DARK MATTER IN THE SOLAR SYSTEM

Scott Tremaine
Canadian Institute for Theoretical Astrophysics
University of Toronto
60 St. George St.
Toronto M5S 1A1
Canada

ABSTRACT. There is little direct dynamical evidence for dark matter in the solar system. However, fairly general models for the formation of the planetary system predict that dark matter in the form of fossil planetesimals is likely to be present in two locations: an extended spherical cloud with semi-major axes between 10^3 AU and 4×10^4 AU, and a flat disk in the region of the outer planets or just beyond them. The outer part of the extended cloud is probably the source of new comets (the Oort cloud), and the disk is probably the source of the Jupiter-family comets. The mass in the extended cloud may exceed the mass in the cores of the giant planets; thus most of the metals in the solar system outside the Sun may be in the form of dark matter. I also review the evidence for a tenth planet, Planet X, and for the companion star Nemesis. Neither is very probable, since the observations that they are invoked to explain (residuals in the orbits of Uranus and Neptune, and periodicities in the cratering and extinction records) are only marginally significant. In addition, the proposed objects have unusual properties: Planet X must be at least ten times more distant and massive than any other planet, and Nemesis must have a very improbable orbit. Finally, I summarize dynamical constraints that strongly suggest that the Sun has no companion more massive than about $0.03\,\mathcal{M}_\odot$.

Introduction

Most of us are mainly interested in baryonic dark matter because of the role it plays in galactic structure and cosmology. Thus I begin by describing why dark matter in the solar system is relevant to this broader context.

First, of course, all of the dynamical arguments for dark matter in galaxies are based on the assumption that Newton's laws of motion and law of gravity are correct, and the solar system provides by far the most accurate tests of these laws on large scales.

A second reason is that some forms of baryonic dark matter may be most easily detected in the solar system. For example, if brown dwarfs comprise most of the dark matter in the Galactic disk, and if they form binaries at the same rate as main sequence stars, then there may very likely be one or more brown dwarfs orbiting the Sun.

It seems certain that dark matter in galaxies is intimately related to galaxy formation;

D. Lynden-Bell and G. Gilmore (eds.), Baryonic Dark Matter, 37–65.

similarly, I shall argue that the nature and distribution of dark matter in the solar system provides both insights and puzzles that bear on the formation of the solar system.

Finally, recall that there have been two major dynamical puzzles in the solar system since 1800: the unexplained residuals in the orbit of Uranus that led to the discovery of Neptune, and the anomalous precession of Mercury's perihelion that was explained by general relativity. One puzzle was explained by a new but commonplace object, and the other by radical new physical laws. These two classes of explanation are also seen in one of the main issues addressed by this meeting: will dark matter in galaxies be explained by commonplace objects such as brown dwarfs, or by novel physical laws and exotic new particles?

1. Dark Matter in the Planetary System

Dark matter located between the planets can be detected by its gravitational influence on planetary orbits. The best limits come from Kepler's law, which states that in the absence of dark matter the period P and semi-major axis a of a planet of negligible mass are related by

$$\frac{4\pi^2 a^3}{P^2} = G\mathcal{M}_\odot. \tag{1}$$

The perturbing effects of the planetary masses are straightforward to account for and will not be discussed here.

Distances are measured in terms of the astronomical unit (AU), which is defined by the relation

$$G\mathcal{M}_\odot = (0.01720209895)^2 \frac{\mathrm{AU}^3}{(\mathrm{day})^2}. \tag{2}$$

The constant in the brackets is approximately 2π divided by the number of days in a year, so that the Earth's semi-major axis $a_\oplus$ is very nearly 1 AU.

Direct range measurements by radar and spacecraft tracking give both the value of 1 AU and the semi-major axes to the other planets in centimeters. The standard IAU (1977) value for the astronomical unit is $1\,\mathrm{AU} = 1.49597870 \times 10^{13}\,\mathrm{cm}$.

If there is dark matter in the system, the semi-major axes a_r for other planets that are deduced from range measurements will differ from the semi-major axes a_p deduced from the orbital periods and Kepler's law (1). If for simplicity I assume that the dark matter is distributed spherically, with mass $\Delta M(r)$ interior to radius r, then the apparent mass of the Sun determined from equation (2) and the value of the astronomical unit will be $\mathcal{M}_\odot{}' = \mathcal{M}_\odot + \Delta M(a_\oplus)$. Then the semi-major axis to another planet as deduced from ranging, and the semi-major axis as deduced from Kepler's law, will be given by

$$\frac{4\pi^2 a_r^3}{P^2} = G[\mathcal{M}_\odot + \Delta M(a_r)], \qquad \frac{4\pi^2 a_p^3}{P^2} = G\mathcal{M}_\odot{}', \tag{3}$$

which yields for $\Delta M \ll \mathcal{M}_\odot$

$$\frac{a_r}{a_p} \equiv 1 + \tfrac{1}{3}\epsilon, \qquad \text{where} \qquad \epsilon \simeq \frac{\Delta M(a_r) - \Delta M(a_\oplus)}{\mathcal{M}_\odot}. \tag{4}$$

The values of ϵ determined from least-squares fits to solar system ephemerides are all consistent with zero (Talmadge *et al.* 1988, Anderson *et al.* 1989) and the 1σ upper limits

Table 1 Limits on dark mass in the planetary system

planet	method	distance (AU)	$\epsilon(1\sigma)$
Mercury	radar	0.4	$< 3 \times 10^{-8}$
	Mariner 10		
Venus	radar	0.7	$< 3 \times 10^{-8}$
Mars	Mariner 9	1.9	$< 1 \times 10^{-9}$
	Viking		
Jupiter	Voyager	5.2	$< 2 \times 10^{-7}$
Saturn	Voyager	9.5	–
Uranus	Voyager	19.2	$< 3 \times 10^{-6}$
Neptune	Voyager	30.0	–

NOTES: Data from Talmadge *et al.* (1988) and Anderson *et al.* (1989). Dashes indicate that data exist but are not yet analyzed and published.

are shown in Table 1. These provide approximate limits on the dark mass in solar masses contained between the planets, although of course any dark mass contained inside Mercury's orbit would not be detected.

Limits on dark mass can also be derived by comparing observed perihelion precession rates with those predicted from mutual planetary perturbations and general relativistic effects. This test is sensitive to the gradient $d\Delta M(a)/dr$ at the planetary orbit rather than the difference $\Delta M(a) - \Delta M(a_\oplus)$. For $\Delta M(r) \propto r$ the limits (from the orbits of Mercury and Mars) are somewhat less sensitive than those from ranging (Talmadge *et al.* 1988).

The main known source of dark matter in the planetary system is the asteroid belt. The total mass of the belt is quite uncertain, due to the unknown contribution from faint asteroids, but is at least $2 \times 10^{-9}\, \mathcal{M}_\odot$ (Hughes 1982). The masses of the largest few asteroids can be directly determined from their perturbations on the orbit of Mars (Standish and Hellings 1989).

2. Residuals in the Orbits of Uranus and Neptune

Residuals in the orbit of Uranus of $\sim$ 100" led to the prediction of Neptune by LeVerrier and Adams in 1845. Neptune was discovered in 1846 within 1° of LeVerrier's predicted position.

Remaining residuals in Uranus's orbit of a few arc seconds led to predictions of a tenth planet by Lowell in 1915 and Pickering in 1928. Lowell initiated a systematic survey for a tenth planet which led to the discovery of Pluto by Tombaugh in 1930, within 6° longitude of the predicted positions (see Hoyt 1980 for a history).

The Pluto mass needed to remove the residuals is between $0.5 M_\oplus$ and $5 M_\oplus$ (see Duncombe and Seidelmann 1980 for a review of Pluto mass estimates; $M_\oplus$ = 1 Earth mass= $3.04035 \times 10^{-6}\, \mathcal{M}_\odot$). However, discovery of Pluto's satellite Charon (Christy and Harrington 1978) showed that the mass of the Pluto-Charon system was only $0.002 M_\oplus$, far too small to have any detectable effect on the orbits of Uranus and Neptune. Since Pluto

cannot account for the residuals, the discovery of Pluto close to the predicted position must have been accidental; thus, if the residuals are real, it is likely that some undiscovered dark mass is responsible.

The best available theory for the orbits of Uranus and Neptune is the JPL ephemeris DE 200, which is based on fitting observations to 1978. The residuals between DE 200 and the observations show three distinct anomalies (Seidelmann and Harrington 1988):

(i) Observations of the right ascension of both Uranus and Neptune made since 1978 already deviate systematically from the DE 200 predictions, by 0.5" after only 10 years.

(ii) Uranus observations prior to 1900 cannot be fit in right ascension. It is possible that systematic errors in the early observations are the source of this discrepancy.

(iii) Prediscovery observations of Neptune by Lalande in 1795 and Galileo in 1613 cannot be fit. However, the deviations are only four times the estimated error for the Lalande observations, and the scale of Galileo's drawing showing Neptune is uncertain (Kowal and Drake 1980).

While the residuals are larger than expected from the known observational errors, it is probably prudent to regard them as setting upper limits on, rather than providing firm evidence for, dark matter in the solar system.

Constraints on dark matter are also available from two other types of probe at similar distances (20 to 40 AU).

The Pioneer 10 spacecraft shows no evidence of unmodelled acceleration to a level $|\Delta\ddot{\mathbf{r}}| \lesssim 5 \times 10^{-9}\,\mathrm{cm\ sec^{-2}}$ out to about 35 AU (Anderson and Standish 1986). If I set $|\Delta\ddot{\mathbf{r}}| = G\Delta M/r^2$, $r = 40\,\mathrm{AU}$, then I obtain a crude limit on dark mass

$$\Delta M \lesssim 1.3 \times 10^{-5}\,\mathcal{M}_\odot = 4M_\oplus. \tag{5}$$

In the near future this limit will be improved substantially by the addition of data from Pioneer 11 and Voyager 1 and 2, and the distance will increase to $\sim 50\,\mathrm{AU}$.

Comets are excellent probes of distant dark mass, both because their aphelion distances can exceed those of any planet, and because their eccentricity and inclination are large so that precession is easier to detect. Limits on the anomalous precession of the orbit of Halley's comet since 1835 yield (Hamid, Marsden and Whipple 1968, Yeomans 1986)

$$\Delta M \lesssim 0.3M_\oplus \quad \text{at} \quad 40\,\mathrm{AU}. \tag{6}$$

There are reasons to treat the limits (5) and (6) with caution. They are obtained from the largest extra acceleration that can be added to a fixed solar system model without introducing unacceptable errors in the trajectory of the spacecraft or comet. However, the mass causing this acceleration would also perturb the planet orbits and change the best-fit values for the planetary masses and orbital elements. Thus the models are not self-consistent. A proper approach would require a simultaneous fit to both planetary data and the data from the spacecraft or comet. The limits (5) and (6) could be substantially too low if there is a large covariance between the dark mass and one or more of the other free parameters.

Another concern is that the solution for the orbit of Halley's comet contains free parameters that model the nongravitational acceleration due to mass loss (Yeomans 1986). These parameters may mask acceleration due to dark mass.

Conclusion. The residuals in the orbits of Uranus and Neptune are poorly understood, and it is not clear whether they are large enough to require the presence of dark mass. However, if dark mass is the major cause of the residuals, then the needed amount is at most about $5M_\oplus$ located at about 40 AU (since this is the largest Pluto mass that dynamicists have invoked to explain the residuals). The Pioneer spacecraft gives a similar limit; Halley's comet gives a substantially lower but less rigorous limit.

Since the perturbing effects of a distant mass M at radius r scale as M/r^3, the limit given by residuals in the orbits of the outer planets can be written as

$$\Delta M(r) \lesssim 5M_\oplus \left(\frac{r}{40\,\mathrm{AU}}\right)^3. \tag{7}$$

2.1 PLANET X

The most popular explanation for residuals in the Uranus and Neptune orbits is an undiscovered Planet X that orbits beyond Pluto.

The search that led to the discovery of Pluto still provides the best limits on the brightness of any undiscovered planets (Tombaugh 1961). Tombaugh covered 75% of the sky, mostly to a limiting magnitude $V = 17$, almost two magnitudes fainter than Pluto. Unless Planet X is on a highly inclined orbit and happened to be near the ecliptic pole at the time of the search, then it must be substantially fainter than Pluto.

The apparent brightness of a distant planet is proportional to pR^2/r^4, where p is the planet's albedo, R is its radius and $r \gg 1$ AU is its distance from the Sun. Pluto's geometric albedo is high (0.6; see Tholen *et al.* 1987), and the albedo of Planet X may be lower; since both the albedo and the exact limiting magnitude of Tombaugh's search are uncertain, I will simply assume that Tombaugh's failure to detect Planet X implies that its ratio R^2/r^4 must be smaller than Pluto's. Thus

$$\frac{R_x^2}{r_x^4} < \frac{R_p^2}{r_p^4} \quad \text{or} \quad \frac{M_x}{r_x^6} < \frac{M_p}{r_p^6}, \tag{8}$$

where Pluto's distance and mass are $r_p = 40$ AU and $M_p = 0.002M_\oplus$, and I have assumed that Planet X and Pluto have the same density. If Planet X is to explain the residuals in the orbits of Uranus and Neptune, its mass must be close to the upper limit (7), so I shall write

$$M_x = \eta \times 5M_\oplus \left(\frac{r_x}{40\,\mathrm{AU}}\right)^3, \tag{9}$$

where η cannot be much smaller than unity.

Combining equations (8) and (9), I find

$$r_x \gtrsim 500\eta^{1/3}\,\mathrm{AU}, \qquad M_x \gtrsim 10^4\eta^2 M_\oplus. \tag{10}$$

Thus, even for η as low as 0.3, Planet X must be at least a factor of ten more distant than all the known giant planets and more massive than any of them ($r_x \simeq 300$ AU, $M_x \simeq 10^3M_\oplus = 0.003\mathcal{M}_\odot$; Jupiter's mass is only $314M_\oplus$). In fact, the required mass is so large that the object should probably be regarded as a degenerate dwarf star rather than a planet. A brown dwarf with this mass and distance would probably be visible in the IRAS Point Source Catalog (Chester 1986; the luminosity and effective temperature are given by

Stevenson 1986), but only at high Galactic latitude. More distant brown dwarfs satisfying equation (9) would be even more massive and even brighter in the infrared; they would also violate other dynamical constraints (§5).

There have been several attempts to determine the location of Planet X from the residuals. Since it is probably very distant, its orbital motion is negligible, so all that can be determined is its position (to within a reflection through the Sun) and the ratio M_x/r_x^3. At present there is no consensus on a predicted location (Harrington 1988, Gomes 1989; see Seidelmann and Harrington 1988 for a comparative review).

My own opinion is that the residuals in the orbits of Uranus and Neptune are not due to a tenth planet. Either the observational errors have been underestimated, or the accuracy of the theoretical models has been overestimated, or some unmodelled dynamical effect is responsible.

Our understanding of the Uranus and Neptune residuals will improve in the future. The Voyager encounters have now provided much more accurate masses for both planets as well as direct ranges. More accurate angular positions of the outer planets will become available through ring occultations and Hubble Space Telescope observations of their satellites. There will also be general improvements to solar system ephemerides from millisecond pulsars, VLA observations of satellites, and spacecraft such as Magellan, Galileo, and Cassini (Standish 1986).

3. The Outer Regions of the Solar System

Before discussing specific models for dark matter beyond the planetary system, it is worthwhile to review some general aspects of dynamics in the outer solar system.

3.1 THE ROCHE SURFACE

Tidal forces from the Galaxy establish an outer boundary to the solar system (Antonov and Latyshev 1972). If I approximate the Galaxy as axisymmetric and assume that the Sun travels on a circular orbit in the Galactic plane, then a test particle orbiting in the combined field of the Sun and Galaxy conserves its Jacobi integral (e.g. Binney and Tremaine 1987)

$$E_J = \tfrac{1}{2}\mathbf{v}^2 - \frac{G\mathcal{M}_\odot}{r} - \tfrac{1}{2}|\boldsymbol{\Omega}\times(\mathbf{R}_0+\mathbf{r})|^2 + \Phi_G(\mathbf{R}_0+\mathbf{r}) \equiv \tfrac{1}{2}\mathbf{v}^2 + W(\mathbf{r}). \qquad (11)$$

Here $\mathbf{R}_0$ and $\mathbf{r}$ are the positions of the Sun relative to the Galactic center and the test particle relative to the Sun, Φ_G is the potential due to the Galaxy, and $\boldsymbol{\Omega}$ is the angular velocity of the Sun in its Galactic orbit. The velocity $\mathbf{v}$ is measured in a frame rotating with the Sun at $\boldsymbol{\Omega}$. I shall write $\mathbf{r} = (x, y, z)$ where the x-axis points away from the Galactic center and the z-axis to the Galactic pole. I also write $\Phi_G = \Phi_G(R, z)$ where R measures distance from the Galactic center in cylindrical coordinates. Then

$$\boldsymbol{\Omega} = \Omega(R_0)\mathbf{e}_z, \qquad \text{where} \qquad \Omega^2(R) = \frac{1}{R}\frac{\partial\Phi_G(R,0)}{\partial R}. \qquad (12)$$

Since $\mathbf{v}^2$ is non-negative, the region accessible to a particle with a given Jacobi integral is bounded by the inequality $W(\mathbf{r}) \le E_J$. Expanding $W(\mathbf{r})$ to second order in the small quantity r/R_0 and eliminating $\Omega(R_0)$ using equation (12) yields

$$W(\mathbf{r}) = -\frac{G\mathcal{M}_\odot}{r} + \frac{1}{2}\left[\frac{\partial^2\Phi_G}{\partial R^2} - \frac{1}{R}\frac{\partial\Phi_G}{\partial R}\right]_{(R_0,0)} x^2 + \frac{1}{2}\left.\frac{\partial^2\Phi_G}{\partial z^2}\right|_{(R_0,0)} z^2 + \text{constant}. \qquad (13)$$

The quantity in square brackets is just $Rd\Omega^2(R)/dR$; when evaluated at R_0 this is equal to $-4A(A-B)$ where A and B are the usual Oort constants (e.g. Binney and Tremaine 1987). Also, $\partial^2\Phi_G/\partial z^2$ can be eliminated using Poisson's equation $\nabla^2\Phi_G = 4\pi G\rho(R,z)$, where ρ is the local density of Galactic material. Thus, equation (13) becomes (Antonov and Latyshev 1972, Heisler and Tremaine 1986)

$$W = -\frac{G\mathcal{M}_\odot}{r} - 2A(A-B)x^2 + [2\pi G\rho_0 + A^2 - B^2]z^2 + \text{constant}, \tag{14}$$

where $\rho_0 = \rho(R_0, 0)$.

The last closed surface $W =$ constant surrounding the Sun is known as the Roche surface, $W \equiv W_R$. The Roche surface is a natural measure of the boundary of the solar system in that (i) a particle at rest that is inside the Roche surface will always remain inside; (ii) any particle outside the Roche surface is not prevented from escaping by the Jacobi integral or any other analytic integral of motion. Numerical calculations (Hénon 1970) confirm that the Roche surface provides a qualitative measure of the boundary of a mass subject to tidal forces. Hénon found that the fractional volume of phase space at a given Jacobi integral E_J that was occupied by escape orbits increased rapidly once $E_J > W_R$, although a few retrograde orbits can remain bound at arbitrarily large distances.

The Roche surface intersects the coordinate axes at $\pm x_L$, $\pm y_L$, $\pm z_L$, where

$$\begin{aligned} x_L &= \left[\frac{G\mathcal{M}_\odot}{4A(A-B)}\right]^{1/3}, \quad y_L = \tfrac{2}{3}x_L, \\ \frac{G\mathcal{M}_\odot}{z_L} &- [2\pi G\rho_0 + A^2 - B^2]z_L^2 = \tfrac{3}{2}(G\mathcal{M}_\odot)^{2/3}[4A(A-B)]^{1/3}. \end{aligned} \tag{15}$$

The Roche surface has roughly the shape of a triaxial ellipsoid, except for cusps at the points $(\pm x_L, 0, 0)$, which are saddle points of W (the collinear Lagrange points).

For $\rho_0 = 0.15\,\mathcal{M}_\odot\,\mathrm{pc}^{-3}$ (the mean from Bahcall 1984 and Kuijken and Gilmore 1989), $A = 14.4\,\mathrm{km\,s^{-1}\,kpc^{-1}}$, $B = -12.0\,\mathrm{km\,s^{-1}\,kpc^{-1}}$ (Kerr and Lynden-Bell 1986), the Roche surface crosses the coordinate axes at

$$x_L = 1.41\,\mathrm{pc}, \qquad y_L = 0.94\,\mathrm{pc}, \qquad z_L = 0.67\,\mathrm{pc}.$$

Thus the Roche surface is about $(2-3)\times 10^5\,\mathrm{AU}$ from the Sun. Numerical calculations of the stability of orbits in the Galactic tidal field are roughly consistent with this result, yielding a maximum aphelion distance of about $2\times 10^5\,\mathrm{AU}$ (Smoluchowski and Torbett 1984).

3.2 EVAPORATION

Large orbits are unlikely to survive because they are perturbed onto escape orbits by stochastic gravitational forces from passing stars and giant molecular clouds (GMCs).

The half-life of bound test particles of semi-major axis a in a background of passing stars has often been discussed in the context of survival of binary star systems. Heggie (1975) estimated that half of an ensemble of particles bound to the Sun would escape in a time

$$t_{1/2} = 6\times 10^{10}\,\mathrm{yr}\left(\frac{2\times 10^4\,\mathrm{AU}}{a}\right). \tag{16}$$

Heggie considered only "catastrophic" encounters, that is, encounters strong enough to disrupt the system completely. In fact, weaker "diffusive" encounters that lead to a gradual random walk of the test particle to less tightly bound orbits are more important than catastrophic encounters. Including both types of encounter, Bahcall, Hut and Tremaine (1985) found

$$t_{1/2} = 4.5 \times 10^9 \,\mathrm{yr} \left(\frac{2 \times 10^4 \,\mathrm{AU}}{a} \right), \tag{17}$$

roughly a factor of ten shorter than Heggie's estimate.

These calculations neglect the effects of the Galactic tidal field, which increases the escape rate. Weinberg, Shapiro and Wasserman (1987) account crudely for the tidal field by assuming that the binary system is disrupted when its semi-major axis exceeds 1 pc and find

$$t_{1/2} = 4 \times 10^9 \,\mathrm{yr} \left(\frac{2 \times 10^4 \,\mathrm{AU}}{a} \right)^{1.4} \qquad \text{for } 1.5 \times 10^4 \,\mathrm{AU} \lesssim a \lesssim 8 \times 10^4 \,\mathrm{AU}. \tag{18}$$

The disruptive effects of GMCs are much harder to estimate, both because the parameters of the clouds are poorly known and because the encounters are relatively rare (see Hut and Tremaine 1985 for a discussion of the uncertainties). Using a plausible set of GMC parameters, Weinberg *et al.* (1987) found that perturbations from GMCs and stars led to a half-life for a binary with $a = 2 \times 10^4$ AU that was about a factor of two shorter than the half-life due to perturbations from stars alone.

These results suggest that the half-life of dark matter bound to the Sun is less than the age of the solar system if its initial semi-major axis exceeds about $(1 - 2) \times 10^4$ AU. Once the initial semi-major axis of the dark matter exceeds 5×10^4 AU, the probability that it will still be bound to the Sun is substantially less than 10%. Thus perturbations from stars and GMCs establish an effective outer boundary to the distribution of dark matter that is several times smaller than the Roche distance of $(2 - 3) \times 10^5$ AU.

3.3 THE OUTER BOUNDARY OF THE ECLIPTIC

The orbits of most of the planets lie within a few degrees of a common plane known as the ecliptic, reflecting the formation of the planets from a flat disk of dust and gas. However, at larger distances, the torque exerted by Galactic tides over the lifetime of the solar system is large enough to destroy any disk structure that was originally present.

The strongest component of the Galactic tide is described by the potential (Morris and Muller 1986, Heisler and Tremaine 1986)

$$\Phi_G = 2\pi G \rho_0 z^2 \tag{19}$$

(cf. eq. 14; I use the fact that $A^2, B^2 \ll 2\pi G\rho_0$). This potential exerts a torque $\mathbf{N} = -\mathbf{r} \times \nabla \Phi_G$ which causes the angular momentum vector $\mathbf{L}$ of a particle orbit to precess around the Galactic pole. The Galactic pole is tipped by an angle $i = 60.2°$ from the ecliptic pole. Thus a flat ecliptic disk of collisionless particles on circular orbits is converted into a fattened spheroidal distribution whose symmetry axis is the Galactic pole; the thickness of the distribution can be measured by $\langle z^2 \rangle / (\langle x^2 \rangle + \langle y^2 \rangle) = \sin^2 i / (2 - \sin^2 i) = 0.604$, not far from the value 0.5 expected for a spherical distribution. The time scale for fattening-up or isotropizing the disk is $t_{\mathrm{iso}} \approx |\mathbf{L}|/|\mathbf{N}|$; thus disk structure can only survive out to a radius

$r_{\rm iso}$ where $t_{\rm iso}$ equals the age of the solar system, $t_{\rm ss} = 5 \times 10^9\,{\rm yr}$. To order of magnitude I find

$$r_{\rm iso} \approx \frac{\mathcal{M}_\odot{}^{1/3}}{(2\pi\rho_0 t_{\rm ss})^{2/3} G^{1/3}} = 5 \times 10^3\,{\rm AU}. \tag{20}$$

At larger distances the weaker components of the Galactic tidal field, as well as perturbations from passing stars and GMCs, will isotropize the dark matter distribution even further.

4. A Minimal Theory of Solar System Dark Matter

It is useful to ask whether theories of formation of the solar system naturally give rise to dark matter, and, if so, what its nature and distribution would be.

We believe that the Sun was originally surrounded by a gas disk from which the planets formed. If most of the disk mass in non-volatile material was incorporated into the cores of the giant planets, then the surface density in non-volatiles can be estimated by smearing each giant planet core into an annulus reaching halfway to the next planet (see Figure 1). The resulting distribution can be fit to a power law

$$\Sigma(r) \simeq 1.4\,{\rm g\,cm^{-2}} \left(\frac{10\,{\rm AU}}{r}\right)^2. \tag{21}$$

As the gas disk cools, the non-volatile material condenses into grains that settle into a thin disk. Once enough material is gathered in this disk, it becomes gravitationally unstable. The most unstable wavelength is (Goldreich and Ward 1973, Binney and Tremaine 1987, eq. (6-54); see Safronov 1960 for an early discussion)

$$\lambda_u = \frac{2\pi^2 \Sigma r^3}{\mathcal{M}_\odot}, \tag{22}$$

so the dominant mass of the condensations ("planetesimals") may be written as

$$m_u = \Sigma(f\lambda_u)^2 = 3 \times 10^{20}\,{\rm g} \left(\frac{f}{0.5}\right)^2 \left(\frac{\Sigma}{1\,{\rm g\,cm^{-2}}}\right)^3 \left(\frac{r}{10\,{\rm AU}}\right)^6, \tag{23}$$

where f is a dimensionless number less than unity. A natural choice is $f = 0.5$, since the overdense region is half a wavelength in size. However, f may be much less than unity, if instabilities develop before all the non-volatile material has settled. Notice also that m_u is only weakly dependent on distance from the planet so long as the radial density distribution is close to that of equation (21).

The orbits and sizes of the planetesimals evolve through a variety of processes including collisions, fragmentation, accretion, gas drag and gravitational scattering. Many are eventually incorporated into the cores of the giant planets. Residual planetesimals that have survived to the present time are a possible source of dark matter in the solar system, and in the rest of this section I will discuss dynamical constraints on the distribution of residual material of this kind.

Orbits in the solar system can be grouped into two classes: regular and chaotic (see Lichtenberg and Lieberman 1983). Regular orbits are characterized by linear divergence of

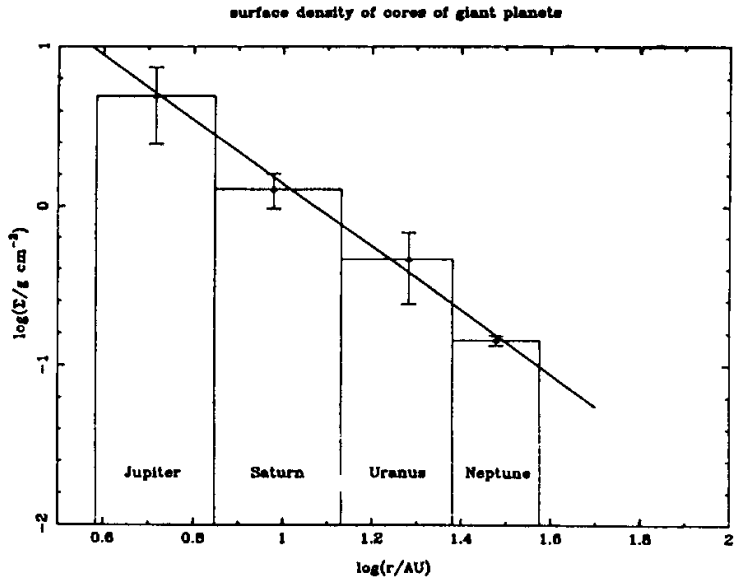

Figure 1. Estimated surface density distribution of non-volatile material in the protoplanetary disk. The masses contained in the rock-ice cores of the giant planets (Stevenson 1982) have been spread uniformly over annuli reaching halfway (logarithmically) to the adjacent planets. The solid line is the power-law fit given in equation (21).

nearby trajectories and discrete power spectra (i.e. they are quasiperiodic). Examples of objects on regular or nearly regular orbits include the planets, their satellites, and most of the asteroids. The evolution of regular orbits can be predicted to high accuracy using the standard methods of celestial mechanics. Chaotic orbits are characterized by exponential divergence of adjacent trajectories and power spectra with a continuous component. The evolution of chaotic orbits is very sensitive to the initial conditions and hence is difficult to predict. For example, planet-crossing orbits are generally chaotic because they can be strongly affected by a chance close encounter with the planet. Pluto's orbit is nearly regular, even though it crosses Neptune's, because the 3:2 Neptune-Pluto resonance prohibits a close encounter. In practical terms, a solar system orbit can only be said to be regular over a given time scale, since it may exhibit weak chaos on a much longer time scale. For example, Sussman and Wisdom (1988) argue that the orbit of Pluto is actually weakly chaotic, even though its orbital elements appear to vary quasiperiodically over time scales as long as several hundred million years.

4.1 PLANETESIMALS ON REGULAR ORBITS

Planetesimals can survive if their orbits are regular or nearly regular over time scales comparable to the age of the solar system, $t_{ss} = 5 \times 10^9$ yr. Unfortunately, the behavior of orbits over these time scales is not well understood. As an example, consider the evolution of a planetesimal in a circular orbit in the ecliptic plane, subjected to perturbations from the planets. What will the orbit look like after a time t_{ss} as a function of the initial radius? Is the orbit stable, in the sense that it remains nearly circular? Will its eccentricity grow, so that it crosses the orbit of an adjacent planet and eventually is ejected by a close encounter?

Or will it, like Pluto, evolve into a planet-crossing orbit that is stabilized by a low-order resonance?

We have only fragmentary answers to these basic questions. Direct numerical integrations show that most near-circular orbits between Jupiter and Saturn are ejected in $\lesssim 10^7$ yr (Franklin, Lecar and Soper 1989). On the other hand, the existence of the asteroid belt strongly suggests that most near-circular orbits with semi-major axes between about 2.2 AU and 3.2 AU are stable over an interval t_{ss}—with the exception of orbits near some low-order resonances with Jupiter (Wisdom 1982, 1983). Investigations using an approximate area-preserving map instead of the accurate equations of motion suggest that most near-circular orbits between the planets are stable over a time t_{ss}; however, ejection occurs for orbits in narrow bands around each planet, for most orbits between Jupiter and Saturn, and for some orbits between Uranus and Neptune (Duncan, Quinn and Tremaine 1989a).

Bands of semi-major axis in which near-circular orbits are stable may host substantial populations of planetesimals. However, no solar system objects other than the planets and their satellites have been found so far on stable orbits, except for the asteroids between Mars and Jupiter; nor is there any dynamically detectable extra mass between the planets (§1). This lack of debris in the planetary system is a remarkable fact that so far has eluded explanation. Perhaps some aspect of the formation process swept the planetesimals out of the system, or perhaps accurate orbit integrations would show that most near-circular orbits are weakly chaotic, so that almost all planetesimals are ejected on 10^9 yr time scales.

An interesting speculation is that the estimate (21) for the density of the protoplanetary disk may be valid to distances much larger than Neptune's semi-major axis of 30 AU. If so, and if I assume that the formation of Neptune depleted the protoplanetary disk only out to about 35 AU, then there should be a residual mass $\Delta M(r) \approx 30 M_\oplus \log(r/r_0)$ inside radius $r > r_0 = 35$ AU, probably in the form of planetesimals. This hypothetical disk was discussed by Kuiper (1951) and is sometimes called the Kuiper belt. It is amusing that the mass $\Delta M(r)$ derived in this way is just below the dynamically determined upper limit (7).

4.1.1 The Kuiper Belt as a Source for Jupiter-family Comets. Indirect evidence for the Kuiper belt is provided by the Jupiter-family comets (comets with orbital period $P < 20$ yr). These have a much flatter distribution of orbits than comets with longer period: their median inclination relative to the ecliptic is only 10° whereas comets with much longer periods have a roughly isotropic distribution of inclinations (Marsden 1983). Also, the Jupiter-family comets are much more numerous than comets with longer period: there are 104 known Jupiter-family comets, compared with only 17 in the period range 20 yr $< P <$ 200 yr (this may however partly be a selection effect, since there are more chances to discover a comet with a shorter period). Numerical simulations show that the inclination distribution and orbital period distribution can only be reproduced if the Jupiter-family comets originate on low-inclination planet-crossing orbits with perihelia in the outer planetary system (Duncan *et al.* 1988, 1989b). Roughly 10–20% of these objects are then scattered by the giant planets to perihelia $q \lesssim 2$ AU where they become visible. It is likely that these comets come from the Kuiper belt, and that they are either the original planetesimals that condensed out of the protoplanetary disk, or possibly fragments of planetesimals broken up by collisions (see Greenberg *et al.* 1984 for a discussion). Thus there are two distinct sources of comets: those with periods $P \lesssim 20$ yr come from the Kuiper belt (Edgeworth 1949, Kuiper 1951,

Whipple 1972), while those with $P \gtrsim 20\,\mathrm{yr}$ come from a separate isotropic source, the Oort cloud (see below).

The lifetimes of planet-crossing orbits are generally short compared with the age of the solar system. Thus Jupiter-family comets must spend most of their lives on nearly regular orbits in the Kuiper belt, and evolve onto chaotic planet-crossing orbits at a slow, steady rate. The mechanism of this process is still uncertain. One possibility is that weak instabilities gradually convert regular near-circular orbits onto chaotic planet-crossing ones; simple dynamical models suggest that such instabilities are likely to be present for near-circular orbits between Uranus and Neptune (Duncan *et al.* 1989a).

4.1.2 Chiron. The object Chiron was discovered by C. Kowal during the Palomar Solar System Survey (Kowal, Liller and Marsden 1979, Kowal 1989). Chiron is unresolved with magnitude $m_{pg} = 18$, suggesting a radius $150\,\mathrm{km}(p/0.05)^{-1/2}$, where p is its geometric albedo. Chiron is the only known minor planet with perihelion well beyond Jupiter's orbit. It orbits between Saturn and Uranus ($a = 13.7\,\mathrm{AU}$, $e = 0.38$, $i = 6.9°$) on a chaotic trajectory that is expected to lead to ejection by Saturn or a close encounter with Jupiter in $10^5 - 10^6\,\mathrm{yr}$ (Oikawa and Everhart 1979). The short lifetime of Chiron's present orbit suggests that it evolved from a more stable orbit in the recent past. It is probable that Chiron spent most of its life on a nearly regular orbit in the Kuiper belt, and that it is simply a large comet following the same evolutionary path that produces the Jupiter-family comets.

4.1.3 Detection of Objects in the Kuiper Belt. The detection of belt objects would offer considerable insight into the formation of the solar system.

To obtain a simple "straw man" model for comparison with the observational limits, I shall assume that the total belt mass is $\Delta M = 1 M_\oplus$, that the density of belt objects is $\rho = 1\,\mathrm{g\,cm^{-3}}$, that the total angular extent of the belt normal to the ecliptic is $2\Delta\theta = 0.2$ radians, and that the belt material is concentrated at a distance $r = 40\,\mathrm{AU}$ from the Sun. I assume that the number of objects in the belt with radii between R and $R + dR$ is $n(R)dR$, where

$$\begin{aligned} n(R) &= \frac{K}{R_0}\left(\frac{R_0}{R}\right)^{b_1+1}, \quad R < R_0, \\ &= \frac{K}{R_0}\left(\frac{R_0}{R}\right)^{b_2+1}, \quad R > R_0. \end{aligned} \tag{24}$$

I set $b_1 = 2$ since this reproduces the size distribution of cometary nuclei (Shoemaker and Wolfe 1982). This index is also consistent with (i) the inferred size distribution of the objects responsible for cratering the Galilean satellites (Shoemaker and Wolfe 1982); (ii) experimental measurements of the size distribution resulting from fragmentation (Hartmann 1969); (iii) the observed size distribution of the particles in Saturn's rings (Cuzzi *et al.* 1984). I shall also take $R_0 = 10\,\mathrm{km}$, since this is close to the maximum observed size of cometary nuclei.

The power-law index b_2 at the high-mass end is quite uncertain. I shall consider two values, $b_2 = 3.5$, roughly the largest value occurring in sources such as gravel, crushed rock, crater fragments, etc. (Hartmann 1969), and $b_2 = 7$, which occurs in a numerical model of planet growth by accretion of planetesimals (Greenberg *et al.* 1984, Figure 12).

One approach is to search optically for slow-moving objects, much as Tombaugh searched for Pluto. The apparent magnitude of a belt object will depend on its optical geometric albedo p, which I take to be 0.05 since albedos in the outer solar system are usually low. The expected number of objects per unit solid angle brighter than visual magnitude V is then

$$\begin{aligned} &0.04\ \mathrm{deg}^{-2} h \left(\frac{R_0}{10\,\mathrm{km}}\right)^{0.5} \left(\frac{p}{0.05}\right)^{1.75} \left(\frac{40\,\mathrm{AU}}{r}\right)^{7} 10^{-0.7(20-V)} \quad \text{for } b_2 = 3.5, \\ &7 \times 10^{-8}\ \mathrm{deg}^{-2} h \left(\frac{R_0}{10\,\mathrm{km}}\right)^{4.5} \left(\frac{p}{0.05}\right)^{3.5} \left(\frac{40\,\mathrm{AU}}{r}\right)^{14} 10^{-1.4(20-V)} \quad \text{for } b_2 = 7, \end{aligned} \tag{25}$$

where $h = (\Delta M/1M_\oplus)(1\,\mathrm{g\,cm^{-3}}/\rho)(0.1\ \mathrm{rad}/\Delta\theta)$.

A number of searches for slow-moving objects in the ecliptic have been carried out, and the limits from these searches are plotted in Figure 2, along with the estimates from equation (25). Tombaugh (1961) carried out a visual search covering 1530 $(\mathrm{deg})^2$ to an approximate limiting magnitude $V = 17.5$. Luu and Jewitt (1988) searched 300 $(\mathrm{deg})^2$ to $V = 20$ using Schmidt plates, and 0.34 $(\mathrm{deg})^2$ using a CCD detector to $V = 24.5$. Kowal's (1989) Solar System Survey covered 6400 $(\mathrm{deg})^2$ to $V = 20$ using Schmidt plates. Some of these magnitude limits may be over-optimistic since faint trailed objects are harder to detect than point sources (e.g. Gehrels 1981). Asteroid surveys also provide useful limits, although it is not certain that asteroid observers would always notice the short trails that an object at 40 AU would produce in a typical exposure. Major surveys include the McDonald Survey (Kuiper *et al.* 1958) of 14,400 $(\mathrm{deg})^2$ to $V = 16$; the Palomar-Leiden Survey (van Houten *et al.* 1970) of 216 $(\mathrm{deg})^2$ to $V = 19.5$; and the Kiso Schmidt Survey (Ishida *et al.* 1984) of 1944 $(\mathrm{deg})^2$ to $V = 18.4$. Additional searches for slow-moving objects are presently being undertaken by M. Duncan and H. Levison using a CCD and by R. Webster, A. Żytkow, and me using Schmidt plates. None of these searches has yielded any slow-moving objects except for Tombaugh's discovery of Pluto and Kowal's discovery of Chiron.

Figure 2 shows that theoretical predictions of the density of Kuiper belt objects are so uncertain that they provide very little guidance. However, it is encouraging that one of our estimates predicts that Kuiper belt objects should already have been discovered. The discovery of even a few belt objects would be so informative that it is worthwhile to strive to improve the observational limits.

Since the optical albedo of belt objects is likely to be low, it is tempting to look for the belt in the infrared. The thermal emission from a black body at 40 AU peaks (i.e. $\nu B_\nu(T)$ is maximized) at $\lambda = 80\mu$. The IRAS satellite detected emission from the ecliptic plane at 60μ and 100μ, but it is difficult to remove the effects of interplanetary dust at smaller radii to measure the contribution from the Kuiper belt (see Jackson and Killen 1988 for a discussion). The expected contribution is also difficult to estimate, since the thermal emission is dominated by very small particles in a belt with $b_1 \geq 2$.

Planned space-based infrared telescopes such as SIRTF could detect individual large bodies in the belt. However, the advantage of higher emissivity is more than offset by the limited aperture, limited telescope time, and limited number of pixels in infrared array detectors.

Belt objects will occasionally occult stars (Bailey 1976, McClintock 1985). The size of the first Fresnel zone in the V band is $\sqrt{\lambda r} = 1.8\,\mathrm{km}(r/40\,\mathrm{AU})^{1/2}$, so that objects smaller than a few kilometers will not produce sharp shadows. For simplicity, let me then

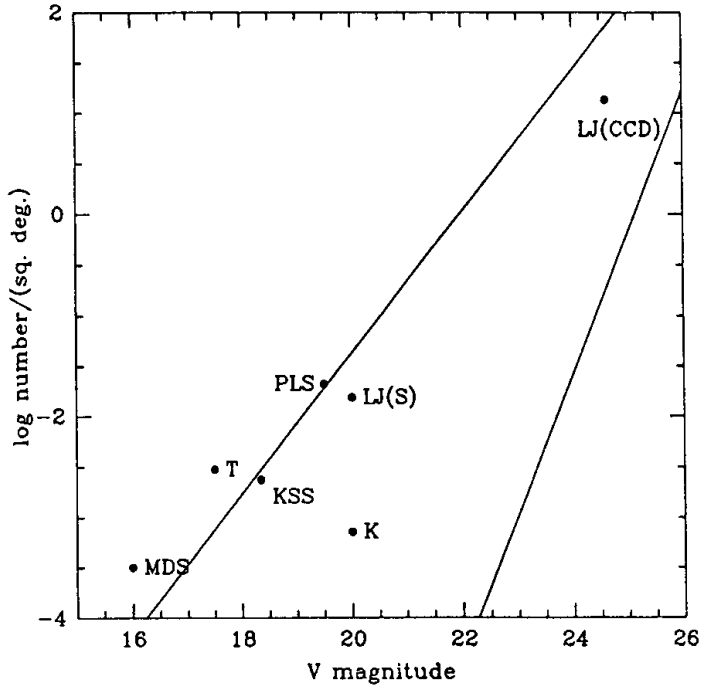

Figure 2. Limits on the number density per square degree of Kuiper belt objects brighter than visual magnitude V. The data points are 99% confidence upper limits. The papers referred to by the symbols are: MDS=Kuiper *et al.* (1958); T=Tombaugh (1961); PLS=van Houten *et al.* (1970); KSS=Ishida *et al.* (1984); LJ=Luu and Jewitt (1988); K=Kowal (1989). The upper and lower lines represent estimates from the top and bottom lines of equation (25).

concentrate on objects with radii $\geq R_0 = 10\,\mathrm{km}$, for which geometrical optics should be a good approximation. Early-type stars are better candidates for occultations than late-type stars, since they have smaller angular diameters for a given apparent magnitude. The angular diameter of the star must be substantially smaller than the angular diameter of the belt object (7×10^{-4} arcsec for a 10 km body at 40 AU); for example, A0 stars satisfy this constraint if their apparent magnitude is substantially fainter than $V = 2$. At opposition, the apparent angular speed of the belt objects mainly reflects the Earth's orbital speed $v_\oplus = 30\,\mathrm{km\ s^{-1}}$; thus the duration of an equatorial occultation is $\Delta t = 2R/v_\oplus = 0.7\ \mathrm{s}(R/10\,\mathrm{km})$. For the assumed size distribution of belt objects (eq. 25), the mean interval between occultations by belt objects larger than R_0 is

$$1.1 \times 10^5\ \mathrm{hr} \left(\frac{1 M_\oplus}{\Delta M}\right) \left(\frac{\rho}{1\,\mathrm{g\,cm^{-3}}}\right) \left(\frac{R_0}{10\,\mathrm{km}}\right)^2 \left(\frac{r}{40\,\mathrm{AU}}\right)^2 \left(\frac{\Delta\theta}{0.1\ \mathrm{rad}}\right), \qquad (26)$$

for both $b_2 = 3.5$ and $b_2 = 7$. The detection rate can be greatly improved by monitoring a field rich in blue stars (e.g. a nearby open cluster) with a CCD, and also by searching for partial occultations by bodies whose size is comparable to the size of the Fresnel zone. Thus the expected occultation rate could be as high as one every few hundred hours. With two telescopes separated by about 1 km, the signature of an occultation would be unmistakable, and the size distribution of the occulting bodies could also be estimated. Thus detection of belt objects by occultations appears to be technically feasible.

4.2 PLANETESIMALS ON CHAOTIC ORBITS

Planetesimals on chaotic orbits can be lost through collision with the Sun or a planet or through escape from the solar system. I will focus here on some interesting aspects of

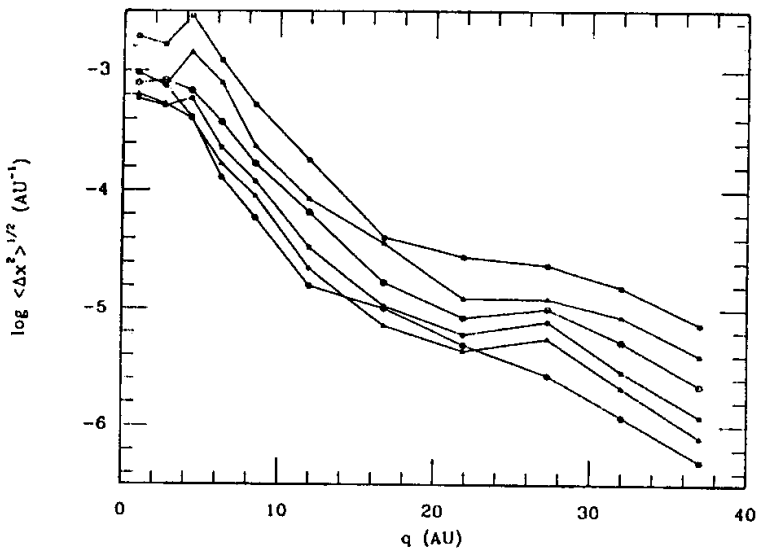

Figure 3. The rms energy change per perihelion passage due to planetary perturbations, as a function of perihelion distance. The light squares, light triangles, light circles, dark squares, dark triangles and dark circles correspond to inclination ranges relative to the ecliptic of $0° - 30°$, $30° - 60°, \ldots, 150° - 180°$. Each point is determined from 3000 passages of parabolic orbits with random argument of perihelion. From Duncan *et al.* (1987), following a figure in Fernández (1981).

the escape process, partly because they have direct observational consequences and partly because the escape process is relatively well-understood. The following arguments are mostly taken from Duncan *et al.* (1987).

In most cases, escape does not result from a single close encounter with a planet, but rather occurs through a gradual random walk or diffusion to less and less tightly bound orbits (this is also true for escape of stars from star clusters and is a consequence of the long range of the gravitational force; see Binney and Tremaine 1987, §8.4.1). During this random walk the perihelion of the planetesimal remains fairly constant, so that the orbit becomes more and more eccentric as it becomes less tightly bound. The planetesimal receives an energy kick each time it passes through the planetary system near perihelion; the orbit is chaotic because each kick depends on the detailed configuration of the planets at that time and hence is effectively random.

For convenience I shall parameterize the energy using the variable $x \equiv 1/a$, where a is the semi-major axis (the usual energy is $E = -\frac{1}{2}G\mathcal{M}_\odot m/a$). Let $\langle(\Delta x)^2\rangle^{1/2} \equiv D_x$ be the rms change in x per perihelion passage (the "diffusion coefficient"); D_x is a function of the perihelion distance q and inclination i (see Figure 3) but is almost independent of a since all high-eccentricity orbits are nearly parabolic near perihelion. The diffusion coefficient decreases by a factor of 100 between $q \lesssim 5\,\mathrm{AU}$ and $q \simeq 20\,\mathrm{AU}$, reflecting the fact that Jupiter dominates the perturbations of orbits with $q \lesssim 5\,\mathrm{AU}$ while the smaller planets Uranus and Neptune dominate for $q \gtrsim 20\,\mathrm{AU}$.

The characteristic time scale t_d for evolution of the semi-major axis is related to the square of the rms energy change per orbit, as is usual for a random walk:

$$t_d = P\frac{x^2}{D_x^2} = 1 \times 10^6\,\mathrm{yr}\left(\frac{10^4\,\mathrm{AU}}{a}\right)^{1/2}\left(\frac{10^{-4}\,\mathrm{AU}^{-1}}{D_x}\right)^2, \quad (27)$$

where P is the orbital period (eq. 1). The diffusion time is generally much less than the age of the solar system, and thus almost all planetesimals on chaotic planet-crossing orbits should reach escape energy long before the present epoch.

However, diffusion toward zero energy does not lead inevitably to escape. The torque from the Galactic potential (19) changes the angular momentum L and thus the perihelion q of the planetesimal orbits. Once the perihelion exceeds $q_{\max} \simeq 35\,\mathrm{AU}$ (i.e. once there are no longer close encounters with the planets) planetary perturbations become ineffective, energy diffusion stops, and the orbital energy is frozen. Since $L = \sqrt{2G\mathcal{M}_\odot q}$ for highly eccentric orbits and $dL/dt \approx 4\pi G\rho_0 a^2$, the characteristic time for the perihelion to reach $q_{\max}$ is $t(q_{\max})$, where

$$t(q) = \frac{(2G\mathcal{M}_\odot q)^{1/2}}{4\pi G\rho_0 a^2} = 6 \times 10^7\,\mathrm{yr}\left(\frac{10^4\,\mathrm{AU}}{a}\right)^2. \quad (28)$$

The energy is frozen in at the semi-major axis a_f where $t(q_{\max})$ first becomes smaller than the diffusion time t_d,

$$a_f = 1.5 \times 10^5\,\mathrm{AU}\left(\frac{D_x}{10^{-4}\,\mathrm{AU}^{-1}}\right)^{4/3}. \quad (29)$$

Many planetesimals escape on the next orbit once the energy $x \lesssim D_x$. Thus most planetesimals on highly eccentric orbits will eventually escape if $x_f = a_f^{-1} \lesssim D_x$; otherwise Galactic tides will remove most of them from planet-crossing orbits before they escape. The criterion that most orbits do not escape is therefore

$$D_x \lesssim 3 \times 10^{-5}\,\mathrm{AU}^{-1}, \quad (30)$$

which holds for $q \gtrsim 15\,\mathrm{AU}$ (Figure 3).

These results show that planetesimals on chaotic, highly eccentric orbits with perihelia in the Jupiter-Saturn region mostly escape, in a time much less than the age of the solar system. However, most planetesimals on orbits with initial perihelia in the Uranus-Neptune zone will remain bound to the solar system, on orbits with semi-major axes $\approx (5-10) \times 10^3\,\mathrm{AU}$ (based on equation 29, with the diffusion coefficient $D_x = 10^{-5}\,\mathrm{AU}^{-1}$, which Figure 3 shows is typical for orbits with random inclinations and $20\,\mathrm{AU} \lesssim q \lesssim 30\,\mathrm{AU}$).

Thus it is likely that the solar system is surrounded by a cloud of planetesimals at semi-major axes $a \approx (5-10) \times 10^3\,\mathrm{AU}$. The cloud is slowly disrupted by encounters with passing stars and GMCs (eq. 18) and by the removal of planetesimals that happen to re-enter the planetary system, but neither of these processes will disrupt the cloud by the present time.

These crude arguments are confirmed by Monte Carlo simulations of the evolution of eccentric orbits subject to planetary perturbations, the Galactic tide, and perturbations from passing stars (Duncan *et al.* 1987). In particular, the simulations confirm that:

(i) The survival of planetesimals on very eccentric orbits depends strongly on their initial perihelion distance q. Less than 6% of planetesimals with $q < 10\,\mathrm{AU}$ remain bound

to the solar system after a time $t_{ss} = 5 \times 10^9$ yr, but 30 – 40% of planetesimals with perihelia in the Uranus-Neptune region (20 AU $< q <$ 30 AU) are still bound after this time. The survival probability may be decreased by the uncertain effects of GMCs.

(ii) The surviving planetesimals are distributed in an extended cloud that surrounds the Sun, with a median semi-major axis 5×10^3 AU. More than 90% of the planetesimals in the cloud have semi-major axes between 1×10^3 AU and 4×10^4 AU.

(iii) The planetesimal cloud is roughly spherical for semi-major axes $\gtrsim 5 \times 10^3$ AU and is flattened towards the ecliptic at smaller semi-major axes, consistent with equation (20). The planetesimals are uniformly distributed over each energy hypersurface in phase space (i.e. the eccentricity distribution is uniform in e^2).

The formation of this planetesimal cloud is a natural consequence of any model for the formation of the solar system in which the giant planet cores are formed by accreting planetesimals. The distribution of orbits of planetesimals in the cloud is determined by the interplay between the Galactic tide and perturbations from the giant planets. Although the orbital distribution is straightforward to predict, at present we cannot reliably predict the total number or total mass of planetesimals in the cloud, since both the distribution of planetesimal masses and the fraction of planetesimals that evolve to chaotic planet-crossing orbits are unknown.

4.3 COMETS

The most striking feature of the distribution of comet orbits is a sharp peak in the distribution of the energy $x = 1/a$ near zero energy. The peak is centered near $x = x_c \equiv 5 \times 10^{-5}\,\mathrm{AU}^{-1}$, with a width $\pm 5 \times 10^{-5}\,\mathrm{AU}^{-1}$ (see Fernández 1985 and Oort 1986 for reviews and Marsden, Sekanina and Everhart 1978, Everhart and Marsden 1983 for the data; here a refers to the "original" semi-major axis that the comet had before entering the planetary system). The existence of this peak led Oort (1950) to propose that the solar system was surrounded by a spherical cloud of comets with typical semi-major axis $x_c^{-1} = 2 \times 10^4$ AU. Oort pointed out that any comets that we see have sufficiently small perihelion ($q \lesssim 2$ AU) that they receive an rms energy impulse $\langle (\Delta x)^2 \rangle^{1/2} \simeq 10^{-3}\,\mathrm{AU}^{-1}$ (see Figure 3) as they pass through the planetary system; since this is much larger than x_c the comets will not return to the Oort cloud but will either escape (if $\Delta x < 0$) or return on a much more tightly bound orbit (if $\Delta x > 0$). Thus we require a flux of fresh comets from the cloud to resupply the peak; this is provided by stellar perturbations and the Galactic tide, which continually change the perihelion distances of comets in the cloud.

There is every reason to believe that Oort's comet cloud is the same as the planetesimal cloud that was derived on dynamical grounds in the previous subsection.

If this belief is correct, we must explain why the typical semi-major axis of comets in the observed peak is $x_c^{-1} = 2 \times 10^4$ AU, whereas the median semi-major axis in the planetesimal cloud was predicted to have the smaller value 5×10^3 AU. A comet is only visible if its present perihelion $q \lesssim 2$ AU; however, for any perihelion $\lesssim 15$ AU the diffusion coefficient is so large (Figure 3) that diffusion in energy is more rapid than the rate of change of perihelion (t_d from equation 25 is shorter than $t(q = 15\,\mathrm{AU})$ from equation 28). Thus objects from the planetesimal cloud diffuse in energy at roughly constant perihelion once $q \lesssim 15$ AU. Comets can therefore only be detected near energy x_c if their *present* perihelion is $\lesssim 2$ AU but their *last* perihelion was $\gtrsim 15$ AU; this requires that $t(q = 15\,\mathrm{AU})$ from equation (28) is less than the orbital period P (eq. 1). In other words, planetesimals only

reach $q \lesssim 2\,\mathrm{AU}$ with the semi-major axes that they had in the cloud if that semi-major axis exceeds (Heisler and Tremaine 1986, Morris and Muller 1986)

$$a_0 = \left[\frac{(15\,\mathrm{AU})\,\mathcal{M}_\odot{}^2}{2^5\pi^4\rho_0^2}\right]^{1/7} = 2.9 \times 10^4\,\mathrm{AU}, \tag{31}$$

which is in adequate agreement with the location of the observed peak of $2 \times 10^4\,\mathrm{AU}$.

There are several interesting consequences of the identification of the Oort comet cloud with the theoretical planetesimal cloud:

i The theoretical arguments predict the radial distribution of comets, which is not directly accessible to observation since the only detectable comets are those from the outermost part of the cloud, $a \gtrsim 2 \times 10^4\,\mathrm{AU}$. Hills (1981) already noted the coincidence of a_0 with the observed semi-major axes of comets from the Oort cloud and suggested that most of the cloud could be hidden at smaller semi-major axes; however, Hills estimated that the hidden inner cloud ($a < 2 \times 10^4\,\mathrm{AU}$) could contain several hundred times as many comets as the observable cloud ($a > 2 \times 10^4\,\mathrm{AU}$), whereas the numerical simulations of the formation of the planetesimal cloud yield a smaller ratio, about five (Duncan *et al.* 1987).

ii The result implies that comets formed in the protoplanetary disk, either through gravitational instability (eq. 23) or by fragmentation of larger bodies. Moreover, the rate of discovery of new comets from the Oort cloud permits us to estimate the mass of the planetesimal cloud. For this we need three independent numbers: the rate at which new comets brighter than some limiting flux pass through a perihelion $< 1\,\mathrm{AU}$ (4 per year brighter than magnitude $H = 10$ following Everhart 1967 and Weissman 1983); the number of new comets per year passing through a perihelion $< 1\,\mathrm{AU}$ per comet in the cloud (1.1×10^{-12} from simulations by Heisler 1989, assuming that we are not now in a comet shower [see below]); and the total mass in comets per comet brighter than $H = 10$ ($1.2 \times 10^{17}\,\mathrm{g}$ following Weissman 1986). Thus we arrive at a total cloud mass of $70 M_\oplus$, with an uncertainty of at least a factor of three. This is almost the same as the total mass in the giant planet cores ($\simeq 75 M_\oplus$ from Stevenson 1982); in other words, of order half of the mass in metals outside the Sun is likely to be dark.

iii The flux of comets reaching the Earth will not be constant in time, since a close or slow encounter with a passing star will shake the cloud strongly enough that comets from the inner part of the cloud may be thrown onto orbits with perihelion $< 1\,\mathrm{AU}$. For $1 - 2$ Myr after such an encounter, the flux of comets reaching the Earth may increase by a factor of 20 or so (Figure 4). These comet "showers" (Hills 1981) occur every 50 Myr or so. It has been suggested that multiple comet impacts on the Earth during such a shower may cause substantial environmental stress and lead to mass extinctions, including the extinction of the dinosaurs at the Cretaceous-Tertiary boundary 65 Myr ago (Hut *et al.* 1987; see van den Bergh 1989 for a general review).

iv The size and mass of the comet cloud depends strongly on the orbits and masses of the planets, and the strength of the Galactic tide. Thus the comet clouds surrounding other stars are probably very different from the Oort cloud.

5. Has the Sun a Companion Star?

Many stars are members of binary systems, and it is possible that the Sun has a distant

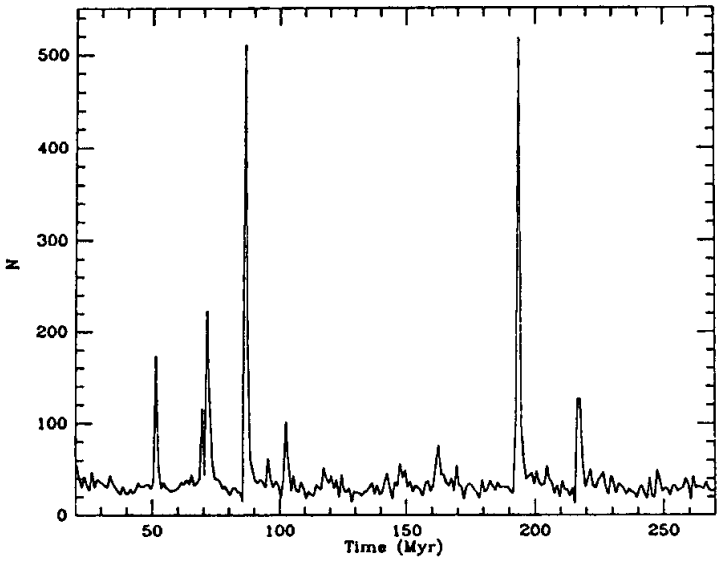

Figure 4. A Monte Carlo simulation of the flux of new comets over a 250 Myr interval. The simulation includes the effects of perturbations from the Galactic tide and passing stars, but not GMCs. The graph shows the total number of comets passing through perihelion with $q < 2\,\mathrm{AU}$ in each 1 Myr interval, from an initial population of 1.44×10^7 comets. The initial distribution of semi-major axes of comets in the Oort cloud is taken from Duncan *et al.* (1987), and the simulation uses a modified version of the code described by Heisler, Tremaine and Alcock (1987). From Heisler (1989).

undiscovered companion, perhaps a neutron star, brown dwarf, or black hole. Optical and infrared searches have not yet revealed any candidate companions (Perlmutter *et al.* 1986, Chester 1986). A hypothetical distant companion called "Nemesis" has been invoked to explain possible periodicities in the record of mass extinctions and impact craters on Earth (see below). Here I review the dynamical limits on the mass M_x and distance r_x of a possible companion. Many of the limits are crude and the discussion here is only intended to produce order-of-magnitude estimates rather than reliable dynamical bounds.

All of the limits derived below are shown on Figure 5.

Constraints from Celestial Mechanics. Residuals in the orbits of the outer planets give the limit (7), which can be rewritten as

$$M_x \lesssim 0.2\,\mathcal{M}_\odot \left(\frac{r_x}{10^3\,\mathrm{AU}}\right)^3. \tag{32}$$

The *Voyager* range to Uranus (Table 1) implies that the mass of any companion inside the orbit of Uranus is

$$M_x \lesssim 3 \times 10^{-6}\,\mathcal{M}_\odot, \qquad r_x \lesssim 20\,\mathrm{AU}. \tag{33}$$

The binary pulsar PSR 1913+16 provides another limit. The orbital period of the system changes at a rate $\dot{P}/P = -8.6077 \times 10^{-17}\ \mathrm{s}^{-1}$, which is 1.010 ± 0.011 times the rate predicted by general relativity (Taylor and Weisberg 1989). The Sun would accelerate toward a companion at a rate $a = GM_x/r_x^2$, which would change $\dot{P}$ by an amount $\delta\dot{P} = -Pa\cos\phi/c$,

where ϕ is the angle between the binary pulsar and the solar companion as seen from the Sun. I replace $|\cos\phi|$ by 0.5, which should be a typical value, and assume that general relativity is correct so that $|\delta\dot{P}/P| < 0.02$ from the timing data, to obtain the limit

$$M_x \lesssim 0.17\,\mathcal{M}_\odot \left(\frac{r_x}{10^3\,\mathrm{AU}}\right)^2. \tag{34}$$

Constraints from Orbital History. These limits are based on the probable evolution of the companion orbit and assume that the companion has been present since the origin of the solar system 5×10^9 yr ($\equiv t_{ss}$) ago.

As shown in §3.2, it is unlikely that a companion can survive perturbations from passing stars and GMCs for a time t_{ss} unless its semi-major axis is $\lesssim 5 \times 10^4$ AU. Thus I assume

$$r_x \lesssim 5 \times 10^4\,\mathrm{AU}. \tag{35}$$

To obtain another constraint, consider a companion whose radius r_x exceeds the distance at which orbits are isotropized by the Galactic tide, $\approx 5\times10^3$ AU (eq. 20). Its orbital angular momentum vector $\mathbf{L}_x$ will not generally be aligned with that of the planetary system $\mathbf{L}_p$. Moreover, $|\mathbf{L}_x|$ will exceed $|\mathbf{L}_p|$, at least for $M_x \gtrsim 10^{-4}\,\mathcal{M}_\odot$, and hence $\mathbf{L}_p$ will precess around $\mathbf{L}_x$ due to the torque exerted on the planets by the companion. The characteristic time in which the direction of $\mathbf{L}_p$ will change by one radian is

$$t_x \approx \frac{4r_x^3}{GM_x}\frac{\sum_i m_i(G\,\mathcal{M}_\odot r_i)^{1/2}}{\sum_i m_i r_i^2} = 2 \times 10^{10}\,\mathrm{yr}\left(\frac{r_x}{10^4\,\mathrm{AU}}\right)^3\left(\frac{\mathcal{M}_\odot}{M_x}\right), \tag{36}$$

where the factor 4 accounts crudely for projection effects at a typical orientation of $\mathbf{L}_x$ relative to $\mathbf{L}_p$, and the sum is over the orbital radii r_i and masses m_i of the giant planets.

Mutual torques between the planets are strong enough that their orbits will precess together so long as $t_x \gtrsim 1 \times 10^6$ yr. Thus it is not surprising that the planetary orbits lie close to a common plane. However, the solar quadrupole moment is so small that its spin angular momentum is almost unaffected by planetary torques. (The precession time for the solar spin due to planetary torques is $10^{10} - 10^{11}$ yr.) Thus the fact that the solar spin axis lies only 7° from the ecliptic strongly suggests that the orientation of the ecliptic has not changed by more than about 0.1 radian since the formation of the solar system. This implies that $t_{ss} \lesssim 0.1t_x$, or,

$$M_x \lesssim 0.4\,\mathcal{M}_\odot \left(\frac{r_x}{10^4\,\mathrm{AU}}\right)^3 \qquad \text{for } r_x \gtrsim 5 \times 10^3\,\mathrm{AU}. \tag{37}$$

Another constraint (Hills 1985) is that no companion more massive than about $M_x = 0.02\,\mathcal{M}_\odot$ could have passed within about 30 AU of the Sun without imparting excessive eccentricities and inclinations to the planetary orbits. The probability that a companion has had no perihelion passage with $q \lesssim 30$ AU is $\exp(-\psi)$, where

$$\psi(q) = \frac{2q}{a}\frac{t_{ss}}{\max[P, t(q)]}. \tag{38}$$

The factor $2q/a$ is the fractional area of the constant energy surface in phase space that is occupied by orbits with perihelion $< q$, and appears because the Galactic tide and stellar

perturbations cause the companion orbit to fill the energy surface uniformly. The second factor involves the orbital period P (eq. 1) and the time $t(q)$ required for the perihelion to change by q (eq. 28). If $t(q) < P$ then consecutive perihelion passages will have perihelion distances that differ by $\gtrsim q$, so the number of independent chances the companion has to hit the target area with perihelion $< q$ is just t_{ss}/P; on the other hand, if $t(q) > P$ there will be of order $t(q)/P$ successive perihelion passages with perihelion distance $< q$, so the number of independent chances is reduced to $t_{ss}/t(q)$.

Equation (38) yields

$$\psi = \min\left[30\left(\frac{10^4\,\mathrm{AU}}{a}\right)^{5/2}, 0.5\left(\frac{a}{10^4\,\mathrm{AU}}\right)\right]. \tag{39}$$

Thus the chance that the companion has passed within 30 AU is large ($\psi > 1$) for $2 \times 10^4\,\mathrm{AU} \lesssim a \lesssim 4 \times 10^4\,\mathrm{AU}$, and I arrive at the constraint (Hills 1985)

$$M_x \lesssim 0.02\,\mathcal{M}_\odot \qquad \text{for} \qquad 2 \times 10^4\,\mathrm{AU} \lesssim r_x \lesssim 4 \times 10^4\,\mathrm{AU}. \tag{40}$$

Constraints from Comets. The model for the formation and evolution of the comet cloud that I described in §4.2 fits the cometary orbit distribution quite well. Any companion must be small enough that this agreement is preserved.

For example, the observed distribution of inverse semi-major axes in the Oort cloud peaks within $x_c = 5 \times 10^{-4}\,\mathrm{AU}^{-1}$ of zero energy (energy measured in units of $-G\,\mathcal{M}_\odot/2$), in good agreement with the dynamical prediction (31). Any companion with $r_x \lesssim x_c^{-1}$ contributes $-GM_x/r_x$ to the measured energy of the comets and hence destroys this agreement unless (Kirk 1978)

$$M_x \lesssim \tfrac{1}{2}\,\mathcal{M}_\odot r_x x_c = 0.025\,\mathcal{M}_\odot\left(\frac{r_x}{10^3\,\mathrm{AU}}\right). \tag{41}$$

I have argued that new comets with $a \lesssim 10^4\,\mathrm{AU}$ are not seen because the time $t(q = 15\,\mathrm{AU})$ required for their perihelia to change by 15 AU exceeds the orbital period P (eq. 31). A companion star exerts an additional torque on the comet orbit given approximately by $N_c = fGM_x \min(a^2/r_x^3, r_x^2/a^3)$ where I take $f \simeq 0.3$ to account for projection effects. If the companion is massive enough, this torque will exceed the torque from Galactic tides, and the time $t(q)$ in equation (28) must be replaced by $t_c(q) = (2G\,\mathcal{M}_\odot q)^{1/2}/N_c$. Then the condition that the comets with $a = (5 - 10) \times 10^3\,\mathrm{AU}$ do not contribute to the flux of new comets is that $t_c(q = 15\,\mathrm{AU}) > P$ in this range of semi-major axis. The resulting constraints on the companion mass are

$$\begin{aligned} M_x \lesssim & 0.04\,\mathcal{M}_\odot\left(\frac{5 \times 10^3\,\mathrm{AU}}{r_x}\right)^2 && r_x \lesssim 5 \times 10^3\,\mathrm{AU}, \\ \lesssim & 0.03\,\mathcal{M}_\odot\left(\frac{1 \times 10^4\,\mathrm{AU}}{r_x}\right)^{1/2} && 5 \times 10^3\,\mathrm{AU} \lesssim r_x \lesssim 10^4\,\mathrm{AU}, \\ \lesssim & 0.03\,\mathcal{M}_\odot\left(\frac{r_x}{1 \times 10^4\,\mathrm{AU}}\right)^3 && 10^4\,\mathrm{AU} \lesssim r_x. \end{aligned} \tag{42}$$

The combination of all these limits (Figure 5) shows that we do not expect the Sun to have any companion of mass $M_x \gtrsim 0.1\,\mathcal{M}_\odot$ and that for most of the possible range of

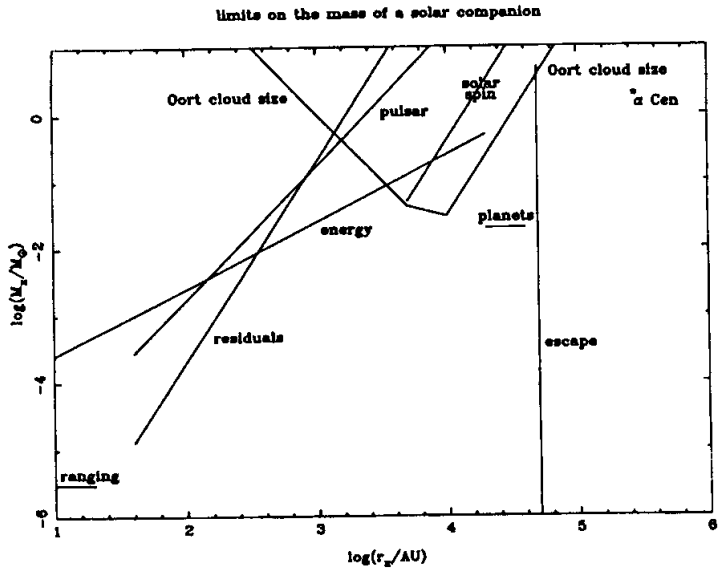

Figure 5. Upper limits on the mass M_x of a solar companion star at radius r_x. The limits come from residuals in the orbits of the outer planets (eq. 32, label "residuals"), ranging to Uranus (eq. 33, label "ranging"), timing the binary pulsar (eq. 34, label "pulsar"), evaporation by passing stars (eq. 35, label "escape"), the alignment of the solar equator with the ecliptic (eq. 37, label "solar spin"), the small eccentricities of the planets (eq. 40, label "planets"), the typical semi-major axes of Oort cloud comets (eq. 41, label "energy"), and the non-observance of new comets with $a \lesssim 10^4$ AU (eq. 42, label "Oort cloud size"). The nearest star, α Cen, is marked for comparison.

distances the limit is closer to $0.03\,\mathcal{M}_\odot$. Thus it is unlikely that the companion is a neutron star (minimum mass $\simeq 0.09\,\mathcal{M}_\odot$, Shapiro and Teukolsky 1983), or a hydrogen-burning main sequence star (minimum mass $\simeq 0.08\,\mathcal{M}_\odot$, D'Antona and Mazzitelli 1985). The only remaining possibilities are black holes, brown dwarfs ($M \gtrsim 0.001\,\mathcal{M}_\odot$, gravitational forces balanced by degeneracy pressure) or planets ($M \lesssim 0.001\,\mathcal{M}_\odot$, Coulomb attraction balanced by degeneracy pressure).

5.1 NEMESIS

Raup and Sepkoski (1984) pointed out that the record of mass extinctions over the past 250 Myr showed a possible periodicity at a period of 26 Myr. This result led Alvarez and Muller (1984) to suggest that a similar period, 28.4 Myr, was present in the record of major impact craters. Both these periodicities may be due to a single astronomical phenomenon: periodic showers of comets. In a strong shower some comets would strike the Earth, leaving large craters and causing substantial environmental stress that may lead to widespread extinctions.

It is difficult to find a mechanism that produces comet showers with this period. The most interesting suggestion is that the Sun has a companion star, "Nemesis", with a semi-major axis of 9.2×10^4 AU (so that the orbital period $P = 28$ Myr), and that Nemesis plunges through the comet cloud at perihelion and thereby triggers a shower (Davis, Hut and Muller 1984, Whitmire and Jackson 1984; see Shoemaker and Wolfe 1986 or Tremaine 1986 for reviews of this and other suggestions).

Unfortunately Nemesis is very vulnerable to perturbations from passing stars. Its semi-major axis and period can change significantly within the duration of the cratering and extinction records (250 Myr), which would erase any detectable periodicity. The probability that the period wanders by less than about 15% during this time (about the maximum allowable) is only $p_1 \simeq 0.14$ (Weinberg *et al.* 1987). In addition, the half-life of a binary with this semi-major axis is only about $t_{1/2} = 5 \times 10^8$ yr (Weinberg *et al.* 1987, without GMCs; with GMCs included the lifetime is even shorter); thus the *a priori* probability that Nemesis was discovered at this special time (i.e. just before it escapes) is $p_2 = t_{1/2}/t_{ss} = 0.1$. The joint probability $p_1 p_2 = 0.014$ is small enough that the Nemesis hypothesis is very unlikely. Weissman (1985) and Shoemaker and Wolfe (1986) reach similar conclusions.

A more prosaic explanation is that the statistical evidence for periodicity is misleading. Many authors have commented on this issue (e.g. Kitchell and Pena 1984, Hoffman 1985, Shoemaker and Wolfe 1986, Grieve *et al.* 1987, Heisler and Tremaine 1989) so here I will mention only two points, relating to the cratering and extinction record respectively.

(i) The statistical significance of the periodicity in the cratering record is based on the null hypothesis that craters are independent events described by a uniform Poisson process. Figure 4 shows that this is not so: a substantial fraction of the total comet flux arrives in brief, intense bursts or showers. The resulting distribution of crater ages may show correlations which could lend spurious extra significance to periodic models.

(ii) The original list of 12 extinction events in Raup and Sepkoski (1984) was revised to a list of 8 events by Sepkoski and Raup (1986). Both the original and revised list were claimed to be periodic at the 99% confidence level. Some problems with the statistical analysis in these two papers were pointed out by Tremaine (1986); after correcting these problems Raup and Sepkoski (1986) concluded that their list of 8 events did not show significant periodicity. However, they then revised the timings of two of the events to obtain a new list of extinction events which was once again significant, this time at the 99.9% level. I cannot comment on the paleontological plausibility of these revisions, but any revisions occurring after the original periodicity hypothesis was introduced may be unintentionally and subjectively biased. This is a well-known problem in psychology experiments, where considerable effort is expended to minimize or eliminate such bias by using blind observers; unfortunately, similar precautions are difficult to apply in this case.

6. Summary

There is little or no direct dynamical evidence for substantial quantities of dark matter in the solar system. However, the observational and theoretical constraints on the amount and distribution of dark matter lead to several interesting conclusions:

(i) The upper limit on dark mass in the region of the terrestrial planets is $\lesssim 3 \times 10^{-8}\,\mathcal{M}_\odot$; within the orbit of Uranus the limit is $\lesssim 3 \times 10^{-6}\,\mathcal{M}_\odot$. On the other hand, we believe that the planets were assembled from billions of planetesimals that condensed out of the gaseous protoplanetary disk. Then why is the planetary system so clean? Was planet formation so efficient that virtually every planetesimal was incorporated into a planet or ejected from the system? Did something sweep the residual planetesimals

out of the system? Or are most orbits in the planetary system weakly chaotic, so that residual planetesimals are ejected on $\lesssim 10^9$ yr time scales?

(ii) The orbits of Uranus and Neptune show unexplained residuals from the best available theoretical models, which may be due to dark matter. Spacecraft and comet trajectories show no unexplained anomalies at a comparable level of accuracy, so the residuals may arise simply from an underestimate of the systematic observational errors. If the residuals are due to an undiscovered Planet X, it must be at least ten times more distant and massive than any known planet, and is more properly called a brown dwarf rather than a planet.

(iii) The outer boundary of the solar system is set by the tidal field of the Galaxy at about 2×10^5 AU from the Sun. However, orbits with semi-major axes exceeding about 5×10^4 AU are unlikely to survive perturbations from passing stars and GMCs for the lifetime of the solar system. Orbits exceeding 5×10^3 AU in size will typically have a random orientation and will not lie in the ecliptic.

(iv) Planetesimals that are perturbed onto Uranus- and Neptune-crossing orbits will naturally evolve into a cloud surrounding the Sun at distances of 10^3 AU to 4×10^4 AU. Objects from the outer part of the cloud can be identified with "new" comets. Thus the existence and properties of the Oort comet cloud follow naturally from fairly general models for the formation of the planetary system.

(v) The mass of the planetesimal cloud may be $100 M_\oplus$ or more; thus the dark mass in the cloud may exceed the total mass in the cores of the giant planets.

(vi) The Jupiter-family comets probably arise from a disk-like source in the outer planetary system. Objects in this source (the Kuiper belt) could be detected by systematic occultation surveys, and may be visible if the mass function decays slowly at the high-mass end. Chiron may be a member of the Kuiper belt that has wandered into the region of the giant planets.

(vii) The hypothetical solar companion Nemesis probably does not exist since its orbit is improbable and the evidence for periodicity in the cratering and extinction records is weak.

(viii) It is unlikely that the Sun has a companion with mass $\gtrsim 0.03 \mathcal{M}_\odot$.

References

Alvarez, W., and Muller, R. A. (1984) 'Evidence from crater ages for periodic impacts on the Earth'. *Nature* **308**, 718-720.

Anderson, J. D., and Standish, E. M. (1986) 'Dynamical evidence for planet X', in R. Smoluchowski, J. N. Bahcall, and M. S. Matthews (eds.), The Galaxy and the Solar System, University of Arizona Press, Tucson, 286-296.

Anderson, J. D., Lau, E. L., Taylor, A. H., Dicus, D. A., Teplitz, D. C., and Teplitz, V. L. (1989) 'Bounds on dark matter in solar orbit'. *Astrophys. J.* **342**, 539-544.

Antonov, V. A., and Latyshev, I. N. (1972) 'Determination of the form of the Oort cometary cloud as the Hill surface in the Galactic field', in G. A. Chebotarev, E. I. Kazimirchak-Polonskaya, and B. G. Marsden (eds.), The Motion, Evolution of Orbits, and Origin of Comets, Reidel, Dordrecht, 341-345.

Bahcall, J. N. (1984) 'K giants and the total amount of matter near the Sun'. *Astrophys. J.* **287**, 926-944.

Bahcall, J. N., Hut, P., and Tremaine, S. (1985) 'Maximum mass of objects that constitute unseen disk material'. *Astrophys. J.* **290**, 15-20.

Bailey, M. E. (1976) 'Can "invisible" objects be observed in the solar system?' *Nature* **259**, 290-291.

Binney, J. J., and Tremaine, S. (1987) Galactic Dynamics, Princeton University Press, Princeton.

Chester, T. (1986) 'A statistical analysis and overview of the IRAS point source catalog', in F. P. Israel (ed.), Light on Dark Matter, Reidel, Dordrecht, 3-22.

Christy, J. W., and Harrington, R. S. (1978) 'The satellite of Pluto'. *Astron. J.* **83**, 1005-1008.

Cuzzi, J. N., Lissauer, J. J., Esposito, L. W., Holberg, J. B., Marouf, E. A., Tyler, G. L., and Boischot, A. (1984) 'Saturn's rings: properties and processes', in R. Greenberg and A. Brahic (eds.), Planetary Rings, University of Arizona Press, Tucson, 73-199.

D'Antona, F., and Mazzitelli, I. (1985) 'Evolution of very low mass stars and brown dwarfs. I. The minimum main-sequence mass and luminosity'. *Astrophys. J.* **296**, 502-513.

Davis, M., Hut, P., and Muller, R. A. (1984) 'Extinction of species by periodic comet showers'. *Nature* **308**, 715-717.

Duncan, M., Quinn, T., and Tremaine, S. (1987) 'The formation and extent of the solar system comet cloud'. *Astron. J.* **94**, 1330-1338.

Duncan, M., Quinn, T., and Tremaine, S. (1988) 'The origin of short-period comets'. *Astrophys. J. Lett.* **328**, L69-L73.

Duncan, M., Quinn, T., and Tremaine, S. (1989a) 'The long-term evolution of orbits in the solar system: a mapping approach'. *Icarus* **82**, (in press).

Duncan, M., Quinn, T., and Tremaine, S. (1989b) 'Planetary perturbations and the origin of short-period comets', submitted to *Astrophys. J.*

Duncombe, R. L., and Seidelmann, P. K. (1980) 'A history of the determination of Pluto's mass'. *Icarus* **44**, 12-18.

Edgeworth, K. E. (1949) 'The origin and evolution of the solar system'. *Mon. Not. Roy. Astron. Soc.* **109**, 600-609.

Everhart, E. (1967) 'Intrinsic distributions of cometary perihelia and magnitudes'. *Astron. J.* **72**, 1002-1011.

Everhart, E., and Marsden, B. G. (1983) 'New original and future comet orbits'. *Astron. J.* **88**, 135-137.

Fernández, J. A. (1981) 'New and evolved comets in the solar system'. *Astron. Astrophys.* **96**, 26-35.

Fernández, J. A. (1985) 'The formation and dynamical survival of the comet cloud', in A. Carusi and G. B. Valsecchi (eds.), Dynamics of Comets: Their Origin and Evolution, Reidel, Dordrecht, 45-70.

Franklin, F., Lecar, M., and Soper, P. (1989) 'On the original distribution of the asteroids. II. Do stable orbits exist between Jupiter and Saturn?' *Icarus* **79**, 223-227.

Gehrels, T. (1981) 'Faint comet searching'. *Icarus* **47**, 518-522.

Goldreich, P., and Ward, W. R. (1973) 'The formation of planetesimals'. *Astrophys. J.* **183**, 1051-1061.

Gomes, R. S. (1989) 'On the problem of the search for Planet X based on its perturbations on the outer planets'. *Icarus* **80**, 334-343.

Greenberg, R., Weidenschilling, S. J., Chapman, C. R., and Davis, D. R. (1984) 'From icy planetesimals to outer planets and comets'. *Icarus* **59**, 87-113.

Grieve, R. A. F., Sharpton, V. L., Goodacre, A. K., and Rupert, J. D. (1987) 'Detecting a periodic signal in the terrestrial cratering record', in G. Ryder (ed.), Proceedings of the 18th Lunar and Planetary Science Conference, Cambridge University Press, Cambridge, 375-382.

Hamid, S. E., Marsden, B. G., and Whipple, F. L. (1968) 'Influence of a comet belt beyond Neptune on the motions of periodic comets'. *Astron. J.* **73**, 727-729.

Harrington, R. S. (1988) 'The location of Planet X'. *Astron. J.* **96**, 1476-1478.

Hartmann, W. K. (1969) 'Terrestrial, lunar, and interplanetary rock fragmentation'. *Icarus* **10**, 201-213.

Heggie, D. (1975) 'Binary evolution in stellar dynamics'. *Mon. Not. Roy. Astron. Soc.* **173**, 729-787.

Heisler, J. (1989) 'Monte Carlo simulations of the Oort comet cloud', in preparation.

Heisler, J., and Tremaine, S. (1986) 'The influence of the Galactic tidal field on the Oort comet cloud'. *Icarus* **65**, 13-26.

Heisler, J., and Tremaine, S. (1989) 'How dating uncertainties affect the detection of periodicity in extinctions and craters'. *Icarus* **77**, 213-219.

Heisler, J., Tremaine, S., and Alcock, C. (1987) 'The frequency and intensity of comet showers from the Oort cloud'. *Icarus* **70**, 269-288.

Hénon, M. (1970) 'Numerical exploration of the restricted problem. VI. Hill's case: non-periodic orbits'. *Astron. Astrophys.* **9**, 24-36.

Hills, J. G. (1981) 'Comet showers and the steady-state infall of comets from the Oort cloud'. *Astron. J.* **86**, 1730-1740.

Hills, J. G. (1985) 'The passage of a Nemesis-like object through the planetary system'. *Astron. J.* **90**, 1876-1882.

Hoffman, A. (1985) 'Patterns of family extinction depend on definition and geological timescale'. *Nature* **315**, 659-662.

Hoyt, W. G. (1980) Planets X and Pluto, University of Arizona Press, Tucson.

Hughes, D. W. (1982) 'Asteroidal size distribution'. *Mon. Not. Roy. Astron. Soc.* **199**, 1149-1157.

Hut, P., and Tremaine, S. (1985) 'Have interstellar clouds disrupted the Oort comet cloud?' *Astron. J.* **90**, 1548-1557.

Hut, P., Alvarez, W., Elder, W. P., Hansen, T., Kauffman, E. G., Keller, G., Shoemaker, E. M., and Weissman, P. R. (1987) 'Comet showers as a cause of mass extinctions'. *Nature* **329**, 118-126.

Ishida, K., Mikami, T., and Kosai, H. (1984) 'Size distribution of asteroids'. *Publ. Astr. Soc. Japan* **36**, 357-370.

Jackson, A. A., and Killen, R. M. (1988) 'Infrared brightness of a comet belt beyond Neptune'. *Earth, Moon, and Planets* **42**, 41-47.

Kerr, F. J., and Lynden-Bell, D. (1986) 'Review of galactic constants'. *Mon. Not. Roy. Astron. Soc.* **221**, 1023-1038.

Kirk, J. (1978) 'On companions and comets'. *Nature* **274**, 667-668.

Kitchell, J. A., and Pena, D. (1984) 'Periodicity of extinctions in the geologic past: deterministic versus stochastic explanations'. *Science* **226**, 689-692.

Kowal, C. T. (1989) 'A solar system survey'. *Icarus* **77**, 118-123.

Kowal, C. T., and Drake, S. (1980) 'Galileo's observations of Neptune'. *Nature* **287**, 277-278.

Kowal, C. T., Liller, W., and Marsden, B. G. (1979) 'The discovery and orbit of (2060) Chiron', in R. L. Duncombe (ed.), Dynamics of the Solar System, Reidel, Dordrecht, 245-250.

Kuijken, K., and Gilmore, G. (1989) 'The mass distribution in the galactic disc. III. The local volume density'. *Mon. Not. Roy. Astron. Soc.* **239**, 651-664.

Kuiper, G. P. (1951) 'On the origin of the solar system', in J. A. Hynek (ed.), Astrophysics: A Topical Symposium, McGraw-Hill, New York, 357-424.

Kuiper, G. P., Fujita, Y., Gehrels, T., Groeneveld, I., Kent, J., van Biesbroeck, G., and van Houten, C. J. (1958) 'Survey of asteroids'. *Astrophys. J. Suppl.* **3**, 289-427.

Lichtenberg, A. J., and Lieberman, M. A. (1983) 'Regular and stochastic motion', Springer-Verlag, New York.

Luu, J. X., and Jewitt, D. (1988) 'A two-part search for slow-moving objects'. *Astron. J.* **95**, 1256-1262.

Marsden, B. G. (1983), Catalog of Cometary Orbits, Enslow, Hillside.

Marsden, B. G., Sekanina, Z., and Everhart, E. (1978) 'New osculating orbits for 110 comets and analysis of original orbits for 200 comets'. *Astron. J.* **83**, 64-71.

McClintock, J. (1985). Private communication.

Morris, D. E., and Muller, R. A. (1986) 'Tidal gravitational forces: the infall of "new" comets and comet showers'. *Icarus* **65**, 1-12.

Oikawa, S., and Everhart, E. (1979) 'Past and future orbit of 1977 UB, object Chiron'. *Astron. J.* **84**, 134-139.

Oort, J. H. (1950) 'The structure of the cloud of comets surrounding the solar system, and a hypothesis concerning its origin'. *B.A.N.* **11**, 91-110.

Oort, J. H. (1986) 'The origin and dissolution of comets'. *The Observatory* **106**, 186-193.

Perlmutter, S., Burns, M. S., Crawford, F. S., Friedman, P. G., Kare, J. T., Muller, R. A.,

Pennypacker, C. R., and Williams, R. W. (1986) 'The Berkeley search for a faint stellar companion to the Sun', in M. C. Kafatos, R. S. Harrington, and S. P. Maran (eds.), Astrophysics of Brown Dwarfs, Cambridge University Press, Cambridge, 87-92.

Raup, D. M., and Sepkoski, J. J. (1984) 'Periodicity of extinctions in the geologic past'. *Proc. Nat. Acad. Sci.* **81**, 801-805.

Raup, D. M., and Sepkoski, J. J. (1986) 'Periodic extinction of families and genera'. *Science* **231**, 833-836.

Safronov, V. S. (1960) 'On the gravitational instability in flattened systems with axial symmetry and non-uniform rotation'. *Ann. d'Astrophys.* **23**, 979-982.

Seidelmann, P. K., and Harrington, R. S. (1988) 'Planet X—the current status'. *Cel. Mech.* **43**, 55-68.

Sepkoski, J. J., and Raup, D. M. (1986) 'Periodicity in marine extinction events', in D. K. Elliott (ed.), Dynamics of Extinction, Wiley, New York, 3-36.

Shapiro, S. L., and Teukolsky, S. A. (1983) Black Holes, White Dwarfs, and Neutron Stars, Wiley, New York.

Shoemaker, E. M., and Wolfe, R. F. (1982) 'Cratering time scales for the Galilean satellites', in D. Morrison (ed.), Satellites of Jupiter, University of Arizona Press, Tucson, 277-339.

Shoemaker, E. M., and Wolfe, R. F. (1986) 'Mass extinctions, crater ages and comet showers', in R. Smoluchowski, J. N. Bahcall, and M. S. Matthews (eds.), The Galaxy and the Solar System, University of Arizona Press, Tucson, 338-386.

Smoluchowski, R., and Torbett, M. (1984) 'The boundary of the solar system'. *Nature* **311**, 38-39.

Standish, E. M. (1986) 'Numerical planetary and lunar ephemerides: present status, precision and accuracies', in J. Kovalevsky and A. Brumberg (eds.), Relativity in Celestial Mechanics and Astrometry, Reidel, Dordrecht, 71-83.

Standish, E. M., and Hellings, R. W. (1989) 'A determination of the masses of Ceres, Pallas, and Vesta from their perturbations upon the orbit of Mars'. *Icarus* **80**, 326-333.

Stevenson, D. J. (1982) 'Formation of the giant planets'. *Planet. Sp. Sci.* **30**, 755-764.

Stevenson, D. J. (1986) 'High mass planets and low mass stars', in M. C. Kafatos, R. S. Harrington, and S. P. Maran (eds.), Astrophysics of Brown Dwarfs, Cambridge University Press, Cambridge, 218-232.

Sussman, G. J., and Wisdom, J. (1988) 'Numerical evidence that the motion of Pluto is chaotic'. *Science* **241**, 433-437.

Talmadge, C., Berthias, J.-P., Hellings, R. W., and Standish, E. M. (1988) 'Model-independent constraints on possible modifications of Newtonian gravity'. *Phys. Rev. Letters* **61**, 1159-1162.

Taylor, J. H., and Weisberg, J. M. (1989) 'Further experimental tests of relativistic gravity using the binary pulsar PSR 1913+16'. *Astrophys. J.* **345**, 434-450.

Tholen, D. J., Buie, M. W., Binzel, R. P., and Frueh, M. L. (1987) 'Improved orbital and physical parameters for the Pluto-Charon system'. *Science* **237**, 512-514.

Tombaugh, C. W. (1961) 'The trans-Neptunian planet search', in G. P. Kuiper and B. Middlehurst (eds.), Planets and Satellites, University of Chicago Press, Chicago, 12-30.

Tremaine, S. (1986) 'Is there evidence for a solar companion star?', in R. Smoluchowski, J. N. Bahcall, and M. S. Matthews (eds.), The Galaxy and the Solar System, University of Arizona Press, Tucson, 409-416.

van den Bergh, S. (1989) 'Life and death in the inner solar system'. *Publ. Astron. Soc. Pac.* **101**, 500-509.

van Houten, C. J., van Houten-Groeneveld, I., Herget, P., and Gehrels, T. (1970) 'The Palomar-Leiden survey of faint minor planets'. *Astr. Astrophys. Suppl.* **2**, 339-448.

Weinberg, M. D., Shapiro, S. L., and Wasserman, I. (1987) 'The dynamical fate of wide binaries in the solar neighborhood'. *Astrophys. J.* **312**, 367-389.

Weissman, P. R. (1983) 'The mass of the Oort cloud'. *Astron. Astrophys.* **118**, 90-94.

Weissman, P. R. (1985) 'Dynamical evolution of the Oort cloud', in A. Carusi and G. B. Valsecchi (eds.), Dynamics of Comets: Their Origin and Evolution, Reidel, Dordrecht, 87-96.

Weissman, P. R. (1986) 'The mass of the Oort cloud: a post Halley reassessment'. *Bull. A. A. S.* **18**, 799.

Whipple, F. L. (1972) 'The origin of comets', in G. A. Chebotarev, E. I. Kazimirchak-Polonskaya, and B. G. Marsden (eds.), The Motion, Evolution of Orbits, and Origin of Comets, Reidel, Dordrecht, 401-408.

Whitmire, D. P., and Jackson, A. A. (1984) 'Are periodic mass extinctions driven by a distant solar companion?' *Nature* **308** 713-715.

Wisdom, J. (1982) 'The origin of the Kirkwood gaps: a mapping for asteroidal motion near the 3/1 commensurability'. *Astron. J.* **87**, 577-593.

Wisdom, J. (1983) 'Chaotic behavior and the origin of the 3/1 Kirkwood gap'. *Icarus* **56**, 51-74.

Yeomans, D. K. (1986) 'Physical interpretations from the motions of comets Halley and Giacobini-Zinner', in Proceedings of the 20th ESLAB Symposium on the Exploration of Halley's Comet (ESA SP-250), 419-425.