs as determined by Stephen Tegler and William Romanishin. The squares and diamonds refer to Kuiper belt objects and Centaurs, respectively. B, V, and R refer to magnitudes measured through blue, green-yellow ("visual"), and red filters, respectively. The error-bars represent 1σ uncertainties. The asterisk represents the color of the Sun. Redder colors are to the upper right of the diagram. Note the two distinct populations.

Figure 6 — The dots and error-bars represent the spectrum of 1993 SC as determined by Robert Brown, Dale Cruikshank, Yvonne Pendleton, & Glenn Veeder. The solid and dotted curves represent the spectra of Neptune's satellite Triton and Pluto, respectively. The absorption features in the spectra have been identified with methane and possibly nitrogen ices.

Figure 7 — The eccentricity of a 'typical' object on a chaotic orbit in the Kuiper belt as a function of time. At first, the object appears to be well behaved as the eccentricity slowly oscillates up and down. Then the behavior quickly changes and the eccentricity gets large. So large, in fact, that the object then suffers a close encounter with Neptune, which gravitationally scatters the object out of the Kuiper belt. It is important to note that this sudden increase in eccentricity is not a result of some large perturbation that the object experienced, such as an encounter with a massive object. It is a result of the slow build up of small perturbations of the planets.

Figure 8 — The dynamical lifetime for small particles in the Kuiper belt derived from 4 billion year integrations. Each particle is represented by a narrow vertical strip of color, the center of which is located at the particle's initial eccentricity and semi-major axis (initial orbital inclination for all objects was 1 degree). The color of each strip represents the dynamical lifetime of the particle. Strips colored yellow represent objects that survive for the length of the integration, 4×10^9 years. Dark regions are particularly unstable on these timescales. For reference, the locations of the important Neptune mean motion resonances are shown in blue and two curves of constant perihelion distance, q, are shown in red.

Figure 9 — The dynamical lifetime for test particles with initial eccentricity of 0.01 derived from one billion year integrations. This plot is similar to Figure 8 except that a different color table was used for the solid bars. In addition, the red and yellow curves show the locations of Neptune longitude of perihelion secular resonances (ν_8) and the Neptune longitude of the ascending node secular resonances (ν_{18}), respectively. The green lines show the location of the important Neptune mean motion resonances.

Figure 10 — The dots show the distribution of semi-major axis (a) and eccentricity (e) of the known Kuiper belt objects which have been observed at more than one opposition. The dashed lines show the locations of Neptune's important mean motion resonances. Notice that inside of the ν_8 secular resonance, which is marked in the figure, all objects are in mean motion resonances. The shaded regions show the locations of stable non-resonant orbits (see Figure 8). Objects that lie outside these regions and that are not in a Neptune's mean motion resonance are not stable for the age of the solar system. The dotted curve is a curve of constant perihelion distance, $q = 35$ AU.

Figure 11 — Same as Figure 10 except the black dots show the distribution of semi-major axis (a) and eccentricity (e) predicted by two models of Kuiper belt formation. The open circles show the orbits of the known Kuiper belt objects which have been observed at more than one opposition. (These objects are the black dots in Figure 10.) A) The distribution predicted by the migrating planet model. B) The distribution predicted by the massive Kuiper belt model.

Figure 12 — The cumulative number of objects in the Kuiper belt between 30 and 50 AU, larger than radius R as a function of R. The filled circles indicate the known constraints as listed in the text, and arrows indicate upper/lower limits.

Figure 13 — A strawman model of the Kuiper belt. The dark area at left marked "Planets" shows the distribution of solid material in the outer planets region. Notice that it follows a power law with a slope of about -1.5 versus heliocentric distance. The dashed curve is an extension into the Kuiper belt of the power law found for the outer planets and illustrates the likely initial surface density distribution of solid material in the solar nebula. The dashed curve has been scaled so that it is twice the surface density of the outer planets,

under the assumption that planet formation was 50% efficient. The Kuiper belt is divided into three regions. In Region A, the Kuiper belt has been shaped by collisions and by dynamical perturbations by the planets. Most of the material in Region A has been eroded away (both dynamically and physically) because of the large orbital eccentricities excited by the planetary perturbations. The solid curve shows a model of the mass distribution in Region A by Martin Duncan, Harold Levison, and collaborators. In Region B collisions have been important, but gravitational perturbations by the planets have not. Orbital eccentricities were small enough that accretion was the dominant process in Region B. In Region C, neither the gravitational effects of the planets nor collisions have been important. The Kuiper belt has remained virtually unchanged in this region over the age of the solar system. The dotted curves illustrate the unknown shape of the surface density distribution in Region B.

Figure 14 — The surface number density (on the plane of the ecliptic) of ecliptic comets as determined from numerical integrations by Martin Duncan and Harold Levison. There are approximately 10^6 comets in this population.

Figure 15 — The temporal behavior of the orbit of a long-lived member of the scattered disk. The black curve shows the behavior of the comet's semi-major axis. The gray curve shows the perihelion distance. The three dotted curves show the location of the 3:13, 4:7, and 3:5 mean motion resonances with Neptune. This object initially underwent a random walk in semi-major axis due to encounters with Neptune. At about 7×10^7 years it was temporarily trapped in Neptune's 3:13 mean motion resonance for about 5×10^7 years. It then performed a random walk in semi-major axis until about 3×10^8 years, when it was trapped in the 4:7 mean motion resonance, where it remained for

3.4×10^9 years. Notice the increase in the perihelion distance near the time of capture into the resonance. While trapped in this resonance, the particle's orbital eccentricity became as small as 0.04. After leaving the 4:7 resonance, it was trapped temporarily in Neptune's 3:5 mean motion resonance for $\sim 5 \times 10^8$ years and then went through another random walk in semi-major axis for the remainder of the simulation.

FIGURE 1

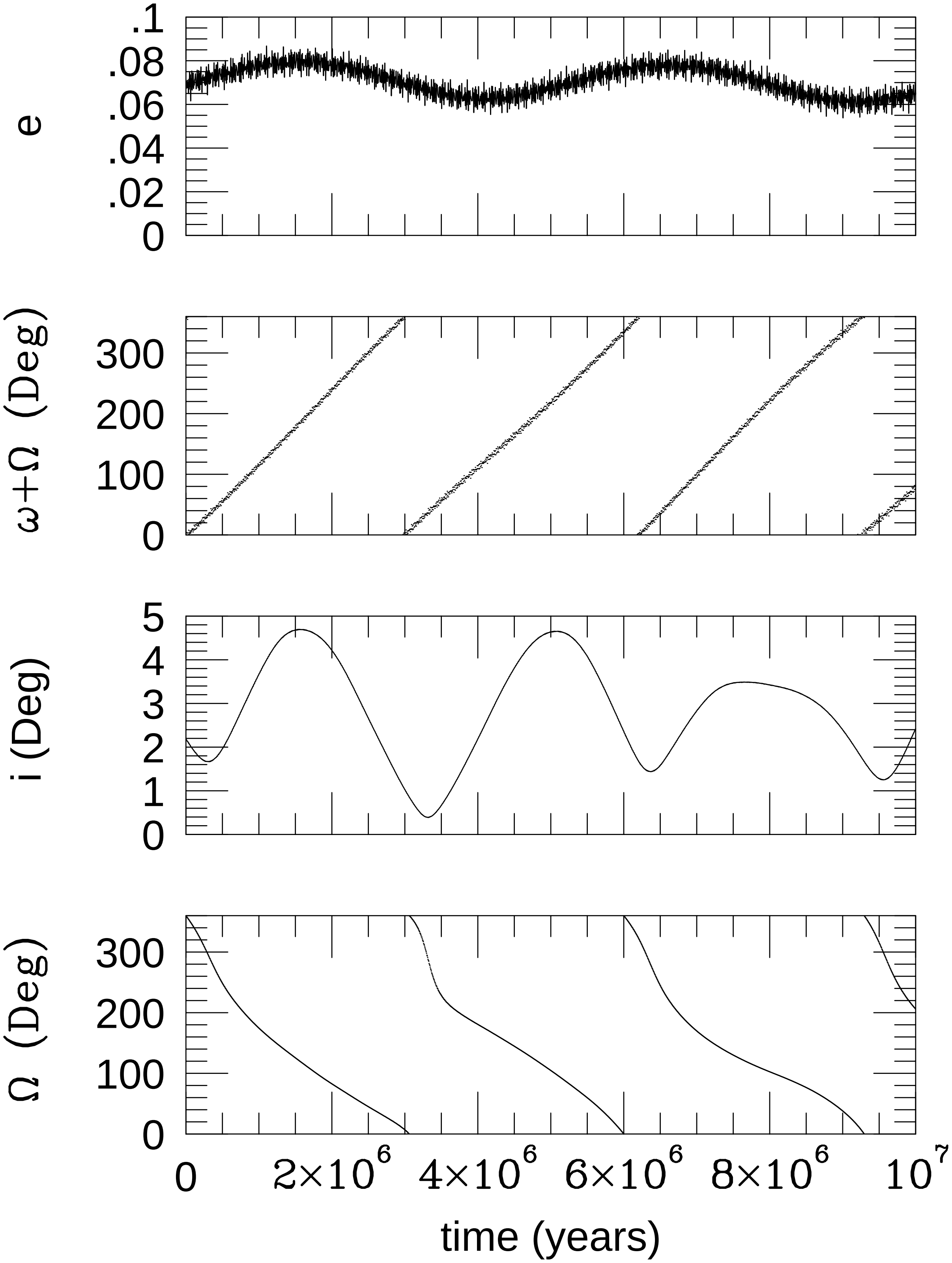

e
.1
.08
.06
.04
.02
0
ω+Ω (Deg)
300
200
100
0
i (Deg)
5
4
3
2
1
0
Ω (Deg)
300
200
100
0
0
2×10^6
4×10^6
6×10^6
8×10^6
10^7
time (years)

FIGURE 2

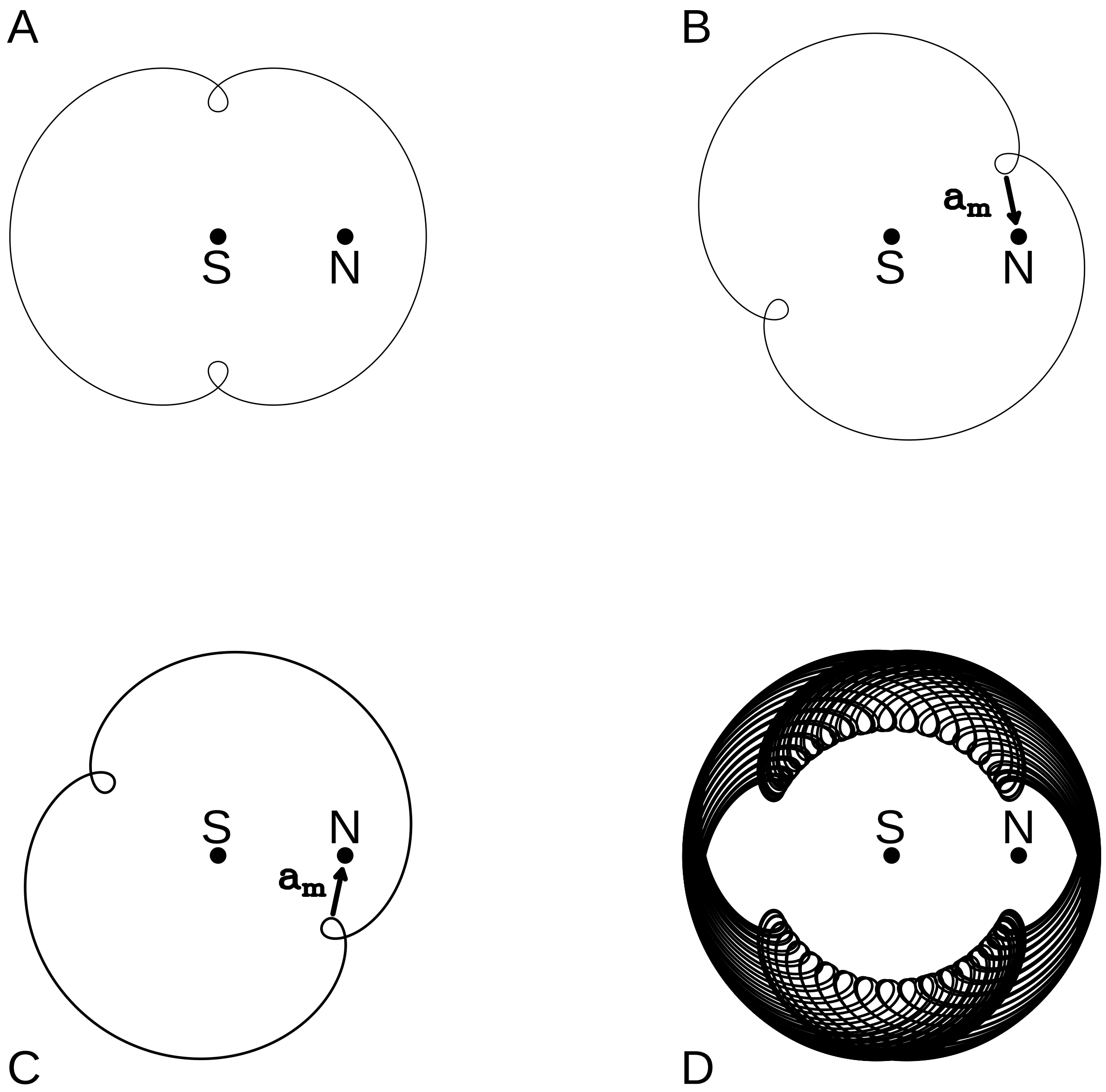
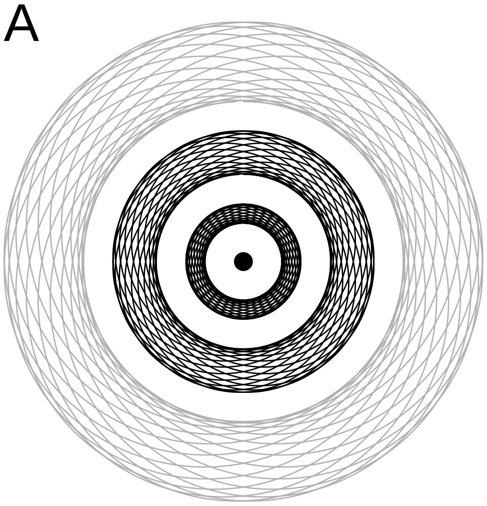


FIGURE 3

A



B

C

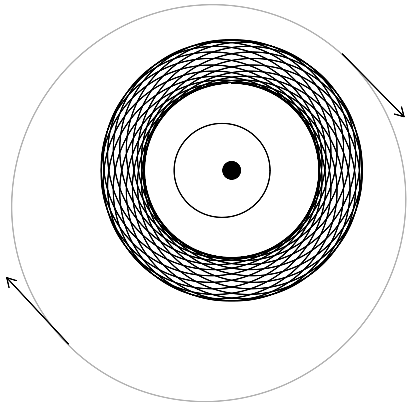
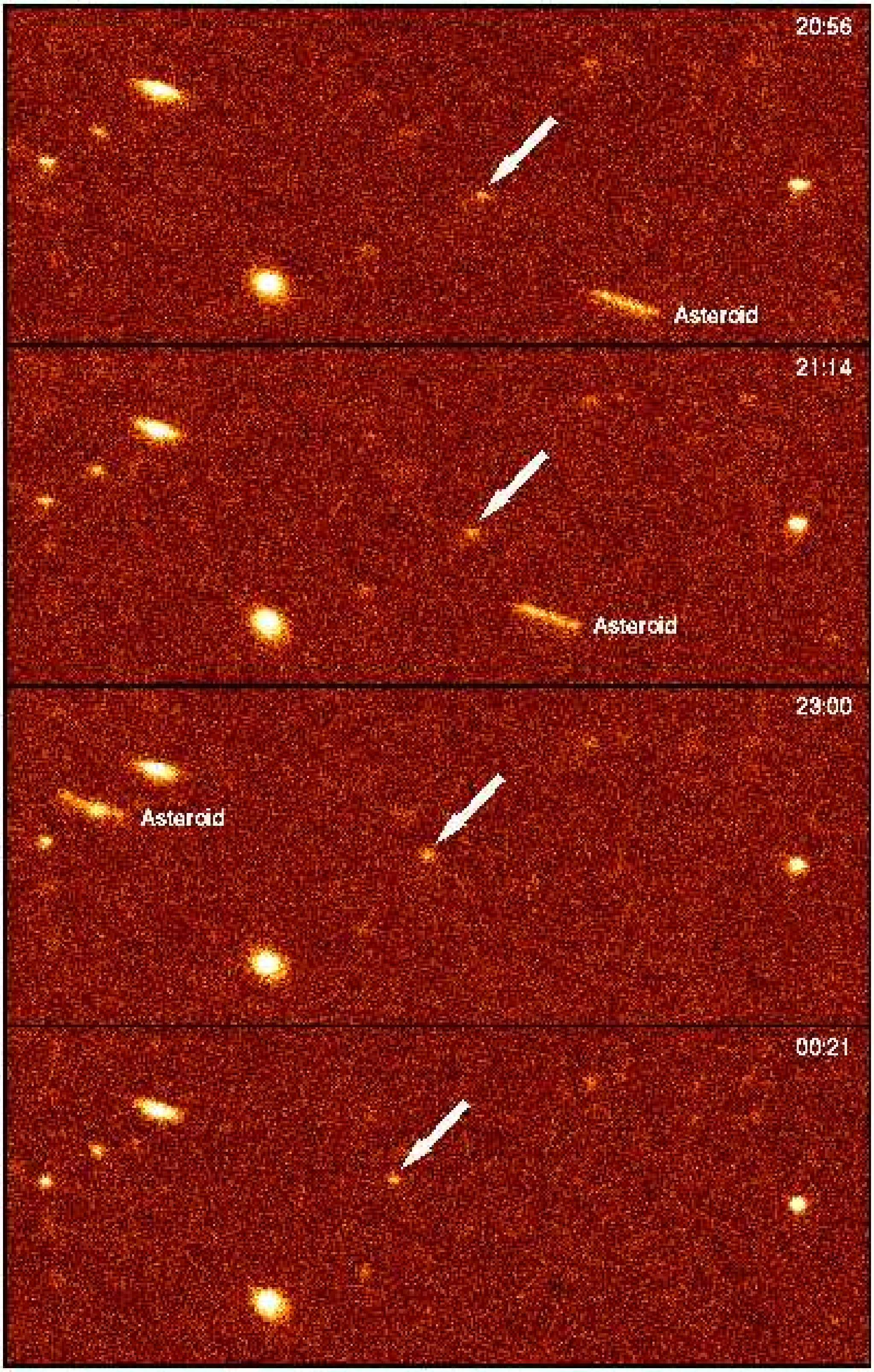


FIGURE 4



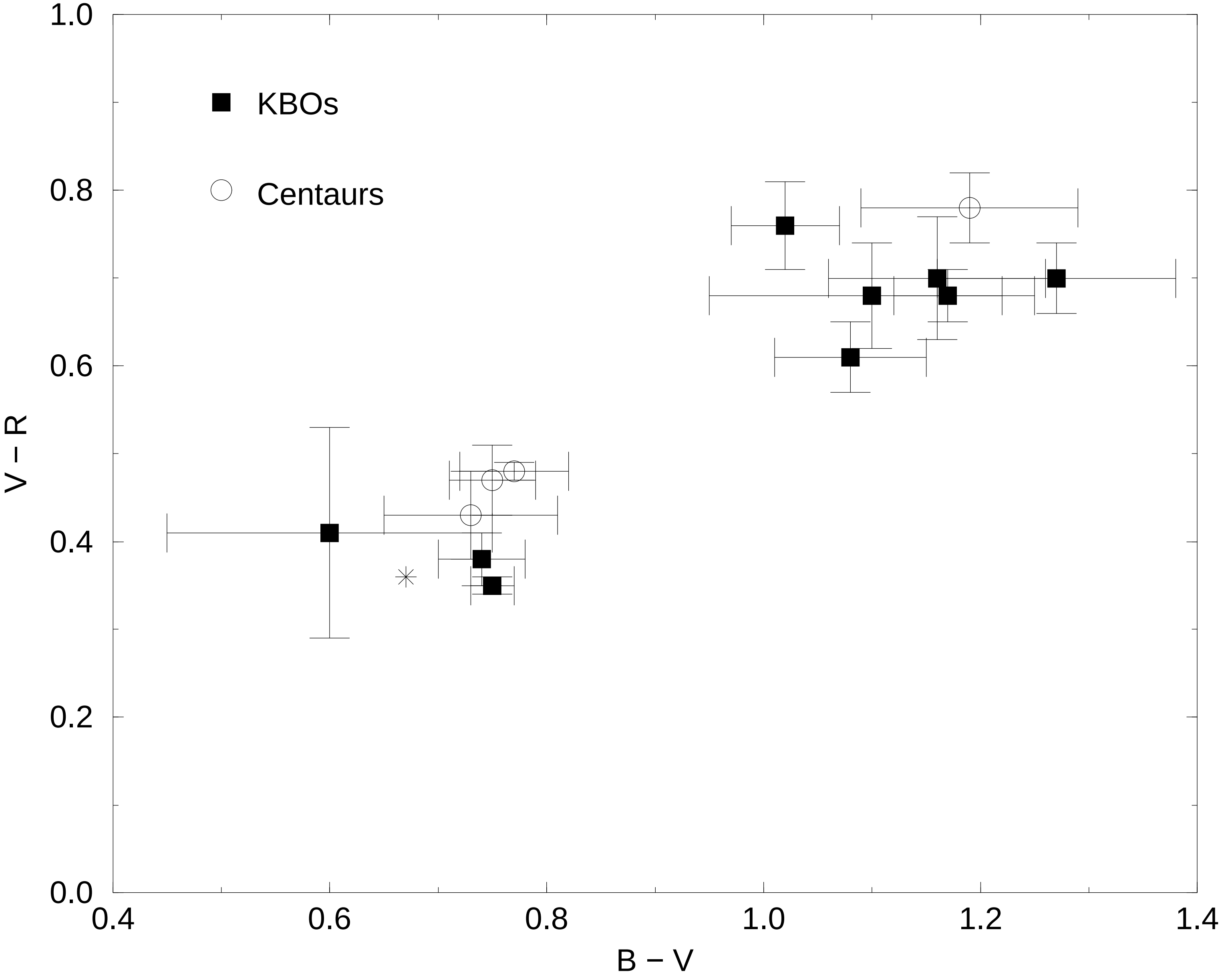
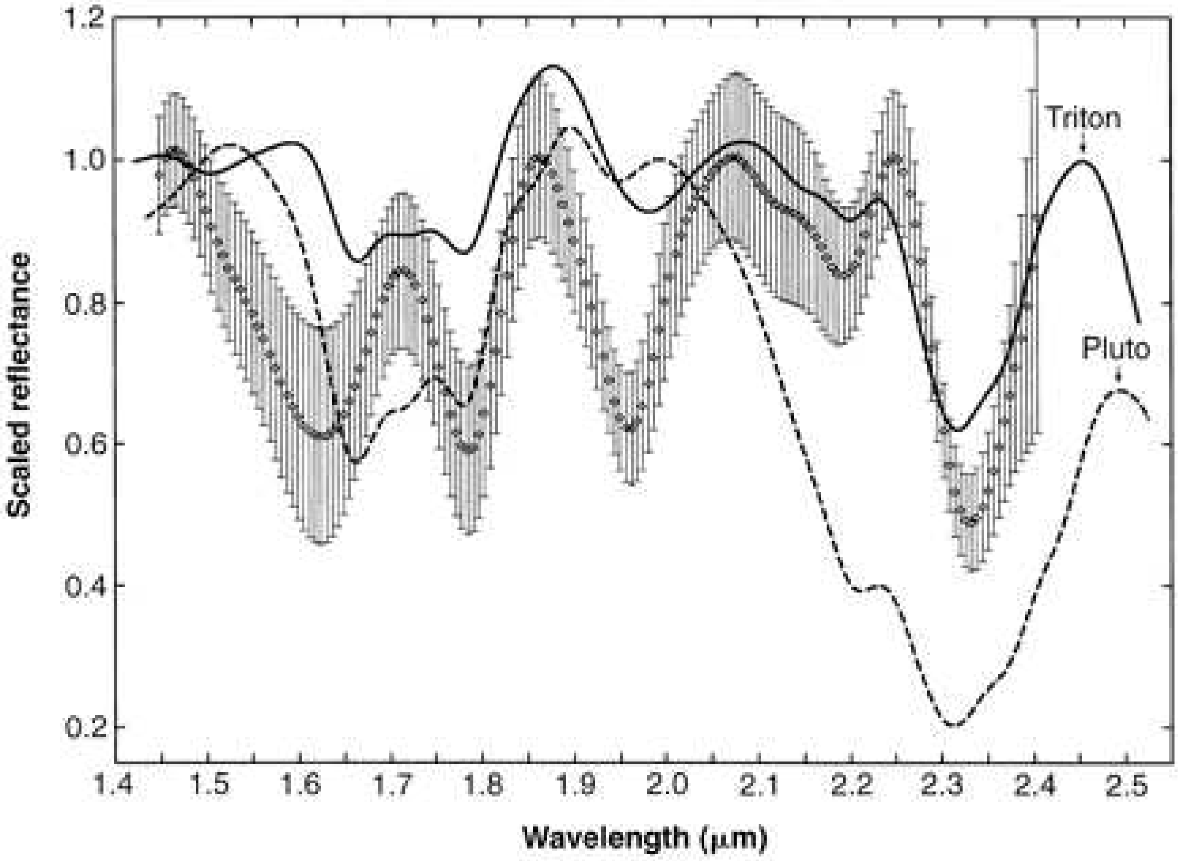


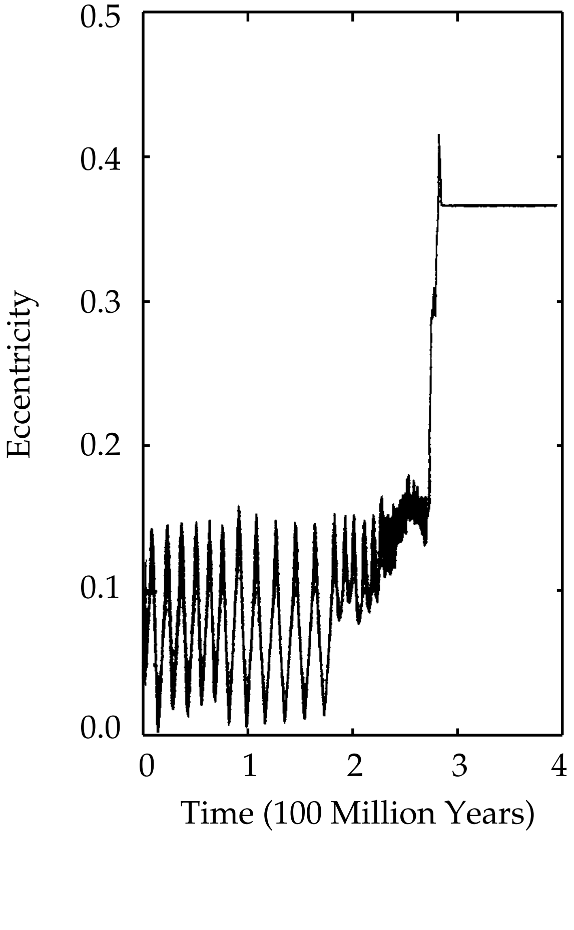
FIGURE 5

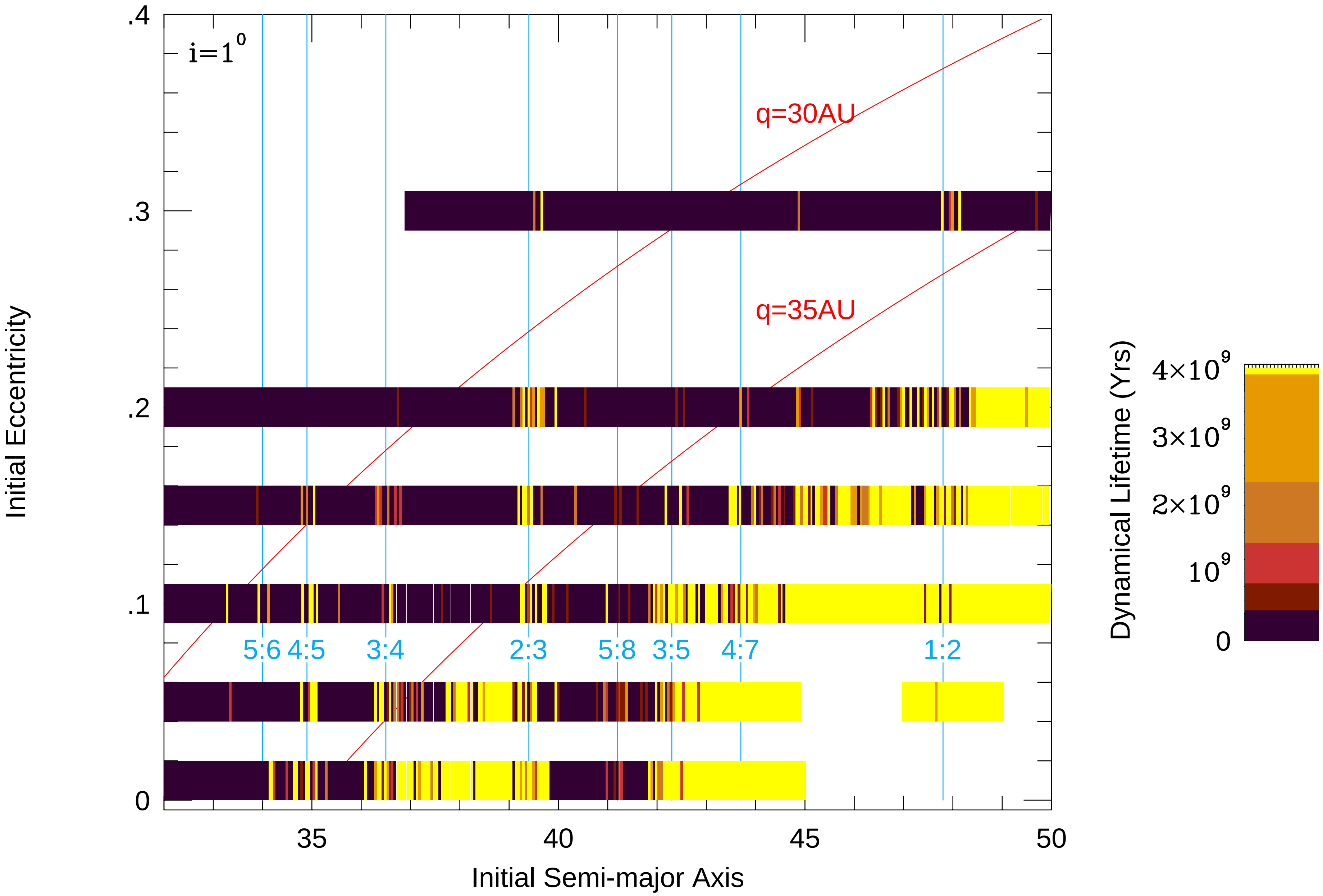
FIGURE 6



Triton
Pluto
Scaled reflectance
Wavelength (μm)
1.2
1.0
0.8
0.6
0.4
0.2
1.4
1.5
1.6
1.7
1.8
1.9
2.0
2.1
2.2
2.3
2.4
2.5

FIGURE 7

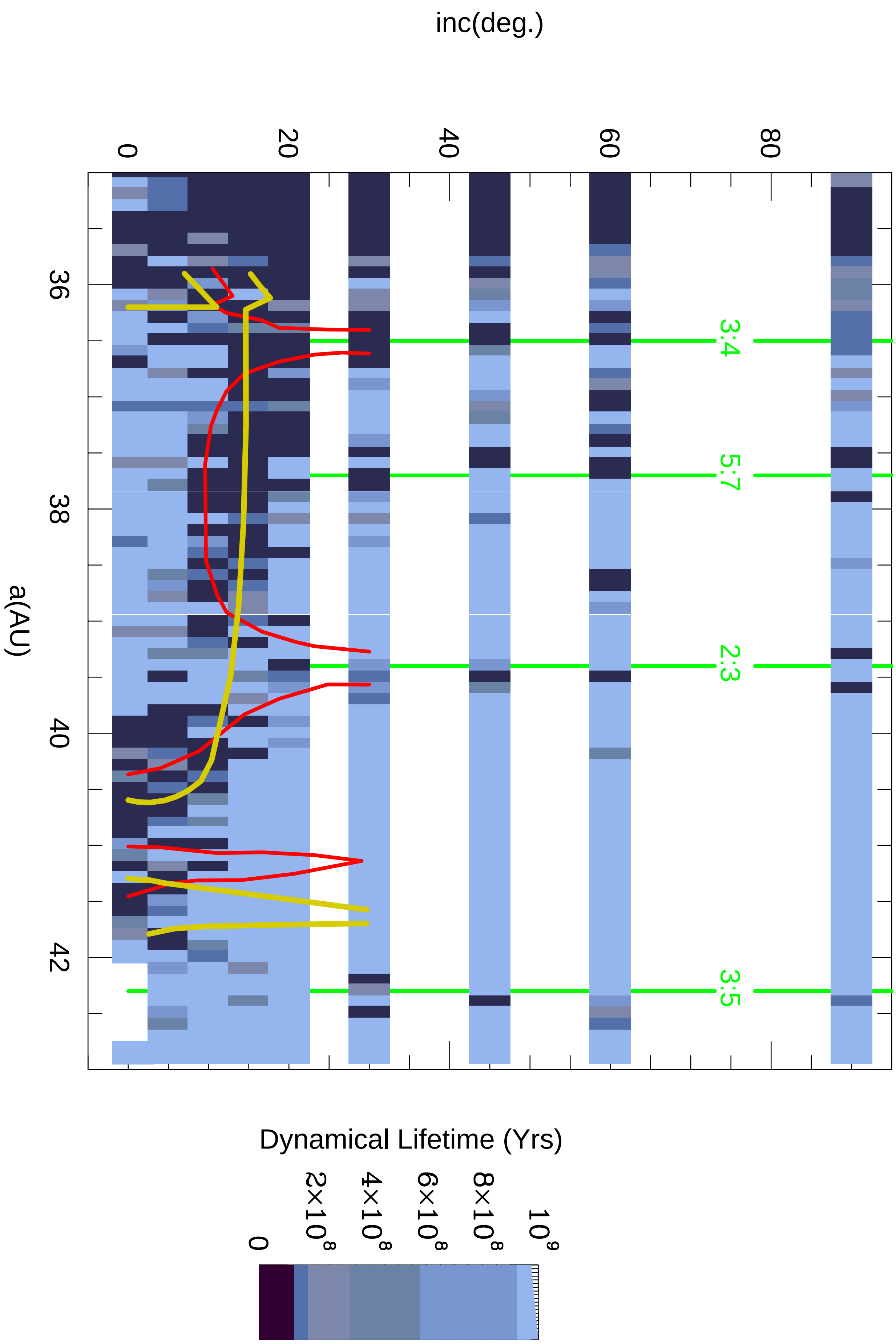




i=1°
q=30AU
q=35AU
5:6 4:5
3:4
2:3
5:8
3:5
4:7
1:2
.4
.3
.2
.1
0
35
40
45
50
Initial Semi-major Axis
Initial Eccentricity
Dynamical Lifetime (Yrs)
4×10^9
3×10^9
2×10^9
10^9
0

FIGURE 8

FIGURE 9



inc(deg.)
a(AU)
3:4
5:7
2:3
3:5
Dynamical Lifetime (Yrs)
0
2×10^8
4×10^8
6×10^8
8×10^8
10^9

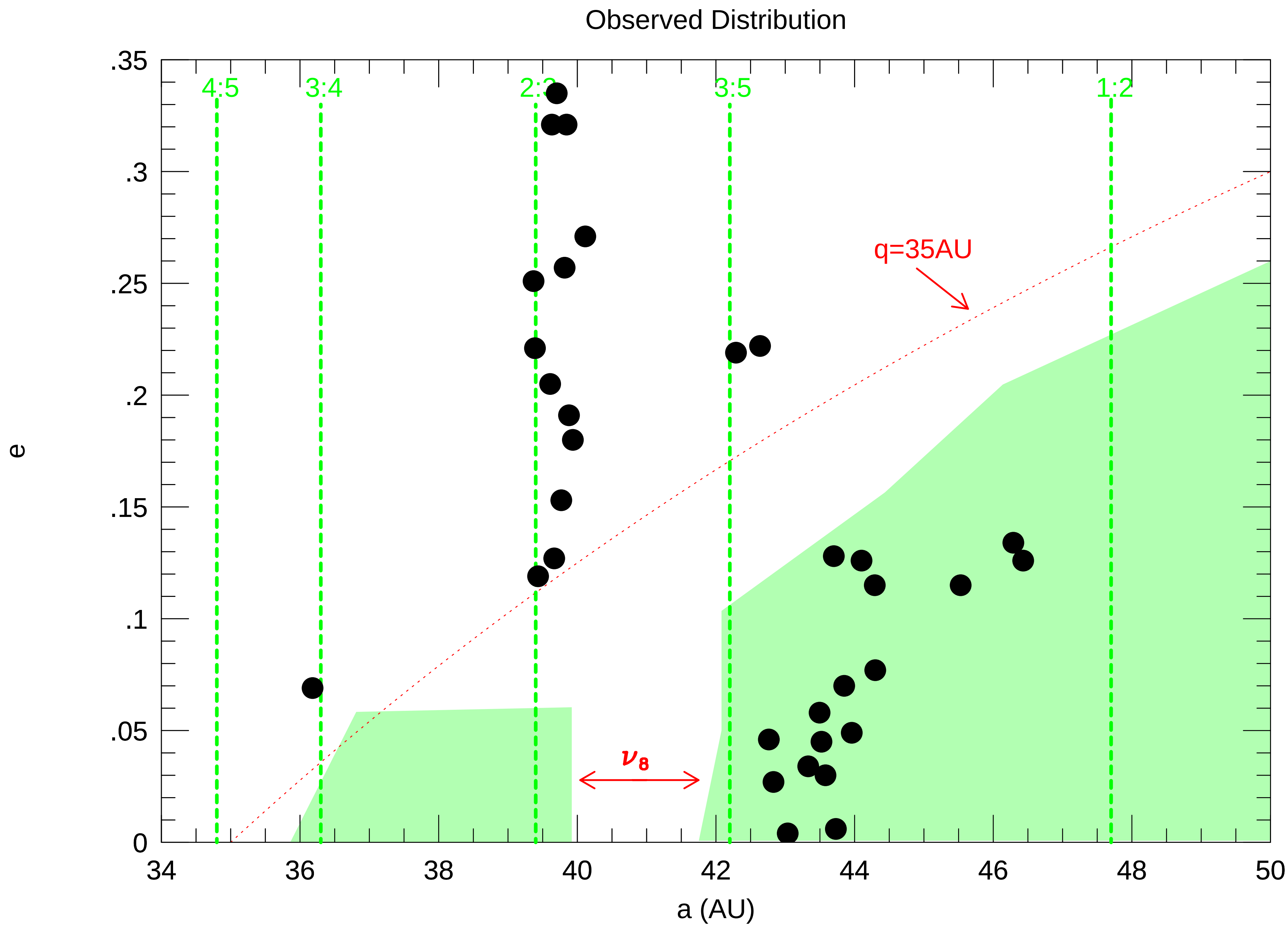

Observed Distribution
4:5
3:4
2:3
3:5
1:2
q=35AU
ν_8
e
a (AU)
.35
.3
.25
.2
.15
.1
.05
0
34
36
38
40
42
44
46
48
50

FIGURE 10

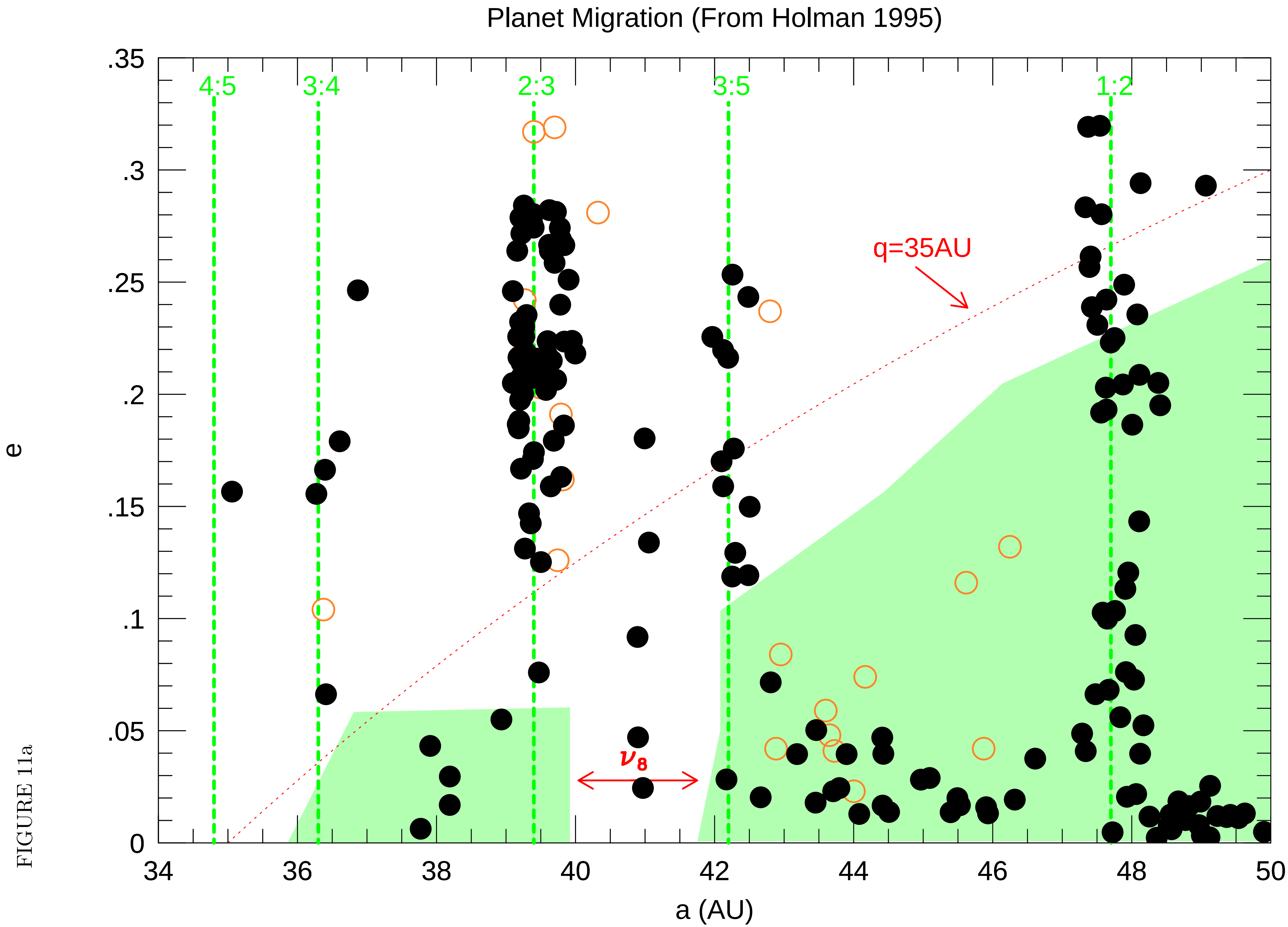
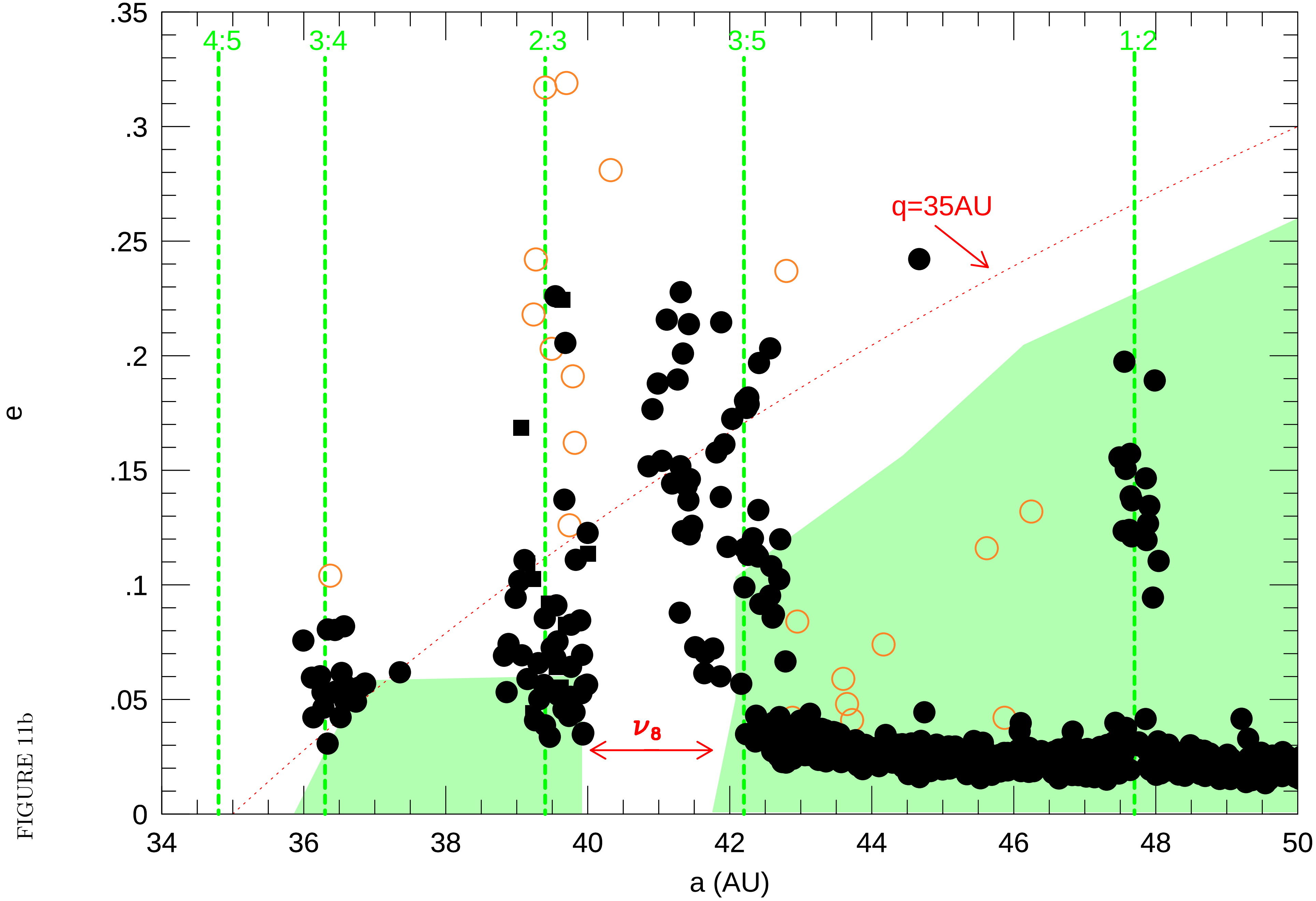

Planet Migration (From Holman 1995)
4:5
3:4
2:3
3:5
1:2
q=35AU
ν_8
e
a (AU)
.35
.3
.25
.2
.15
.1
.05
0
34
36
38
40
42
44
46
48
50

FIGURE 11a



.35
.3
.25
.2
.15
.1
.05
0
e
34
36
38
40
42
44
46
48
50
a (AU)
4:5
3:4
2:3
3:5
1:2
q=35AU
ν_8

FIGURE 11b

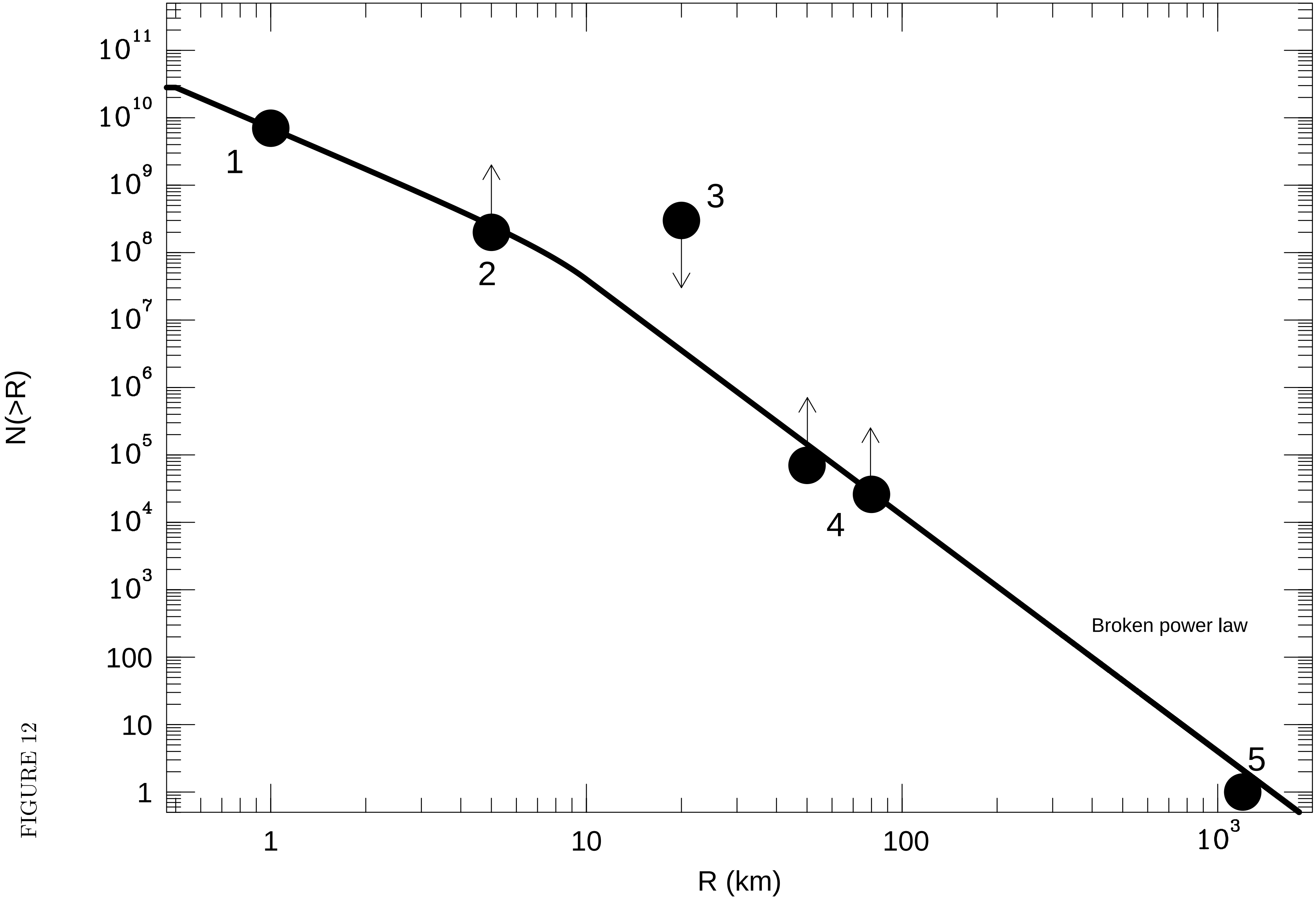
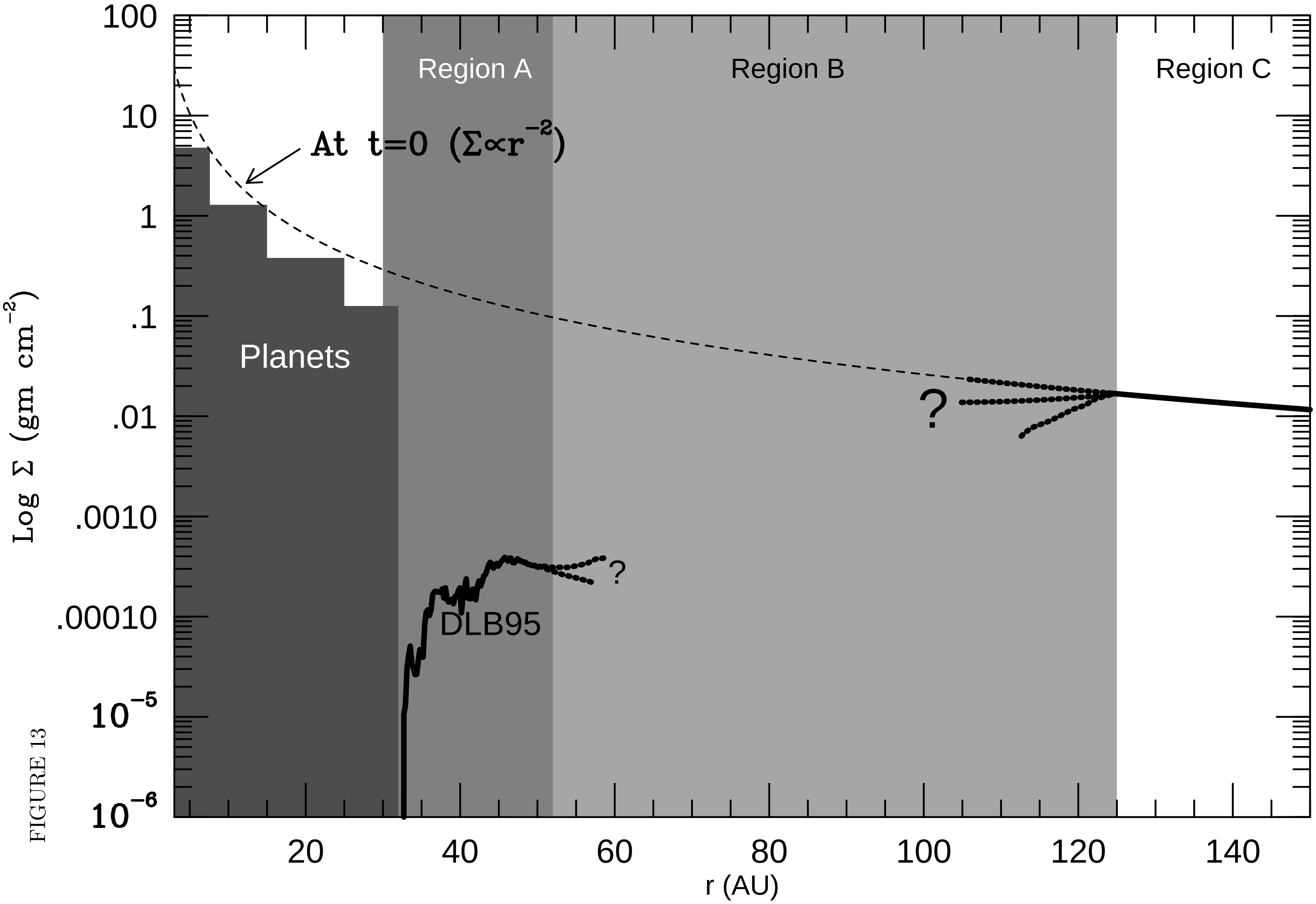


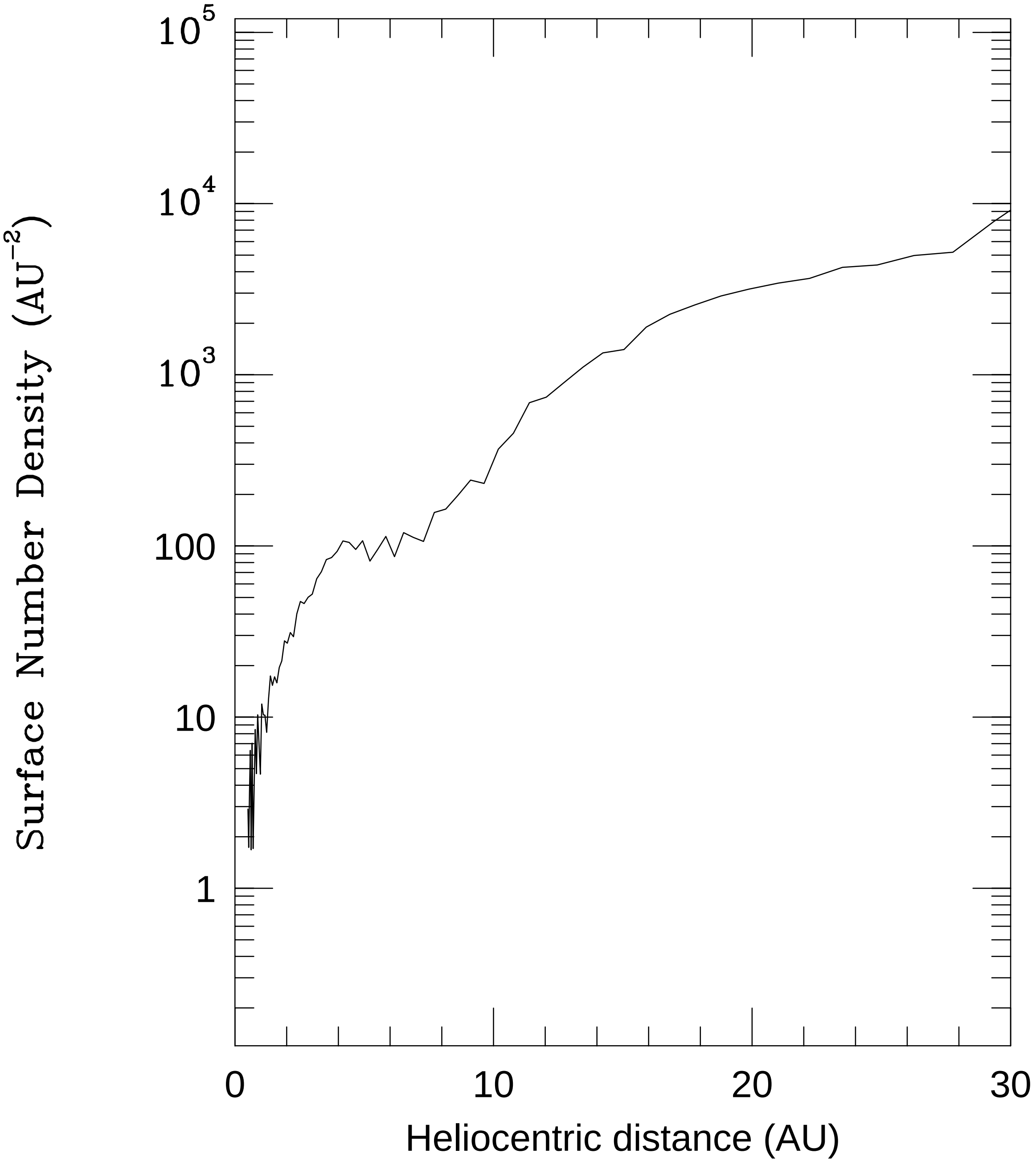
FIGURE 12



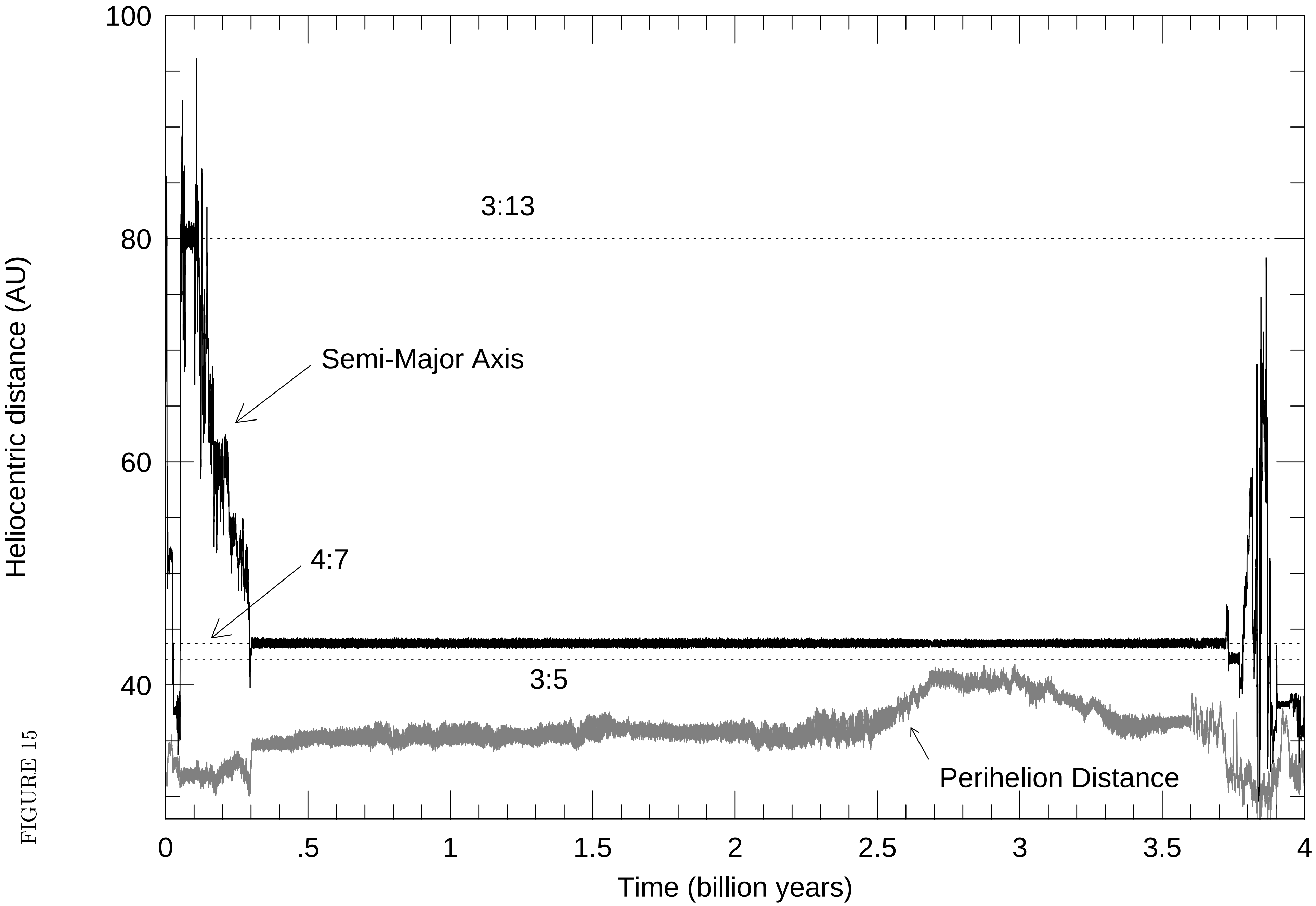
Region A
Region B
Region C
At t=0 ($\Sigma \propto r^{-2}$)
Planets
DLB95
?
?
100
10
1
.1
.01
.0010
.00010
10^{-5}
10^{-6}
20
40
60
80
100
120
140
r (AU)
Log Σ (gm cm^{-2})

FIGURE 13

FIGURE 14



10^5
10^4
10^3
100
10
1
0
10
20
30
Heliocentric distance (AU)
Surface Number Density (AU^{-2})



100
80
60
40
Heliocentric distance (AU)
3:13
Semi-Major Axis
4:7
3:5
Perihelion Distance
0
.5
1
1.5
2
2.5
3
3.5
4
Time (billion years)

FIGURE 15