Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

C. R. Physique ●●● (●●●●) ●●●–●●●

COMPTES RENDUS PHYSIQUE

New Frontiers in the Solar System: Trans-Neptunian Objects/Les Nouvelles Frontières du Système Solaire : les objets transneptuniens

The discovery and exploration of the trans-Neptunian region

John Keith Davies

UK Astronomy Technology Centre, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK

Presented by Pierre Encrenaz

Abstract

First predicted qualitatively in the 1940s, quantitatively in the 1980s and finally discovered in the 1990s, the planetesimals beyond Neptune provide a fossil record of the early history of the solar system. A decade of observations have shown that the region is far more complicated, both dynamically and compositionally, than originally suspected and it continues to challenge both observers and modellers who attempt to understand it. This region of space provides an observational link between evolved planetary systems like the solar system and the disks of material recently detected around other nearby Sun-like stars. ***To cite this article: J.K. Davies, C. R. Physique ●●● (●●●●).***

Résumé

La découverte et l'exploration de la région trans-neptunienne. Prédits qualitativement dans les années 1940, quantitativement dans les années 1980 et finalement découverts dans les années 1990, les planétésimaux sités au delà de Neptune constituent un enregistrement fossile des débuts du système solaire. Une décennie d'observations a montré que cette région est bien plus compliquée qu'initialement envisagé, tant en ce qui concerne sa dynamique que sa composition, et elle continue à défier les observateurs aussi bien que les modélisateurs qui tentent de la comprendre. Cette région de l'espace fournit un chaînon évolutionnel entre les systèmes planétaires évolués, tel le système solaire, et les disques de matière récemment détectés autour d'étoiles quasi-solaires voisines. ***Pour citer cet article : J.K. Davies, C. R. Physique ●●● (●●●●).***

Keywords: Kuiper belt; Trans-Neptunian objects; Plutinos; Scattered disk objects

Mots-clés : Ceinture de Kuiper ; Objets trans-neptuniens ; Plutinos ; Objets du disque dispersé

---

1. Introduction

In papers published in 1943 and 1949 [1,2] retired soldier and amateur astronomer Kenneth Essex Edgeworth propounded his thoughts about the formation of the solar system and suggested that there should exist a population of small icy condensations beyond Neptune. In the first of these papers he even remarked that, from time to time, one of these condensations might wander into the inner solar system to become a comet. Edgeworth's speculations, which were very qualitative in nature, were not taken up at the time but somewhat similar ideas were propounded by Gerard Kuiper in 1951 [3].

The concept of a trans-Neptunian comet belt was also discussed by Fred Whipple in 1964 [4]. Whipple considered whether comets in such a cloud might be detected individually or via the integrated effects of their combined reflectivity in the form of a second 'zodiacal light' originating beyond the conventional planetary region. His conclusion was that even extremely large comet nuclei, i.e., objects 100 km diameter, would shine at only magnitude 22 and so be very hard to detect with the technology of the day. Approaching the problem from the other direction he estimated that if the likely mass of the hypothetical

---

E-mail address: jkd@roe.ac.uk (J.K. Davies).

1631-0705/$ – see front matter © 2003 Published by Éditions scientifiques et médicales Elsevier SAS on behalf of Académie des sciences.
doi:10.1016/j.crhy.2003.09.008

comet belt was distributed as a multitude of 1 km diameter objects, then their combined light would contribute a diffuse glow of only 8.5 magnitudes per square degree, almost 100 times fainter than the night sky. On this basis he concluded that "direct observational evidence for the existence of a cometary belt may not be available for some time to come" [5]. Having reached this conclusion Whipple and collaborators attempted to determine the mass of this hypothetical comet belt dynamically by searching for perturbations in the orbits of several periodic comets whose aphelia took them close to the trans-Neptunian region [6]. They detected no such perturbations and so placed upper limits on the mass of a comet belt as 0.5 M_{EARTH} at a distance of 40 AU and 1.3 M_{EARTH} within 50 AU.

Although briefly revived by Fernádez in 1980 [7] the concept of a comet belt then languished for two decades until simulations by Duncan, Quinn and Tremaine [8] showed that the number and inclination distribution of the short period comets was not consistent with the capture of long period comets entering the solar system from the Oort Cloud. Having ruled out an Oort cloud origin, they showed that the best source for the short period comets lay in leakage inwards from a hitherto unobserved low inclination disc of small icy bodies just beyond the orbit of Neptune. They called this hypothetical structure the 'Kuiper Belt'.

Although Duncan et al.'s paper made testable predictions about the existence of a comet belt, there were only a small number of attempts to search for a trans-Neptunian population during the late 1980s. These attempts were all unsuccessful, although in some cases this was due more to poor luck than to any fundamental problems with the search strategy. However rapid improvements in the size and sensitivity of CCD detectors during the late 1980s and early 1990s eventually led to the discovery, by Jewitt and Luu, of a distant slow moving object initially designated 1992 QB_1 [9] and now given the minor planet number (15 760). Follow-up observations soon established that (15 760) 1992 QB_1 was of order 250 km in diameter and in an orbit ranging from 41–47 AU from the Sun. Over the next few years a few dozen broadly similar objects were discovered until the development of still larger CCDs and automated software to search the images resulted in a rapidly increasing discovery rate after mid-1998.

By 2003 of order 1000 such objects had been discovered. Many of these now have secure orbits, indeed some of them have already been assigned minor planet numbers and names. The pioneering discovery phase is now over and astronomers are moving on the essential follow-up and characterisation of this newly discovered region of the solar system. This requires a multi-disciplinary approach, blending dynamical studies of the population as a whole with characterisation of individual objects and statistical investigations of their physical properties.

2. The Centaurs

Although not strictly 'trans-Neptunian', the Centaurs are a population of icy objects in unstable, planet crossing orbits in the outer solar system. They are bodies which have escaped from the trans-Neptunian region and whose orbits are evolving under the gravitational influence of the giant planets. Centaurs have dynamical lifetimes of only $\sim 10^7$ years before they are either ejected from the solar system or perturbed inwards to join the Jupiter family comets. The first Centaur, 2060 Chiron, was discovered in 1977 but it was not until 1992 that the second example, 5145 Pholus, was discovered. Since then, like the trans-Neptunians, the known population of Centaurs has increased rapidly.

Since they are escaped trans-Neptunians which are closer to Earth and hence brighter and easier to study, it is common to use observations of Centaurs to infer the properties of their more distant cousins. However, such interpretations must be done with care since the Centaurs are a dynamically and evolutionarily different population. For example there are several case of cometary activity amongst the Centaurs (most notably 2060 Chiron and C/NEAT 2001 T4) which must surely have had a significant influence on their surface properties. The Centaur population is indeed worthy of study, both for its own sake and for what it can reveal about the trans-Neptunian region, but space prevents more than passing references to these objects in this review.

3. Searches and structure

Searches for trans-Neptunian objects may be characterised as wide and shallow, covering large areas of sky with relatively low sensitivity, or pencil beams, going to very high sensitivity over small areas of sky. Both approaches have merit. Wide and shallow searches such as that of Trujillo and co-workers [10] discover the small number of brighter, and presumably larger, objects which are well suited for physical observations and which bridge the size gap between traditional planets and asteroids. Pencil beam searches using large telescopes have much higher sensitivity to faint objects and so discover objects which are on average smaller and more distant. By co-adding multiple images of the same region which have been shifted at the likely rates of slow moving objects, these pencil beam surveys can reach very faint limiting magnitudes, typically $R = 27$–28, which corresponds to $a \sim 25$ km object at a distance of 40 AU. A decade of such search programmes have now mapped out the broad structure of the trans-Neptunian region and shown that the outer solar system is much more complicated than originally anticipated.

The first two discoveries, (15 760) 1992 QB_1 and 1993 FW, were in what has since become known as the classical Kuiper Belt. This comprises a population of objects in near circular orbits with semi-major axes around 45 AU. Such orbits are stable against gravitational perturbations by Neptune over the age of the solar system and this region most closely resembles the concepts first propounded by Edgeworth, Kuiper and Whipple. The classical objects represent the majority of the presently known trans-Neptunian objects. A second population was recognised when Marsden and others realised that several objects discovered in 1993 were in or near 3 : 2 mean motion resonance with Neptune. The similarity of these orbits to that of Pluto, whose orbit crosses Neptune and which only survives by virtue of being in 3 : 2 resonance with Neptune, has led to members of this population being known as 'Plutinos' [11]. The first member of a third dynamical class of objects was discovered in 1996. Initially designated 1996 TL_{66}, minor planet (15 874) occupies a highly eccentric orbit which at perihelion is inside the classical Kuiper Belt but at aphelion is 134 AU from the Sun. (15 874) 1996 TL_{66} was the prototype of what has become known as the scattered disk and represents a population of objects gravitationally ejected by Neptune. Note that this scattered population also includes objects with perihelia inside Neptune's orbit, e.g., (29 981) 1999 TD_{10} which has a perihelion distance of 12.3 AU but whose aphelion is at 184.6 AU.

All three of these populations must be explained by any dynamical model of solar system formation. Planetary migration, the gradual movement outwards of the still forming proto-Neptune as it exchanged angular momentum during gravitational interactions with the numerous but much smaller planetesimals which surrounded it, would have caused the 3 : 2 resonance of Neptune to sweep across a region likely to contain many smaller objects. Once stabilised by this resonance these 'proto-Plutinos' would have been carried along as the resonance evolved outwards, allowing the Plutino population to grow as it did so [12]. However, the basic resonance sweeping hypothesis cannot be the complete answer to the existence of the population of resonant objects since although it allows the 3 : 2 resonance to be filled, it cannot explain why Neptune's 2 : 1 resonance is less well populated. As the 3 : 2 resonance was sweeping up Plutinos, the 2 : 1 resonance should have been moving through what is now the classical Kuiper Belt depleting this region and becoming populated itself. That the 2 : 1 resonance is not well populated may indicate that Neptune's outward migration occurred too quickly for a significant number of objects to be captured into this region.

Two other conclusions have emerged from a decade of searches and the equally vital but unglamorous work of astrometric follow-up. Firstly the Kuiper Belt is not confined to the solar system's invariable plane. Even though searches tend to concentrate near the plane of the solar system, for it is here that discoveries are most likely, it is now clear that the half thickness of the disk is at least 20 degrees. Furthermore the average eccentricity of the orbits is unusually high. Such a wide range of inclination and eccentricity suggests that some mechanism 'pumped' orbital energies of the objects in the disk to higher values.

There is also increasing evidence for an edge to the classical Kuiper Belt. The use of larger telescopes able to reach fainter limiting magnitudes should by now have discovered a significant number of objects in quasi-circular orbits beyond about 50 AU. However, almost no such objects have been found. Unless one invokes unrealistic assumptions about sudden changes in reflectivity or in the size distribution of an outer belt population which would make such objects much fainter and so more difficult to discover, then it becomes clear that the belt has been truncated at some point. This truncation could be explained by the scattering of objects from the disk by one or more large (Earth or Mars mass) planetary embryos which were later ejected themselves. An alternative and more widely favoured explanation is that another star passed within $\sim$150 AU of the forming solar system and affected the outer regions of the protoplanetary disk at a quite early point in the planet building process. Such a close passage would be unlikely today but, assuming that the Sun was born in a cluster with a dissolution time of the order 10^8 years, then it is quite conceivable in the distant past. Such an encounter would pump up the velocity dispersion in the outer parts of the disk such that collisions in this region would become erosive, halting the growth of planetesimals and hastening their eventual destruction by mutual disruptive collisions [13].

Despite the remaining uncertainties in the models it is clear that the present structure of the trans-Neptunian region is a fossil which provides vital clues to the formation of the solar system. For more details see the review by A. Morbidelli and H. Levison elsewhere in this issue.

4. Size distributions

A fairly reliable size distribution for the objects in the Kuiper belt has now been established. The population is believed to be described by a differential power law of the form:

$$N(r)\,\mathrm{d}r = \Gamma r^{-q}\,\mathrm{d}r,$$

where $N(r)\,\mathrm{d}r$ is the number of objects with radii between r and $r + \mathrm{d}r$ and Γ and q are constants. Various surveys are in general agreement that the value of q is of order 4, implying that the trans-Neptunian region contains approximately 10^{10} objects greater than 1 km in radius and about 10 with radii greater than 1000 km. These surveys also suggest that there are 1–3 objects similar in size to Pluto still awaiting discovery. The smaller end of the distribution is still being probed by deep searches using 8 m class telescopes but the controversial HST result of Cochran et al. [14] remains at odds with all the other data. The density of a typical trans-Neptunian object is not well constrained but taking the assumption of a density equal to 1 (approximately midway between comet nuclei and the density of Pluto) leads to a total mass of trans-Neptunians equal to about 0.08 M_{EARTH}. To place these numbers in perspective, the trans-Neptunian region has several hundred times the mass of the asteroid belt and about 500 times as many objects with diameters larger than 200 km.

The number of comparatively large objects in the trans-Neptunian region leads to another interesting result. With the large volume of space involved, the growth of objects by accretion must be slow, and calculations [15] have shown that it is impossible to grow the larger objects unless the original density of the region was many times higher than it is today. Only in a dense primordial belt can collisions occur frequently enough to grow objects the size of the larger trans-Neptunians in the 10^8 years before the growth of Neptune inhibits further accretion. This in turn requires that most of the original mass of the region must have been removed. Scattering of objects following gravitational interaction with Neptune can remove some mass, but is probably not sufficient to deplete the region to the degree seen today. A favoured mechanism for this removal is mutual collisions grinding down the objects and releasing dust which can then be removed from the solar system by cosmic ray sputtering, radiation pressure, Poynting–Robertson drag and so on. For further discussion of this issue see the paper by S.A. Stern and S. Kenyon on accretion and erosion elsewhere in this issue.

5. Colours and trends

In view of the faintness of the first trans-Neptunian objects discovered initial attempts at determining their physical properties were made via broad band photometry. One early result of these observations was that the BVRI colours of the trans-Neptunian population were quite diverse. This was rather surprising since the expectation was that all of the objects would have very old, ‘space weathered’ surfaces with broadly similar properties. The colour diversity implied either that the trans-Neptunian population had a wide range of initial bulk compositions or that some mechanism was gradually changing the colours of the objects with time. Given that the range of temperatures across the protoplanetary disk beyond Neptune was small, only about 10 K, significant bulk chemical diversity of objects formed in-situ within the trans-Neptunian region seemed unlikely and so initially an impact resurfacing model was favoured.

In the impact resurfacing scenario objects composed primarily of low molecular weight ices such as H_2O, CH_4, CO, CO_2 and NH_3 would begin life with relatively clean surfaces which would be bright and neutrally reflective. Over time cosmic ray and solar photon bombardment would cause chemical changes in the surface ices, gradually forming a refractory layer of more complex organic materials which would cause the surface to redden and become darker. These red surfaces would subsequently be disrupted by impacts which would excavate fresh ices, returning some or all of the surfaces to their original, neutrally reflective, state. Thus the colour of an object at any given time would be determined by the relative rate of the impacts which resurface the body and the gradual reddening under the influence of radiation.

Although conceptually attractive, the impact resurfacing model fails to meet some critical tests. Firstly, detailed laboratory studies of the effects of irradiation of ices has shown that the situation is more complicated than a simple scenario of blue = fresh ice, red = old refractory organics. The end states of irradiation of likely outer solar system surfaces depends on their composition and various combinations of ices and pre-existing refractory compounds can produce a bewildering variety of possible colours.

Equally telling is that there is no clear evidence of large scale colour changes on the surfaces of trans-Neptunians with rotation. A clear prediction of the impact resurfacing model is that objects which have recently been subjected to a large impact should have large scale colour diversity across their surfaces. That is to say one hemisphere may have a large spot of fresh ice ejected from an impact which covers the ancient reddened surface while the opposite face retains its red colour. Few such colour changes with rotation have been convincingly detected to date.

There remains a lively debate about the origin of the colour diversity. The initial data was taken with relatively small telescopes and on a small number of objects, but the increasing availability of time on 4, 8 and 10 m telescopes has allowed the sample to be expanded significantly while at the same time drastically reducing the observational errors on the individual measurements. Using these expanded datasets, various groups have sought trends in colours with such parameters as inclination, eccentricity and heliocentric distance, see the article by A. Doressoundiram and H. Boehnhardt elsewhere this issue for more details. To date most of these trends have only been demonstrated at the 2–3 sigma level but evidence from colour and absolute magnitude data is increasingly suggesting that the classical Kuiper Belt may comprise two separate populations. These are an energetically cold population of objects with inclinations below about 7° which are intrinsically red and a population of, on average, larger and bluer objects at higher inclinations.

Although first recognised in earlier data sets, this conclusion of a hot and a cold population is supported by the results of wide and shallow surveys which are detecting disproportionately large numbers of bright objects in high ($i > 7^\circ$) inclination orbits. This population of objects at high inclination presents a challenge to the resonance sweeping models used to explain the population of the resonant objects such as the Plutinos. The large, high inclination blue objects may have originated closer to the Sun and been injected into the trans-Neptunian region after non-reversible gravitational encounters with the migrating proto-Neptune [16].

6. Spectra and compositions

Broad band colours, e.g., BVRI in the optical and JHK in the infrared, are useful for classification, but the wide bandpasses of the filters limits the diagnostic value of colour data for compositional studies. The extreme faintness of the trans-Neptunians has meant that spectroscopic information has been difficult to obtain, but in recent years observations from the 10 m Keck telescope and the 8 m VLT has begun to open this important area of study. Of most relevance are observations in the interval 1–2.5 microns since this is where absorption features of water ice and common organic molecules are found. Spectroscopic observations and their interpretation are discussed in detail elsewhere in this issue and it is sufficient to mention that spectroscopy has tended to confirm the diversity of the trans-Neptunian population. Some objects have infrared spectra which are essentially featureless while others show broad absorption features at around 1.5 and 2.0 microns which are typical of water ice. In some cases there is evidence that the depth of these features varies as the objects rotate, suggesting that the ices may be distributed in a patchy manner across their surfaces, although HST observations of such an effect on the Centaur object 8405 Asbolus have not been supported by more recent ground based time resolved spectroscopy from the VLT.

In a small number of cases there are other features which suggest the existence of hydrated materials which, if confirmed, would indicate that aqueous alteration has occurred on some of these bodies.

7. Size and albedo

The sizes of trans-Neptunian objects are usually estimated using their visible magnitude, however optical measurements provide only the product of the albedo and physical cross section. A priori the albedos of the objects are not known and most diameter estimates have traditionally been made by assuming an albedo of 4%, which is typically that of short period comet nuclei, but which remains not-proven for trans-Neptunians.

Only in the case of 50 000 Quaoar has a diameter, and hence albedo, been established directly. Comparisons of the point spread functions of 50 000 Quaoar and of a nearby star in 14 HST images allowed, after convolution of the motion vector of the minor planet with the PSF, the angular size of Quaoar to be established as 40 milli-arcsec [17]. This leads to a diameter of Quaoar of 1260 ± 190 km and implies an albedo of 10%, rather higher than usually assumed.

In cases, where the objects are not resolved, sizes and albedos can determined using thermal model techniques developed for the study of main belt asteroids. These methods require a quasi-simultaneous determination of the reflected (visible) light and the emitted (thermal) radiation. For an object in thermal equilibrium these two quantities must equal the energy being received from the Sun and the use of asteroid thermal models allows both the diameter and albedo of the object to be calculated. Unfortunately the temperature of the objects in the trans-Neptunian region is low, around 40 K, and consequently the peak of their thermal emission is at wavelengths around 60 microns, a region of the electromagnetic spectrum which is not accessible from ground based telescopes due to absorption by the Earth's atmosphere.

Two alternative strategies exist to circumvent this problem and both have had some success. Space based observations of a small number of trans-Neptunians were made using the ISO satellite, but the low angular resolution of its 60 cm telescope and the confusing diffuse backgrounds from interplanetary and interstellar dust made these observations challenging and difficult to interpret. Only two results, with one of them having un-explained astrometric uncertainties, were published [18]. The planned SIRTF space telescope, with its new generation of infrared detectors and a slightly larger mirror, offers considerable potential to re-attempt this approach.

An alternative observational technique is to use sub-mm and radio-telescopes to detect the Rayleigh–Jeans tail of the thermal emission using ground based instruments. In this case the much larger collecting area of the telescopes is balanced by the weaker emission from the objects and the difficulties of observing through the atmosphere. Despite this, a few of the larger trans-Neptunians have been observed in this way and, for example, the albedo of 20 000 Varuna has been determined as 7% using sub-mm observations [19].

Although the number of secure results is small, the overall picture being painted by these measurements is that the average albedo of the trans-Neptunians is rather higher than the value of 4% which has generally been assumed to date. Although these values are much smaller than that of Pluto, which has areas in which the albedo approaches 50–70%, Pluto's high albedo is the result of surface frosts being deposited from a tenuous atmosphere. Objects such as Varuna are much less massive than Pluto and so unable to retain an atmosphere, making a global surface frost unlikely.

One other technique which has the potential to add considerable new information to the size-albedo issue is the observation of stellar occultations by trans-Neptunian objects. Stellar occultations by Pluto led to the discovery of that planet's atmosphere and have since been used to probe its atmospheric structure. If other trans-Neptunians have atmospheres then occultations offer a means to detect and study them, as well as providing direct measurements of diameter. Such observations will be challenging as they require ephemerides with sufficient precision to predict the track of the objects along the Earth's surface and the availability of sufficiently large telescopes along that track to obtain high signal-to-noise observations. However the experience gained by several groups in deploying mobile telescopes along the tracks of Pluto, its satellite Charon, Neptune's large moon Triton and Centaur 2060 Chiron, plus the likely availability of the SOFIA airborne observatory with its 2.5 m telescope, suggest that such opportunities may soon become realisable.

8. Rotation rates and shape

The combination of more bright objects to study and improved access to medium sized (2–4 m) telescopes has meant that rotation periods and lightcurve parameters are being determined for more and more trans-Neptunian objects. The most comprehensive programme has been by Sheppard and Jewitt [20] who find that about one third of the trans-Neptunians they observe exhibit systematic brightness variations greater than 0.15 magnitudes. Since there is no evidence for colour variation with rotation, they attribute this variability to the changing projected area of rotating, non-spherical objects. This observation, which is supported when the dataset is expanded to include photometric data from the literature, can shed light on the macroscopic physical structure of individual trans-Neptunian objects.

Lightcurve studies require considerable amounts of telescope time which is generally only available on medium sized telescopes (most of Sheppard's work has been done on a 2.2 m telescope). This in turn means that observations tend to be targeted on the brighter, and presumably larger, objects. These objects, which are expected have diameters in excess of 250 km, would be expected to be spherical in shape due to gravitational self-compression and so to exhibit low amplitude rotational lightcurves. However, this is not generally the case, several of Sheppard's objects have peak to peak lightcurve amplitudes of order 0.5 magnitudes and rotation periods of less than 12 hours. If this variation is the result of non-spherical shapes then these objects must be quite highly elongated.

As a specific example, the large trans-Neptunian 20 000 Varuna was found to have a rotation period of 6.34 hr with a peak to peak amplitude of 0.42 magnitudes [21] leading to the conclusion that Varuna was elongated, with the ratio of its axes projected onto the plane of the sky being 1.5 : 1. If this is so then for plausible material strengths the conclusion is that Varuna cannot be a solid body but is a rotationally distorted 'rubble-pile' with a bulk density close to 1. This weak internal structure is presumably the result of a history of fracturing and possibly disruption and re-assembly by impacts and provides further evidence for a period of an intense collisional epoch in the history of the trans-Neptunian region. Comparison with objects in the main asteroid belt confirms that statistically the trans-Neptunian objects are less spherical and have higher specific angular momentum, a feature which is presumably a relic of their epoch of formation.

9. Binaries

Charon, the large satellite of Pluto was discovered in 1977 and since the barycentre of the Pluto–Charon system lies outside Pluto itself, the system is better described as a binary 'double planet' rather than a planet and its satellite. Despite the obvious power of observations of binarity to probe directly the masses of trans-Neptunian objects, the first binary, 1999 WW_{31} was discovered serendipitously [22]. Follow up observations from both ground based telescopes and the HST established that the system has an orbital period of 574 ± 10 day and a highly eccentric ($e > 0.8$) orbit with a semi-major axis of 22 300 km. This gives a combined mass of about 2.7×10^{18} kg, approximately 5500 times less than the Pluto–Charon system.

Since the discovery of 1998 WW_{31} eight other binary systems have been found [23]. This leads to the conclusion that the binary fraction amongst the trans-Neptunian population is of order 5%. With such small numbers of discoveries it is hazardous to try and draw detailed statistical conclusions but it is already clear that binaries are found amongst both the resonant (Plutino) and non-resonant (Classical) populations and over a wide range of inclinations.

Binary asteroids are now known to be common amongst both the main asteroid belt and the near Earth asteroid population. However there are fundamental differences between these and several of the trans-Neptunian binaries, notably the wide separations and small size differences amongst some of the trans-Neptunian pairings. These wide spacings present challenges for theorists who attempt to describe the formation mechanism of such pairs. The total angular momentum of the wide pairs rules out a transfer of spin angular momentum to orbital angular momentum. An alternative mechanism for binary formation by tidal disruption during a planetary encounter (which is favoured for the production of binaries amongst the Near-Earth asteroids) is implausible in a region which does not contain large planets. Collisionally formed binaries would produce pairs with small separations, like the Pluto–Charon system, and several such systems, (e.g., 1996 TC_{36} and 1998 SM_{55}), have been found using the HST. In both of these two cases the ratio of brightness of primary and secondary is large (∼2 magnitudes) more in line with that of the Pluto–Charon system.

Various models for producing binary trans-Neptunians exist. These include three body interactions (which require formation when the density of the region was ∼100 higher than at present), close encounters between objects with a retinue of small bodies and two body collisions. As pointed out by Noll [23] each of these scenarios make testable predictions about the binary trans-Neptunian population that might be resolved by observations in the next few years.

Note that at such large heliocentric distances even the most widely separated trans-Neptunian binaries are stable against perturbations by the Sun and other planets although collisions and close approaches may be able to disrupt some of the more weakly bound pairs. Thus the present binaries may be just a remnant of a still larger primordial population.

10. Cometary activity

Since the trans-Neptunian objects are believed to represent a reservoir from which the short period comet population is drawn, it is natural to ask if there might be any evidence for cometary activity amongst the objects in the region. Cometary activity driven by the sublimation of water ice occurs only within the inner solar system, inside about 3 AU, but there is considerable evidence for activity from comet-like objects well outside this distance. For example, comet Hale-Bopp was already active when it was discovered some 7 AU from the Sun in 1995 and Centaur 2060 Chiron undergoes cometary outbursts even when close to its 18.8 AU aphelion.

Although direct searches have been made for comae around trans-Neptunians using the HST, none have ever been detected and preliminary claims of such detections have never been substantiated. Sub-mm observations searching for rotational lines of CO have also failed to detect evidence for comae [24], although this may not be surprising given that some models for trans-Neptunians suggest that their outer layers are severely depleted in CO. However, Hainaut et al. [25] invoked possible cometary activity to explain the photometric behaviour of (19 308) 1996 TO_{66}. Their conclusion was based on observations that the lightcurve shape and amplitude of this object underwent considerable change over the period 1997–1998 while the underlying rotation period remained the same. They explained this by suggesting that in the year-long interval between their observations some event recoated a large area of the surface resulting in a change from a low-amplitude (0.12 mag) double-peaked lightcurve to a single peaked lightcurve of significantly greater amplitude (0.33 mag). Observations in 1999 [26] showed no evidence for a coma at that time (the PSF of the object matched that of stars down to the noise floor of the data at 29 magnitudes/sq arc second) and supported the later single peaked lightcurve, although at a rather lower level (0.21 mag compared with the earlier value of 0.33 mag).

11. Towards the future

Despite a decade of activity much remains to be done before the outer regions of the solar system can be said to be well understood and there is much to look forward to. Conventional studies of individual objects will continue using the power of the present generation of large ground based telescopes, but several projects offer the potential to address specific topics.

After many false starts, NASA selected a mission to Pluto and the Kuiper Belt in 2001. The New Horizons project is scheduled to launch in 2006 and fly through the Pluto–Charon system in 2016 or 2017. Although no specific target has yet been identified, it is statistically likely that the spacecraft will have sufficient fuel to encounter one or more 35 km diameter trans-Neptunian objects some years after the Pluto encounter. Such a flyby of a trans-Neptunian will provide detailed in-situ measurements and provide information on surface geology, bulk composition, surface compositional variegation, albedo and mass. Searches for possible targets are hampered by considerable uncertainty of the volume to be explored as this depends critically on the accuracy of the launch and the amount of fuel expended during post launch trajectory tuning to ensure the desired geometry at the Pluto encounter. Adding to the difficulty of searching for a suitable target is that the region to be searched lies close to the galactic centre where the background star density is very high.

The small end of the trans-Neptunian size distribution can be probed using serendipitous observations of stellar occultations as objects pass in front of background stars. Although it does not allow any astrometric or physical follow-up, this technique does offer a chance to determine the population statistics of very small (∼100 m) objects which could not be detected by any other method. In the case of small trans-Neptunians the angular size of the objects is comparable to the angular size of the background stars and so the situation is more complicated than a simple dimming of the starlight, the effects of diffraction must be taken into account. Diffraction increases the effective size of the object's shadow at the Earth and increases the likely rate of detection compared with that indicted by simple geometric considerations [27]. However the events are short, and the fluctuations of the starlight are rapid, so the best chance of success comes from observing stars of small angular radii with high time resolution.

Several occultation projects are underway or planned. A few attempts have been made using high speed photometry of individual stars in which case the diffraction effects during the occultation may be detectable. In such experiments a second telescope observes the same star to verify the reality of any features. An alternative approach is to monitor larger numbers of stars at lower time resolution to increase the probability of detecting occultations. The Taiwanese–American Occultation Survey will use an array of four small (∼0.5 m) telescopes along a roughly East–West baseline to monitor several thousand stars for occultations. Each telescope will cover the same field to provide essential redundancy of the detections and eliminate false positive signals caused by atmospheric scintillation or the passage of bats, birds etc through the line of sight. The French space mission COROT will carry a 25 cm telescope for parallel projects in astro-seismology and searches for extra-solar planet transits. Thee projects involve long (5 month) staring observations at selected fields containing several thousands of stars. Due to the observing strategy these experiments may detect only small numbers of occultations by trans-Neptunians. However using the satellite for a dedicated search using more frequent observations of a much smaller number of stars might detect as many a few hundred events per day.

A new survey facility called Pan-STARRS being built in Hawaii to search for Near-Earth objects also offers the chance of discovering many more trans-Neptunians. Pan-STARRS will comprise four 1.8 m telescopes working in tandem to provide the light gathering power of a single larger telescope at lower cost and with a shorter development time. As a by-product of its repeated scans of the sky, Pan-STARRS will discover 1000 s of new trans-Neptunian objects. A built in follow-up strategy will ensure the repeated observations necessary to determine reliable orbits for the objects detected and over a decade of operation the survey will build up a huge orbital database with which to challenge the dynamical models of the formation and evolution of the trans-Neptunian region.

The GAIA astrometric mission has recently been approved as one of the next two cornerstones of ESA's space science programme. Launch is expected not later than mid-2012. GAIA can contribute to the study of trans-Neptunian objects in a number of ways. It can provide very accurate astrometry for individual objects and will produce a high quality reference star catalogue. With this much improved astrometry it will be possible to determine accurate orbits for many more trans-Neptunians. This information is essential for understanding the dynamics of the Kuiper Belt and the relative populations of the various resonances. Since, unlike most ground based surveys, GAIA will cover the whole sky rather than being targeted close to the ecliptic plane, it should detect large numbers of the brighter but rare trans-Neptunians in comprehensive survey with well understood biases. This will make it possible to characterise the poorly determined upper end of the size distribution. GAIA may also detect any remaining undiscovered Pluto sized objects, especially if they exist at high inclinations or are distant by virtue of being members of the scattered disk.

12. Conclusion

In just over a decade the 'Kuiper Belt' has gone from a theoretical concept required to explain the number of short period comets to a physically real population of objects which is challenging both observational and theoretical astronomers with many yet to be answered questions. However progress is rapid and in about another decade, 20 years after the discovery of (15 760) 1992 QB_1, we will have reached a level of understanding comparable to that which took 2 centuries to achieve for the far smaller population of much closer main belt asteroids.

Acknowledgements

JKD thanks Catherine de Bergh for her helpful comments on the original manuscript.

References

[1] K.E. Edgeworth, J. Brit. Astron. Assoc. 53 (1943) 181–188.
[2] K.E. Edgeworth, Mon. Not. R. Astron. Soc. 109 (1949) 600–609.
[3] G.P. Kuiper, in: J.A. Hynek (Ed.), Astrophysics – A Topical Symposium, McGraw-Hill, New York, 1951, pp. 357–424.
[4] F.L. Whipple, Proc. Nat. Acad. Sci. 52 (1964) 565–594.
[5] F.L. Whipple, Proc. Nat. Acad. Sci. 51 (1964) 711–718.
[6] S.E. Hamid, B.G. Marsden, F.L. Whipple, Astron. J. 73 (1968) 727–729.
[7] J.A. Fernández, Mon. Not. R. Astron. Soc. 192 (1980) 481–491.
[8] M. Duncan, T. Quinn, S. Tremaine, Astrophys. J. Lett. 328 (1988) L69–73.
[9] D.J. Jewitt, J.X. Luu, Nature 362 (1993) 730–732.
[10] C.A. Trujillo, J.X. Luu, A.S. Bosh, J.L. Elliot, Astron. J. 122 (2001) 2740–2748.

[11] D.J. Jewitt, J.X. Luu, in: T.W Rettig, J.M. Hahn (Eds.), Completing the Inventory of the Solar System, Astron. Soc. of the Pacific, San Francisco, 1996, pp. 255–258.
[12] R. Malholtra, Astron. J. 110 (1995) 420–429.
[13] S. Ida, J. Larwood, A. Burket, Astrophys. J. 528 (2002) 351–356.
[14] A.L. Cochran, H.F. Levison, S.A. Stern, M.J. Duncan, Astrophys. J. 455 (1995) 342–346.
[15] S.A. Stern, Astron. J. 110 (1995) 856–868.
[16] R. Gomez, Earth, Moon and Planets (2003), in press.
[17] M.E. Brown, C.J. Trujillo, Astron. J. (2003), in press.
[18] N. Thomas, S. Eggers, W.-H. Ip, G. Lichtenberg, A. Fitzsimmons, L. Jorda, H.U. Keller, I.P. Williams, G. Hahn, H. Rauer, Astrophys. J. 534 (2000) 446–455.
[19] D.J. Jewitt, H. Aussel, A. Evans, Nature 411 (2001) 446–447.
[20] S.S. Sheppard, D.J. Jewitt, Astron. J. 124 (2002) 1757–1775.
[21] D.J. Jewitt, S.S. Sheppard, Astron. J. 123 (2002) 2110–2120.
[22] C. Veillet, J. Parker, I. Griffin, B. Marsden, A. Doressoundiram, M. Buie, D.J. Tholen, M. Connelley, M. Holman, Nature 416 (2002) 711–713.
[23] K. Noll, Earth, Moon and Planets (2003), in press.
[24] D. Bockelee-Morvan, E. Lellouch, N. Biver, G. Paubert, J. Bauer, P. Colom, D.C. Lis, Astron. Astrophys. 377 (2001) 343–353.
[25] O.R. Hainaut, C.E. Delahodde, H. Boehnhardt, E. Dotto, M.A. Barrucci, K.J. Meech, J.M. Bauer, R. West, A. Doressoundiram, Astron. Astrophys. 356 (2000) 1076–1088.
[26] T. Sekiguchi, H. Boehnhardt, O.R. Hainaut, C.E. Delahodde, Astron. Astrophys. 385 (2002) 281–288.
[27] F. Roques, M. Moncuquet, Icarus 147 (2000) 530–544.