PART VIII

THE ORIGIN OF COMETS

1

THE ORIGIN OF COMETS

A. H. DELSEMME

Empirical data are confronted with different hypotheses on the origin of comets. The hypotheses are classified into three categories: 1) Comets were condensed from the solar nebula and ejected later into the Oort's cloud. 2) Comets were condensed in situ, more or less recently, on their present trajectories; 3) Reversing the arrow of time in the traditional evolution of comets. Only two hypotheses, both from the first category, are found to be in agreement with all empirical data. The first hypothesis explains the origin of the Oort's cloud by the perturbations of the giant planets (mainly Uranus and Neptune and possibly Pluto) on a ring of proto-comets, during the final accretion stages of the solar system. The second hypothesis uses the fast mass loss of the solar nebula to expell an outer ring of proto-comets into elliptic trajectories. Although no empirical evidence requests that the Oort's cloud be older than a few million years, its matter is not likely to be from a different reservoir than solar system stuff, and no satisfactory theory explains its formation more recently than 4.5 billion years ago.

EMPIRICAL DATA

Empirical data on the origin of comets, that are not somewhat model-dependent, are not very numerous; they come either from the physical nature of the nucleus, or from the evolution of the orbits.

The physical nature of the cometary nucleus is still poorly known, and the principal clue comes from its volatile fraction. Since the gravity field of the nucleus is extremely weak (of the order of 10^{-4} g) no gaseous atmosphere can be kept by gravity beyond a few ^{o}K, therefore any volatile fraction would have vaporized and dissipated away had it not been kept at a very low temperature until very recently.

a. PRIMITIVE TEMPERATURE AND ITS EVOLUTION

If we want to transform the previous statement into a quantitative assessment, a model is already needed, but we can keep it simple. For instance, we can compute the time needed at different temperatures, to vaporize *in vacuo* a large homogeneous nucleus of, say, 3 km diameter (Table I). Such a model does not introduce any assumption on the source of heat (whether external or internal), the albedo, or any other physical characteristic. The kinetic theory of gases sets the sublimation rate *in vacuo,* at a given surface temperature (this theory is given p. 232 in Delsemme and Miller 1971). The presence of a gaseous phase (for instance, in a primitive nebula) would slow down the flow rate linearly until a steady state is reached when the vapor pressure equals the actual partial pressure of that gas in the nebula.

TABLE I

CONSTANT TEMPERATURE NEEDED TO SUBLIMATE IN A GIVEN TIME A MEDIUM-SIZED (3 km) COMETARY NUCLEUS IN VACUUM (*)

Characteristic Sublimation Times	Water or Clathrates	Carbon Dioxide
in one year	257 °K	140 °K
in one million years	143 °K	86 °K
in five billion years	117 °K	73 °K

(*) deduced from the kinetic theory of gases, see Delsemme and Miller (1971).

The presence of large amounts of water in comets seems sufficiently established: hence, the temperatures given in Table I for a water ice nucleus set strong empirical *upper limits* for the "primitive" temperature of comets and its subsequent evolution. Here, "primitive" temperature means the earliest nuclear temperature after the nebula has dissipated, or (in Vsekhsvyatsky's model, discussed later) the nuclear temperature after leaving Jupiter's ionosphere.

The three characteristic times listed in Table I can be used for different purposes. One year is a characteristic time to leave the inner solar system; one million years, to go to or to come from the Oort's cloud (it could also be a characteristic time of accretion); whereas five billion years is the probable upper limit for the age of the nucleus.

The presence of a moderate amount of compounds more volatile than water (like the observed HCN and CH_3CN) does not change the previous upper limits, because these compounds may be adsorbed in the molecular cavities of clathrates (solid hydrates of gases); this would also include less certain constituents like CO_2 or CO, provided the total remains within 15 to 17% of the amount of water (this is set by their adsorption limit in clathrates).

However, more volatile ices could also exist. In particular, solid CO_2 could be present in *some* comets, in larger amounts than what could possibly be adsorbed in clathrates (Delsemme 1977). For this reason, the temperatures needed to vaporize a CO_2 nucleus have been also listed in Table I. These temperatures become lower upper limits, although they might still be easier to challenge than those for water ice.

Without any preliminary hypothesis on the origin of comets, Table I makes it clear that the cometary stuff was assembled at cold temperatures. The longer the formation period, the colder the maximum temperature limit in vacuum, because the formation mechanism has to work against the sublimation rate; the assumed partial vapor pressures in primitive nebula models (10^{-4} to 10^{-8} atmosphere) are low enough not to change these temperatures very much; whereas a *minimum* temperature of formation cannot be assessed without using uncertain assumptions on which gases were present but did not condense (Delsemme 1975, Shimizu 1976, Delsemme and Rud 1977). For instance, using the H/O and C/O ratios observed in comet Kohoutek, *and* a condensation model of the primitive nebula, Delsemme (1976) gives 110 ± 60 °K for the accretion temperature of the nucleus of comet Kohoutek. Whipple's (1950) remarkable insight has therefore been entirely confirmed by recent data.

b. ISOTOPIC RATIOS

The $^{12}C/^{13}C$ isotope ratio is the only one that has been measured (so far, in four comets: Vanysek 1977) but the results remain very uncertain. However, the terrestrial ratio (89) is much more likely in comets than the interstellar value (about 40) (hints that it could be even larger than the terrestrial ratio must be accepted so far with caution). The results make it very probable that comets were indeed born from some of that solar-system stuff, that separated from an interstellar cloud five billion years ago, rather than from some interstellar matter more recently scattered by supernova explosions. The argument is important, because it rules out recent ($< 10^8$ years) replenishment of the cometary volatiles by some uncertain mechanism from interstellar material. However, it does not say *when* the event that separated comets from solar system stuff took place (it could be for instance, recent volcanic eruptions of Jupiter, Vseksvyatsky 1977).

c. CHEMISTRY

Meteor data as well as cometary micrometeorites collected in the upper atmosphere suggest pure solar abundances for metals in comets (Millman 1977, Brownlee *et al.* 1977); cometary spectra, combined with dust-to-gas ratios, suggest (within a factor of two) solar abundance for oxygen, probably a mild depletion for carbon and possibly for nitrogen, and certainly a drastic depletion of helium and all uncombined hydrogen (Delsemme 1977). We do not have a single complete molecular analysis of the snowy constituents in any comet, although water, probably CO_2, and possibly CO are the most abundant species; HCN and CH_3CN are the only two minor constituents we are reasonably sure of, although a formal proof that they could not be produced in the coma from other parent molecules, has yet to be found. From the elements' abundances, comets seem to be the most primitive material still existing (at least in such a *weak gravity field*: 10^{-4} g; even if it were possible to make comets by overcoming the strong gravityfield of one of the giant planets, it is unclear how quasi-solar abundances might be preserved). Apart from these general considerations, the chemical evidence remains inconclusive, because it is consistent with different possible scenarios. The previous empirical data could probably be duplicated, as well as by clumping together interstellar grains (Greenberg 1977) as by condensing the solar nebula, in particular if proper chemical kinetics is introduced (as by Anders *et al.* 1977). The difference between these two scenarios is indeed a matter of kinetics: in the first case, the low temperature implies that charge-exchange reactions prevail; the low pressure rules out triple collisions, so that grains are used to store those intermediate steps that will yield larger molecules; a hydrogen deficiency is easily introduced in these large molecules, by the fact that hydrogen does not stick easily upon grains. In the second case, classical thermodynamics, with possible catalysis induced by the presence of grains, prevails at those larger temperatures and pressures of any solar nebula model. The fact that simple-minded condensation models (like Delsemme and Rud's 1977) have almost duplicated observations, whereas charge-exchange chemistry has not, may come rather from conceptual difficulties and remaining uncertainties in interstellar chemistry, than from valid physical reasons. Expected advances in interstellar chemistry, combined with mass-spectrometry data from a first flyby or rendez-vous mission to a comet, could bring a drastic improvement in the present situation within the next decade.

d. ORBITS

No interstellar comet has ever been seen among the 600-odd different comets that have been observed so far. The few nominal hyperbolic comets have velocities so small at infinity, that they still would follow the general motion of the solar system for aeons. Upper limits for interstellar comets have been discussed recently by Whipple (1975). All comets can therefore be assumed to belong to the solar system.

The evolution of an individual orbit can be studied by integrating it forwards and backwards. However, the stunning accuracy of celestial mechanics is lost for durations longer than a few million years. This accuracy is even lost much earlier for those comets entering the inner solar system, because unpredictable non-gravitational forces (stemming from the asymmetric vaporizations) perturb trajectories near perihelion. Extrapolation from a single orbit does not yield evolutionary properties, because the arrow of time can always be changed mathematically along a given orbit by the substitution (-t) instead of (t) for the time variable. In particular, the same trajectory may describe the evolution of a long-period comet into a short period comet, or vice versa. This opens the door to theories reversing the evolution (Vsekhsvyatsky 1977). Of course, the passage of time cannot be detected except statistically for instance by the entropy increase of a closed system. In particular, the spherically random distribution of those comets coming from the Oort's cloud resembles that of the globular clusters surrounding the galaxy, and suggests that the analogy comes from the same cause: its old age. However, the scaling down of the geometry, plus the residual uncertainty on its accurate shape coming from the small number of observed "new" comets, does not imply that it must be much older than 10^7 - 10^8 years. Standing in contrast, the flattening and generally direct rotation of the system of short-period comets does not imply that it is necessarily young, but only that it has *recently* exchanged momentum and energy with the giant planets (which is indeed a known fact).

The distribution of the binding energy of comets (conveniently expressed by their reciprocal semi-major axis a^{-1}, which is proportional to the binding energy per unit mass) is one of those statistical properties from which the arrow of time can be deduced, at least in principle. Because of planetary perturbations, each cometary passage through the inner solar system changes the binding energy of an orbit, by a random (positive or negative) number, whose average absolute value is near 700 (in 10^{-6} AU^{-1} units, Everhart and Raghavan 1970). This random walk of the binding energies produces their diffusion in such a way that, in the long run, the number of comets that pass perihelion per year would become constant in constant intervals of a^{-1} (van Woerkom 1948) whatever the original distribution of the energies. This important result has not always been clearly understood or accepted in all generality because it is based on the peculiar nature of the comet-planet interaction. Perhaps it is useful to illustrate it by the analogy derived from the kinetic theory of gases. The two types of particles in this"gas" (planets and comets) have velocities of the same order, but masses that differ by (crudely) a factor of more than 10^{10}. In other words, planets are 10^{10} times "hotter" than comets, but they do not exchange energy and momentum by direct hits, and the equipartition of energy is long to come. However, it would accelerate comets eventually to velocities 10^5 times as large as those of planets although they will be lost on hyperbolic orbits much earlier. Van Woerkom's distribution is nothing else but the flat Bolzmann distribution for an infinitely high "temperature" which is not a bad approximation for an energy about 10^{10} times larger than the binding energy to the sun. This constant flux of comets for all energies (when a steady state is reached) is shown by the horizontal dashed line in Figure I. If peaks and gaps appear, sources and sinks of comets can be identified, showing the arrow of time in the diffusion process. For instance, the solid line in Fig. I represents a fictitious source of comets (peak A) in the vicinity of 5 AU (for instance, from Jupiter's "volcanoes"). B is the plateau where van Woerkom's steady-state is reached; C shows that we have a sink at $1/a = 0$ (comets that leave on hyperbolic orbits will never come back), and slope D, that we have a sink for very short distances (comets disappear because they decay in the solar heat). The characteristic times of diffusion are given by van Woerkom; they are short in the inner solar system because comets turn fast, but would reach 10^7 - 10^8 years on the left-hand side of Fig. I.

The empirical evidence does not look at all like the fictitious example of Fig. I. Fig. II represents the statistical data, deduced from Marsden's

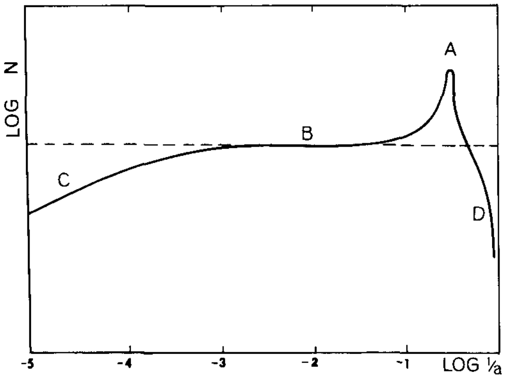

Figure 1 *Fictitious distribution of cometary orbits, to illustrate what would happen if the only source of comets were at Jupiter's distance (peak A at log* $a^{-1} = -0.5$*). The dashed horizontal line represents van Woerkom's distribution. The slope at C is due to the sink of comets for* $a^{-1} \leqslant 0$ *(hyperbolic comets are lost). The slope at D comes from cometary decay by the solar heat. N is the number of comets that pass perihelion per century, per* a^{-1} *unit, for a given range in perihelion distances; a is the semi-major axis of the orbits.*

(1975) catalogue of cometary orbits. N is the number of comets that pass perihelion per century, reduced for a Δq of 1 AU, per $\Delta(a^{-1}) = 1$ AU^{-1} (q is the perihelion distance). The distribution AB for the long-period comets is taken from all of Marsden's (1975) osculating orbits that can be corrected by Everhart and Raghavan's (1970) changes in binding energy. The distribution for the short-period comets BDC has not been corrected (the original orbits have no meaning for comets that do not leave the solar system, and in any case it would not make any visible difference). B is the traditional cutoff (period of 2 centuries) between long-period and short-period comets. This cutoff is an artifact because all comets with periods of more than two centuries have been observed only once. From now on, BD will be called the set of "intermediate" period comets.

Connecting the statistics of long-period and short-period comets is easy: in each interval $\Delta(a^{-1})$, we count the number of *different* comets going through perihelion, per unit of time and normalize per unit of perihelion distance.

The empirical evidence in Fig. II shows that we have at A a source of comets (Oort's *new* comets) for values of a^{-1} between 10^{-5} and 5×10^{-5} AU, and that their diffusion ABD towards the inner solar system is depleted by a very fast decay. The dashed and the dotted lines marked 333 and 52 represents Oort's two models of exponential decay (respectively in 333 and 52 perihelion passages). The lifetime found for a comet depends very much on the model (Whipple 1962) but this problem is not considered here.

The existence of another peak C at the short-period comets' position, apparently suggests the presence of another source of comets near Jupiter. This peak has been the origin of many difficulties and misunderstandings. Its existence is the reason of all theories that have tried to reverse the sense of evolution (Proctor 1884, Crommelin 1910, Vsekhsvyatsky 1933-1977).

First, the gap at D is analogous to one of the Kirkwood gaps for asteroids: there are no comets in the period gap from comet p Neujmin I (P = 17.9 years) to comet p Crommelin (P = 27.9 years). The Jupiter 2:1 resonance has a period of 23.7 years, the Saturn 3:2 resonance, a period of 19.7 years, strongly suggesting a resonance-induced gap. Because of this gap, curve ABD is difficult to extrapolate below C, but it is readily apparent that peak C is approximately three or four orders of magnitude larger than what could be explained by the steady diffusion of the *observed* long-period comets; this is independent of the decay models, since it is based on the empirical curve of Fig. II.

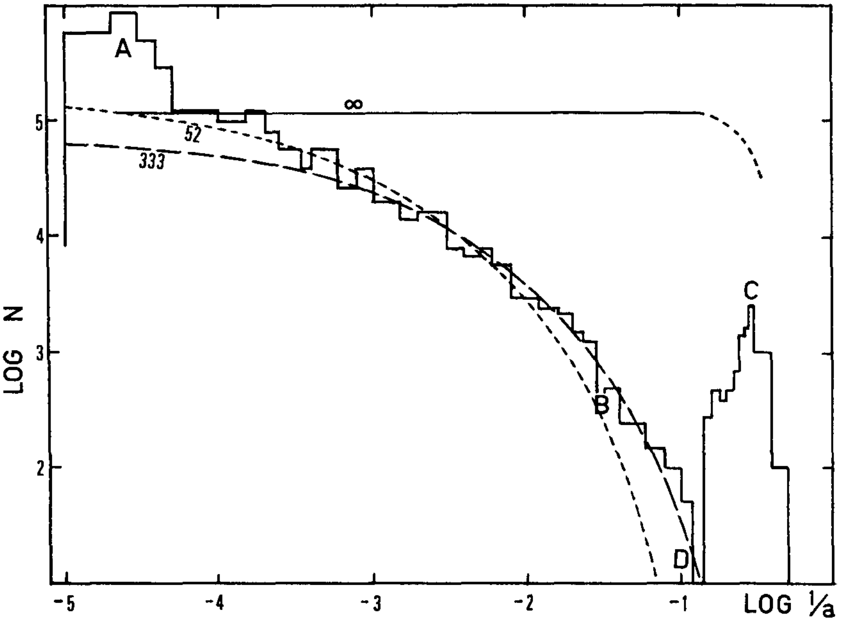

Figure 2 *Actual distribution of all cometary orbits available. N is the number of comets that pass perihelion per century, per a^{-1} unit, per Δq unit; a is the semi-major axis of the orbits, q is the perihelion distance; both are expressed in A.U. Data are reduced to a $\Delta q = 1$ AU (approximately from 0.5 to 1.5 AU). Dashed and dotted lines are Oort's exponential models of decay (in respectively 333 and 52 passages). A is the source of comets coming from the Oort's cloud, B is the traditional cutoff for long-period comets. BD is the range defining the "intermediate-period" comets. D is a Kirkwood-type gap for comets (it contains Jupiter 2:1 and Saturn 3:2 resonances). C is the short-period-comet peak; its source is an unobservable subset of long-period and intermediate-period comets: the prograde comets whose perihelion is near Jupiter. Their flux can be predicted by drawing the horizontal line marked ∞ (they do not decay because they are too far away from the sun).*

However, Everhart's (1972) numerical experiments with orbits has established that 90% of the captures happen in the range of perihelion distances 4 AU $< q <$ 6 AU and of orbit inclination $0^\circ < i < 9^\circ$. The effective range of orbit inclinations is strongly prograde, which explains how a flattened system, turning in the prograde direction (the short-period comets) can be derived from a spherically symmetrical system (the long-period comets) answering neatly one of Vsekhsvyatsky's major arguments against capture. The range of perihelion distances of the capture zone *explains the origin* of the (secondary) source of short-period comets: we do not see the whole set of long-period comets; those we observe are only a subset, whose perihelia are near the earth ($0.5 < q < 1.5$ AU). All our statistics and in particular curve ABC come from this subset. However there is another unobservable subset of long-period comets, whose perihelia are near Jupiter ($4 < q < 6$ AU). *Those comets,* that are always too far away to be seen, *are the source of the short-period comets.* The subset we observe decays fast; because of solar vicinity, the volatile fraction of the nucleus sublimates. The unobservable subset physically decays in times larger than the age of the solar system (Delsemme 1973). If source A of comets has lasted more than 10^8 years, then the orbits of the unobservable subset have reached van Woerkom's distribution and can be represented by a flat line drawn from peak A to the right, down to Jupiter's distance. Oort's (1950) analysis has shown that the number of "new" comets per unit Δq (q = perihelion distance) is a constant. Therefore peak A reaches the same level at 5 AU as at 1 AU. However, to take empirically into account the permanent injection of "new" comets, the flat line of van Woerkom's distribution has been started from about an order of magnitude lower than the top of peak A. The predicted number of intermediate-period comets with perihelion at Jupiter's distance is large enough to explain the number of comets in peak C, from captures by the repeated action of the giant planets (see the detailed computation in Delsemme 1973. The coincidence is good: 84 predicted against 73 observed, but the error bar is very large). Astronomers, who do not like the simplicity of this approach and prefer statistics on individual orbits, have shown how the repeated action of all the giant planets can be used to progressively transform enough long-period orbits into intermediate-period and eventually into short-period orbits (Kazimirchak-Polonskaya 1972, Everhart 1972, 1977). Their efforts must be praised very much, because they transform semi-quantitative estimates into quantitative arguments, but the picture of cometary evolution that seems to emerge remains the same, in particular:

--1. Although the observed short-period comets decay fast, their number is constant because new ones are continuously captured (by the *combined* and *repeated* action of the giant planets) mainly from the subset of the prograde intermediate-period comets whose perihelia are near Jupiter's orbit.

--2. The intermediate-period comets result mainly from the random walk in binding energy of the long-period comets, coming from their repeated passages near the giant planets; all these comets stem from the steady source of "new" comets shown by peak A.

--3. The origin of the steady supply of "new" comets has been explained by Oort (1950) by stellar perturbations on a large reservoir of comets weakly bound to the sun (the Oort's cloud).

The existence of the Oort's reservoir of comets which is the only known explanation of the direction of cometary evolution given by Figure II, has been confirmed by recent work (Marsden 1977); in particular Marsden and Sekanina (1973) have shown that, the more accurately "new" orbits are known, the narrower peak A becomes; unfortunately, the number of extremely accurate orbits becomes vanishingly small. To convince people that peak A is not an artifact coming from the selection of the "best" orbits, it is worth looking at its fine structure *before this* selection. Fig. III shows a direct plot of *all* available orbits. In his original paper, Oort (1950) had a (unresolved) peak of 14 comets in one single box. We have now 37 comets in ten boxes from -0.5 to +1.0 (in 10^{-4} AU^{-1}) that resolve the peak; a total of 42 comets if we count weakly hyperbolic comets, as Oort had done. Correction for non-gravitational forces suggests that most new comets come from between 20,000 and 50,000 AU and that most residual hyperbolic orbits when they are not spurious can be explained by these forces.

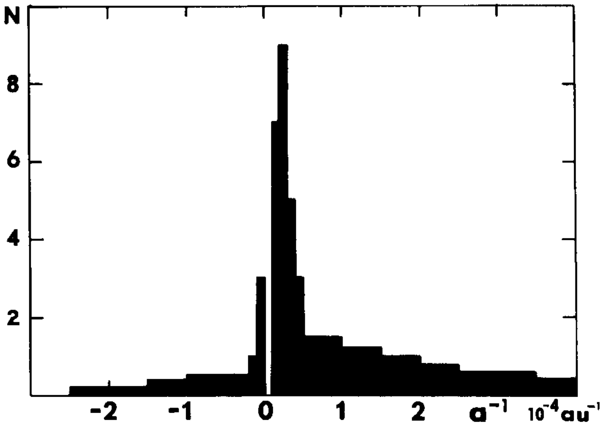

Figure 3 *This blow-up of peak A from Figure 2, with N versus* a^{-1} *(instead of log N versus log* a^{-1}*), makes the source of "new" comets (coming straight from the Oort's cloud) very conspicuous. This is a direct plot of all available orbits, before any selection procedure. N is the observed number of cometary orbits, per interval of* 10^{-5} AU^{-1}. *The boxes in the peak coincide with these intervals from -0.2 to +0.5 (in* 10^{-4} AU^{-1}*); they are enlarged to 5 x* 10^{-5} AU^{-1} *for all places outside of the peak were not enough comets are availble. The difficulty to sieve out in the literature, the small number of genuine parabolic orbits from the large number of parabolic approximations, apparently explains the gap at zero as an artifact. The zero should be shifted to the left by an average of* 10^{-5} *AU, to take the non-gravitational forces into account.*

Lyttleton's (1974) criticism is based on a misconception that has been refuted many times (see Marsden 1977).

THEORIES AND MODELS

The present existence of Oort's cloud does not imply that we understand its origin, or its age, or that we know accurately how many comets it contains. Some 20 new comets per year must reach the region of the giant planets (we see 1 or 2). If (and only if) comets were all created at the origin of the solar system, extrapolating the present steady state suggests a *lower limit* of 20 comets per year x 5 x 10^9 years = 10^{11} comets in the *original* Oort's cloud; or probably several times that amount because of its slow diffusion towards interstellar space, and because it has obviously not yet been emptied of its contents. Nezhinsky (1972) gives 1.1 x 10^9 years for the lower limit of its half-life. Oort's (1950) assessment of the *present* number of comets in the cloud was 1.9 x 10^{11}; however, new data on the observed distribution of comets in the peak (Fig. III, corrected by Marsden and Sekanina's (1973) average shift of the peak) combined with Oort's (1950) Table 4 p. 98, gives a revised number of about 0.9 x 10^{11} comets in the present cloud.

Using Oort's model, the number density at the margin of the cloud can be derived (Fig. IV) from the peak of Fig. III. For those distances where it has been depleted by stellar perturbations, it seems to diminish with an inverse-square law of the heliocentric distance (Fig. 4). However, these results are of course very uncertain. New comets necessarily come only from the margin of the cloud, that is, where it is perturbed by stars. Therefore, the density of Oort's cloud is totally unknown for shorter distances (< 10^4 AU);

but since it has not been depleted there by stellar perturbations, the slope of Fig. 4 might flatten out. This is suggested by the dashed line in Fig. 4. Since the content of the cloud within 10^4 AU cannot change its total mass, the present results combined with Nezhinsky's (1972) half life yields 2 x 10^{12} comets or 30 earth masses for the upper limit of the original cloud.

Since any hypothesis on the origin of comets must now be reduced to a hypothesis on the origin of Oort's cloud, we will now review the principal hypotheses that have been proposed in the past.

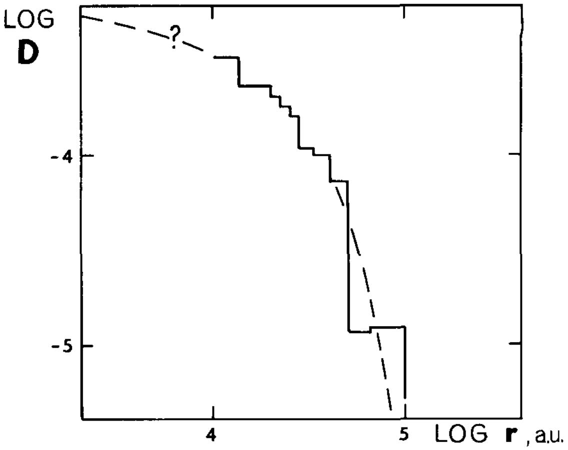

Figure 4 *Number density D of comets per AU^3 in the margin of Oort's cloud, versus heliocentric distance r, deduced from the unmodified peak's structure of Figure 4, combined with Oort's model of the stellar perturbations at distance r. The results would be shifted somewhat to the left, if the action of the non-gravitational forces were taken into account. For distances less than 10^4 AU, stellar perturbations are too weak, therefore no data on the density of Oort's cloud are available. The uncertain distribution is suggested only by the dashed line.*

a. COMETS WERE CONDENSED FROM THE SOLAR NEBULA AND EJECTED LATER INTO OORT'S CLOUD.

1. Oort's (1950) original hypothesis was that comets were born with the asteroids. In his second note added in proof, he discusses the possibility of a rupture of a planet. A belt of comets in the inner solar system, whatever its origin, would steadily diffuse away because of planetary perturbations (as the original belt of asteroids did, or as in the fictitious example of Fig. I that is directly applicable to the early diffusion in Oort's hypothesis). Because of the large value of the binding energy of such a belt, the process is very inefficient, since the objects ejected out of the solar system are several orders of magnitude more numerous than those stored in Oort's cloud. However, Oort's proposal is no more very attractive, mainly because of the need of a low primitive temperature (see Table I) to preserve the volatile fraction: the asteroid belt is and has probably always been too hot to store ices.

2. Kuiper (1951) aware of these consequences of Whipple's (1950) icy conglomerate model proposes that comets are the condensation products of the *outer parts* of the nebula that formed the planets; more specifically, Pluto would eventually disperse away all those small objects that would have been

in the portion of a belt from 38 to 50 AU. An unobservable remnant of this belt could still be present beyond 50 AU, since no perturbing bodies are available to scatter it away. The smaller binding energy of the "Pluto comets" (16 times smaller than for asteroids) considerably improves the efficiency of the process, and its location seems much more compatible with a cool primitive temperature. Kuiper's hypothesis has been strongly supported by Whipple, who tried inconclusively (1964, 1972) to find gravitational evidence for a comet belt that would still be present beyond Neptune.

3. Safronov (1972, 1977) generalizes the previous approach, by showing that the building up of the Oort's cloud through the ejection of minor bodies by the perturbations of *all the giant planets* is a necessary consequence and a natural by-product of their growth by accretion. Most of the mass of the proto-comets was ejected by Jupiter's action but its efficiency to fill up Oort's cloud was very low; because of the large initial binding energy of the minor bodies ejected by this planet, most of them were lost to hyperbolic orbits. The contributions of the giant planets to the mass stored in the Oort's cloud would be: 8% Jupiter, 16% Saturn, 24% Uranus, 52% Neptune, and the total (initial) mass of the cloud is estimated to be about three times that of the Earth (2×10^{11} objects of 10^{17} grams), whereas the total mass lost to infinity would be about three hundred Earth masses. This is based on the low-mass model (5 to 6% of the solar mass) favored by the Soviet scientists for the solar nebula, and on accretion of 1/3 of the gas available in the planet zone (hydrogen and helium) mostly by Jupiter and Saturn. Uranus and Neptune would therefore have accreted from pure cometary stuff (as is indeed suggested by their present hydrogen depletion). Safronov's model could certainly be improved in the future, for instance by comparing in more detail the collision and ejection probabilities, (Weidenschilling 1975) but his basic results are likely to stay (within the constraints of the low-mass solar nebula model). In particular, Öpik (1975) strongly supports them.

4. Cameron (1972, 1977) is uneasy with the idea of cool temperatures inside *his* model of the nebula (this high-mass model is evolving towards larger and larger masses). Consequently, he visualizes first (1972) the formation of comets in many nebular satellites, left over outside of the main solar nebula, during its early collapse. Presumed difficulties in condensing these nebular satellites *in situ* into cometary nuclei, and the development of evolutionary model sequences of the solar nebula (Cameron 1977) lead him recently to a novel hypothesis. His solar nebula has most of its mass lying beyond the orbit of Neptune's formation. The condensed solids in the outer disk (possibly, unprocessed interstellar grains) settle into cometary nuclei; their circular orbital periods will be several times 10^4 years. But the final stage of mass loss of his solar nebula is much shorter than these periods, therefore it transforms impulsively the circular orbits into elongated ellipses, producing the Oort's cloud with a very high efficiency.

Although there may be difficulties in accreting particles into comets fast enough and so far away, this proposal is highly original, but very much model-dependent. The future will tell whether massive models of the solar nebula will prevail.

b. COMETS WERE RECENTLY CONDENSED *IN SITU* OR CAPTURED.

The extremely slow growth of particles in a low-density region makes it very difficult to imagine processes able to grow comets within the Oort's cloud (Opik 1966). For instance, the homogeneous distribution of 1.5 solar mass in a sphere reaching 50,000 AU yields an average density of 10^{-20} g/cm^3. A sticking probability of 100% for each atomic collision (except hydrogen), at T = 100 $^{\circ}$K, ignoring all destruction processes, yields mm grains in 10^{10} years, km spheres in 10^{15} years. For this reason, all successful theories must rely on ingenious scenarios, like making comets near 100 AU but with just enough velocity to reach 50,000 AU before coming back.

For instance, Lyttleton's (1953-1968) accretion theory explains the formation of all long-period comets by several events (many, to provide isotropy; namely, the passage of the sun through many interstellar clouds) that happened a few million years ago (the longest time needed is the transit time back and forth of those comets that have the largest periods). His theory is based on the accretion of interstellar dust at distances from 10 to 1000 AU, because of the gravitational focussing effect of the sun; the orbits of dust grains are hyperbolas with the sun as focus and these orbits all intersect the "accretion axis," a line drawn from the center of the sun to the direction of the relative velocity vector of the cloud. The accretion works 1) if there are collisions among particles, and 2) if these collisions are inelastic, so that the transverse momentum be dissipated into heat. To make collisions probable, turbulent motions in the cloud must be kept small ($<$ 1 m/sec) for sharp focussing, and interstellar grains big (10 microns). Even so, the efficiency is low (10^{-5}). Since the accretion is radial, unspecified secondary effects must give the accreted dust enough rotational momentum not to fall back onto the sun. If collisions are inelastic, heat dissipation vaporizes ices. The model might be compatible with Lyttleton's sand bank model of the cometary nucleus, but both models are at complete variance with presence evidence. In particular, they do not explain either peak A in Fig. 2 or 3, or the existence of an icy conglomerate (Whipple 1950).

Other proposals have not been developed into models. For instance, Reeves (1974) suggests that the shock wave of the solar wind (at rather large distances from the sun) could result in its continuous condensation into comets. The peculiar difficulty here is that the solar wind plasma behaves as a fluid only because of its frozen magnetic field, but there are no direct collisions between atoms in this by-product of the solar corona.

McCrea (1975) proposes to use the compression of an interstellar cloud that was induced by the passage of the solar system through a dust lane bordering the Orion arm of the Galaxy, as recently as 1 or 2 million years ago (he also explains the ice ages). Again, this is not in formal contradiction with the (unknown) age of the Oort's cloud, but order-of-magnitude computation shows that the previous difficulty stands if unrealistic compressions are not claimed. Greenberg (1974, 1977) thinks that the balance explaining the depletion of O, C, N from the interstellar medium, cannot be found in fine dust, but in "snowballs" that could conceivably be as large as cometary nuclei. If countless interstellar comets were indeed formed by some unexplained mechanism, it must be first explained by we have never seen any strongly hyperbolic trajectory.

Witkovski (1972) proposes such an explanation. He assumes first the existence of a spherical gas-and-dust cloud bound to the sun with a radius of 100,000 AU or more and a density of the order of 10^{-18} g cm^{-3}. Only direct motion ensures stability of the cloud for distances between 100,000 and 230,000 AU. An interstellar comet entering the cloud will start to increase its mass. A homogeneous uniform cloud would not capture the comet, but a patchy cloud with large density concentrations would accelerate the process of accretion leading to capture. All observed comets would have been captured by the cloud and would therefore be a mixture of a cosmic and of a planetary component. Bound comets would replenish their supplies of gas and dust by going back through the cloud. Witkovski's cloud contains an unrealistic mass of one thousand suns. This is a measure of the difficulty of capturing interstellar comets as well as of replenishing supplies of gas near aphelion.

c. REVERSING THE ARROW OF TIME FOR COMET'S EVOLUTION

1. Lagrange (1814) Proctor (1884) Tisserand (1890) Newton (1893) Crommelin (1910) had concluded that it was impossible to explain the number of observed short-period comets (peak C of Fig. II) by a *simple capture process* from the field of *long-period* comets. (They were right; to do so, we now use the *multiple and repeated* action of the giant planets (Everhart 1977)

in order to make enough intermediate-period comets first). Since they did not visualize this possibility, Proctor concluded (wrongly) that there was no alternative but to accept that the primitive source of comets was located within or near Jupiter (at peak C, Fig. II). This injection of short-period comets, discussed later many times in the literature, would have taken place in one or several events, as recently as 10^6 - 10^7 years ago; a single big event (the traditional 19th century explanation for the origin of asteroids) would be the disruption of a planet, that would have the advantage of explaining simultaneously the origin of comets, asteroids and meteorites. Another possibility would be grandiose eruptions from Jupiter's volcanoes. Vsekhsvyatsky (1933-1977) has developed this line of reasoning. He established first (1933) that comets have a fast decaying luminosity. Vsekhsvyatsky's conclusions on the average lifetime of decay of the short-period comets have been challenged on the ground that instrumental effects have introduced a systematic error on ancient observations, that make comets decay apparently faster than they really do. However, it is generally accepted that all observed comets decay, if not in a matter of a few centuries, at least in a few millenia. For this reason, Vsekhsvyatsky extends Proctor's hypothesis to the present: volcanic processes are still expelling comets from the giant planets and from their satellites; Saturn's rings are therefore the (unstable and fast disappearing) by-products of these explosions. The dark line seen on Jupiter's equator during the first half of this century was the shadow of its fast disappearing ring (Vsekhsvyatsky 1962). The recent discovery of Uranus rings, as well as the volcanoes on Venus and Mars had been predicted as a consequence. However, no numerical model has ever been developed for the ejection of masses of snow and dust reaching cometary sizes through Jupiter's atmosphere and gravity field. The pristine nature of comets (Delsemme 1977), in particular the solar abundance ratio for metals and the approximate solar ratio of gas-to-dust, seem difficult to preserve if they are derived from a planet whose gravity has presumably induced differentiations. The most serious argument against this hypothesis seems however that there is another primary source of "new" comets (peak A in Fig. II) much more abundant than the assumed (secondary) source at peak C; peak C cannot therefore produce enough comets by its present diffusion to explain peak A and the Oort's cloud, whereas an alternate explanation has been found to derive peak C (through another subset of long-period comets: those with prograde perihelia near Jupiter) from the same source as peak A, namely from the Oort's cloud. In other words, the arrow of time empirically established by the evolution of comets (Fig. II) does not support Vsekhsvyatsky's hypothesis.

Van Flandern (1977) presents in this book arguments supporting the previous existence of a 90 earth-mass planet, at the average distance of the asteroid belt, and its disintegration 5 million years ago into asteroids and comets. Peak C would have to be explained by temporary storage of intermediate-period comets, later captured by Jupiter, but the Oort's cloud would not exist, outside of the transient return of those comets that come back for the first time. In particular, the extreme narrowness of peak A, which appears better on Fig. 3 would be totally inexplicable, because, even in the unlikely hypothesis of a constant and proper velocity for all fragments of the disintegration, their passage once through half of the solar system would have spread their energies 20 times more than the observed width of peak A. Other difficulties and weaknesses of the argument appear in the discussions published after van Flandern's paper.

2. The arrow of time for cometary evolution is also reversed in a different way by Alfvén and Arrhenius (1976). This must be understood in the framework of their independent analysis of the origin and evolution of the solar system, which differs completely from the traditional paradigm. To explain primeval condensations by plasma effects, gas densities are kept low

TABLE II

a. COMETS WERE CONDENSED FROM THE SOLAR NEBULAR AND EJECTED LATER INTO OORT'S CLOUD

Place of Origin	Mechanism of Ejection	Cause	Critical Remarks
2 to 4 AU (Oort 1950)	perturbations	Jupiter	place too hot to keep ices; process is inefficient; (0.1% trapped in Oort's cloud).
5 to 40 AU (Safronov 1972)	perturbations	all giant planets	by-product of planet's accretion (1% trapped in Oort's cloud); major contributors: 1/4 Uranus, 1/2 Neptune
38-50 AU (Kuiper 1951)	perturbations	Pluto	"raking range" of Pluto is large; if Pluto lost satellite, process could be recent
300-3000 AU (Cameron 1977)	sudden weakening of attraction	fast mass loss of solar nebula	faraway ring of comets and short time-scale of mass loss both needed.

b. COMETS WERE RECENTLY CONDENSED *IN SITU*, OR CAPTURED

Place of Origin	Mechanism	Cause	Critical Remarks
10 to 1000 AU (Lyttleton 1953)	gravitational focussing of sun	passage through interstellar cloud	model at complete variance with empirical evidence
10 to 100 AU (Reeves 1974)	shock wave	solar wind	no model; no direct collisions possible
the Oort Cloud (McCrea 1975)	gravitational compression of interstellar clouds	passage through galactic arm	no model; unrealistic compression needed to get collisions.
interstellar medium (Witkovski 1972, Greenberg 1974)	capture from interstellar field of comets	Witkovski's cloud of gas bound to sun	could explain depletion of C,N,O in interstellar medium; but unrealistic mass of Witkovski's cloud (10^3 suns)

c. REVERSING THE ARROW OF TIME IN COMETS' TRADITIONAL EVOLUTION

Place of Origin	Mechanism	Cause	Critical Remarks
asteroids' belt (19th century and van Flandern 1977)	disruption of former planet	unknown; 5 x 10^6 years ago	does not explain observed flux size of "new" comets in narrow energy range
Jupiter (Proctor 1884)	former big eruptions	volcanoes of Jupiter, active 10^6-10^7 years ago	-same-
all giant planets and their satellites (Vsekhsvyatsky 1977)	grandiose eruptions going on	volcanoes observed everywhere	-same-
all meteoroid trajectories (Alfvén and Arrhenius 1976)	accretion from meteoroid streams	focussing by inelastic collisions	-same- also contradicts that older streams are observed wider

and evolutionary times grow in proportion. Later, Kepler motions perturbed by inelastic collisions are given a paramount importance (jet stream effect) in the accretion of asteroids and comets from meteoroids and from dust; in particular: long-period comets accreted from the long-period meteoroid streams; planetary encounters perturb these streams into short-period streams, that accrete eventually into short-period comets (the diffusion of long-period comets into short-period comets is denied on the same ground as the 19th century authors, see first paragraph of this section).

The theory does not explain quantitatively how plasma condenses into long-period meteor streams, in transplanetary space, beyond the influence of the solar magnetic field. The extremely slow growth of particles in a low-density region is emphasized by the need of very low pressures in the inner solar nebula, in order not to quench the ionization of the plasma. The jet stream effect seems at variance with the evolution of the meteor showers (the diffusion of the meteor streams seems to grow with time). In its present state, Alfvén and Arrhenius' approach does not provide any theory for the condensation of an ultra-rarefied gas into a meteor stream, let alone a theory explaining the origin of comets.

CONCLUSIONS

Table II gives a summary of the present discussion on the hypotheses about the origin of comets. The only hypothesis that seems to survive this screening is that comets were condensed some five billion years ago, in the vicinity of the giant planets or beyond, at the outer edge of the solar nebula. The mechanism of their ejection into Oort's cloud is a necessary by-product of their growth by accretion. It is also a necessary by-product of a fast(?) mass loss of the solar nebula within some 1000 AU, if (?) a cometary ring extended beyond. Since there are two question marks in this latter hypothesis, the first ejection mechanism is more likely, although both could have participated in the building up of the Oort's cloud. Although no empirical evidence requests that the Oort cloud be older than a few million years, no convincing theory has been proposed so far that could explain its formation more recently than 4.5 billion years ago. In the present state of our ignorance, the origin of comets is linked with the origin of the solar system, and a better understanding of their chemistry would probably give important clues on the condensation of the early solar nebula.

Partial support of NSF Grant AST 76-07116 and of a NASA grant is gratefully acknowledged.

REFERENCES

Alfvén, H., and Arrhenius, G. 1976, *Evolution of the Solar System*, NASA-SP-345, Washington, D.C.

Anders, E., Hayastu, R., and Studier, M. H. 1977, this volume.

Brownlee, D. E., Rajan, R. S., and Tomandl, D. A. 1977, this volume.

Cameron, A.G.W. 1972, *On the Origin of the Solar System*, p. 57, edit. Reeves, publ., CNRS, Paris.

Cameron, A.G.W. 1977, NATO's Newcastle Meeting on *The Origin of the Solar System;* edit. Dermott (in press).

Crommelin, A. 1910, *Scientia, 7.*

Delsemme, A. H. 1973, *Astron. Astrophys. 29,* 377.
Delsemme, A. H. 1975, in "Comet Kohoutek," p. 219, NASA-SP-355.
Delsemme, A. H. 1976, *Mem. Soc. Roy. Sci. Liège, 6 série. , 9,* 135.
Delsemme, A. H. 1977, this volume.
Delsemme, A. H., and Miller, D. 1971, *Planet. Space Sci., 19,* 1229.
Delsemme, A. H., and Rud, D. 1977, this volume.
Everhart, E. 1972, *Astrophys. Letters, 10,* 131.
Everhart, E. 1977, this volume.
Everhart, E., and Raghavan, N. 1970, *Astronom. J., 75,* 258.
Greenberg, J. M. 1974, *Astrophys. J. Letters, 189,* L81.
Greenberg, J. M. 1977, this volume.
Kazimirchak-Polonskaya, E. I. 1972, IAU Symposium No. 45, p. 373, publ. Reidel-Dordrecht.
Kuiper, G. P. 1951, in *Astrophysics: A Topical Symposium,* p. 357; edit. Hynek, publ. McGraw-Hill, New York.
Lagrange, J. L. 1814, Addendum to *La Connaissance des Temps.*
Lyttleton, R. A. 1953, *The Comets and Their Origin,* Cambridge Univ. Press.
Lyttleton, R. A. 1968, *Mysteries of the Solar System,* Clarendon Press, Oxford.
Marsden, B. G. 1975, Catalogue of Cometary Orbits, Smithson. Astrophys. Obs., Cambridge, Mass.
Marsden, B. G. 1977, this volume.
Marsden, B. G., and Sekanina, Z. 1973, *Astron. J., 78,* 1118.
McCrea, W. H. 1975, *Nature, 255,* 96 and 607.
Millman, P. M. 1977, this volume.
Newton, H. A. 1893, *Mem. Nat. Acad. Sci., 6,* 7.
Nezhinsky, E. M. 1972, IAU Symposium No. 45, p. 335, edit. Chebotarev *et al.*, publ. Reidel/Dordrecht, Holland.
Oort, J. H. 1950, *Bull. Astron. Instit. Netherlands, 11,* 91.
Öpik, E. J. 1966, *Mem. Soc. Roy. Sci. Liège, Sér. 5, 12,* 523.
Öpik, E. J. 1975, *Irish Astron. J., 12,* 1.
Proctor, R. A. 1884, *Knowledge, 6,* 126.
Reeves, H. 1974, *Nature, 248,* 398.
Safronov, V. S. 1972, IAU Symposium No. 45, p. 329, edit. Chebotarev *et al.*, publ. Reidel/Dordrecht, Holland.
Shimizu, M. 1976, *The Study of Comets,* part 2, p. 673, publ. NASA-SP-393.
Tisserand, F. 1890, *Bull. Astron.* 7, 453 (Paris).
Vanysek, V. 1977, this volume.
Van Flandern, T. C. 1977, this volume.
Van Woerkom, A.J.J. 1948, *Bull. Astron. Inst. Netherlands, 10,* 445.
Vsekhsvyatsky, S. K. 1933, *Astron. Zh., 10,* 8.
Vsekhsvyatsky, S. K. 1962, *Astron. Vestn, 39,* 290.
Vsekhsvyatsky, S. K. 1977, this volume.
Weidenschilling, S. J. 1975, *Astron. J., 80,* 145.
Whipple, F. L. 1950, *Astrophys. J., 111,* 375.
Whipple, F. L. 1962, *Astron. J., 67,* 1.
Whipple, F. L. 1964, *Proc. Nat. Acad. Sci., 51,* 711.
Whipple, F. L. 1972, in IAU Symposium No. 45, p. 401, edit. Chebotarev *et al.*, publ. Reidel/Dordrecht, Holland.
Whipple, F. L. 1975, *Astron J., 80.* 525,