# PART II
# THE ORBITAL EVOLUTION OF COMETS



https://doi.org/10.1017/S0252921100069943 Published online by Cambridge University Press

1

# ORBITAL DATA ON THE EXISTENCE OF OORT'S CLOUD OF COMETS

B. G. MARSDEN

*Oort's work on the cometary cloud is reviewed and extended using new data from 99 comets with high-quality orbits. These data clearly show the pronounced pile-up of the "original" reciprocals of the semi-major axes at values of less than 0.000100 $AU^{-1}$. This concentration is found to be even more striking for comets of large perihelion distance, and the possible significance of this is discussed. Lyttleton's criticisms of the concept of the Oort cloud (or, as only he calls it, "shell") are reviewed and dismissed as largely irrelevant. A set of data on 96 comets with orbits of second-class quality is also considered.*

Russell (1920) seems to have been the first to realize that there was something strange about the distribution of the sizes of the orbits of the known long-period comets. He discussed briefly the fact that as the comets continue to revolve around the sun under the gravitational attraction of the planets there ought to be a gradual diffusion of the values of 1/a, the reciprocal of the semimajor axis. Equal ranges in 1/a should therefore contain roughly equal numbers of comets, but this is not what is observed. The matter was later taken up in more detail by van Woerkom (1948), and although his tabulation of 1/a values merely includes the observed osculating orbits of the comets (and even includes parabolic cases), the fact that small values of 1/a are more prevalent than larger values is quite noticeable.

The very thorough treatment of this problem by Oort (1950) was based on 19 accurately determined orbits and referred, not to the osculating values of 1/a derived for an epoch close to perihelion passage, but to the so-called original values $(1/a)_{orig}$, in which allowance is made for the planetary perturbations acting back to the time when the comets were out beyond the orbit of Neptune, and an adjustment of the center of attraction is then made from the sun to the barycenter of the solar system. At large heliocentric distances -- but not so large that the gravitational influence of the stars become important -- a comet will travel about the barycenter in an essentially unchanging conic.

What Oort pointed out was that while 10 of the 19 comets had $(1/a)_{orig} < 0.000050$ $AU^{-1}$ and four more had $0.000050 < (1/a)_{orig} < 0.000100$, each of the next three 0.00050 - $AU^{-1}$ bands contained only one comet. Although the planetary perturbations will cause about half of the comets of small $(1/a)_{orig}$ to leave the solar system on hyperbolic trajectories, those comets that remain for another return will experience an average change of $\Delta \simeq +0.000500$ $AU^{-1}$ in 1/a, and since almost all of the comets in Oort's list have $(1/a)_{orig} < \Delta$, the obvious and basic conclusion is that almost all of them were being observed on their *first* passages near the sun. Lyttleton (1963) has criticized this conclusion on the grounds that almost all of the 19 comets had osculating orbits that were hyperbolic, so a bias toward small values of $(1/a)_{orig}$ was to be expected. While this is certainly true -- the most accurate computations available to Oort had their nucleus in the work of Strömgren (1914) and had mainly been undertaken in order to demonstrate that no comet actually originated outside the solar system -- it is important to note that Oort did include in his analysis many comets of larger 1/a, the values of $(1/a)_{orig}$ being derived with the aid of the approximate computations made by Fayet (1906).

Nevertheless, it is desirable to extend the accurate computations to a more representative set of comets. Such an extension can now be made rather easily, thanks to the tabulation by Everhart and Raghavan (1970) of the quantity $u_b$ =

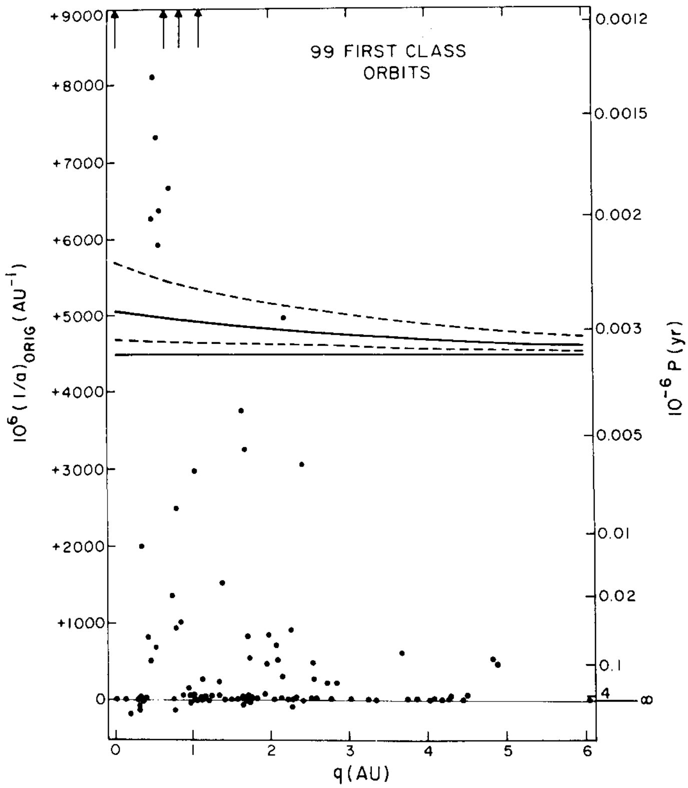


*Figure 1.* *Plot of $10^6$ $(1/a)_{orig}$ versus q for 99 comet orbits of first-class quality. The errors in the positions of the points are rarely larger than the points themselves. The horizontal line near the center of the diagram represents the fact that, because of planetary perturbations, 50 percent of the values of 1/a can be expected to increase during one revolution of the comets about the sun, and 50 percent to decrease. The solid curve above it shows the average expected change of 1/a of the comets for which this quantity increases, while the broken curves indicate the perturbations which should be exceeded by 20 percent and 80 percent of these comets, respectively. The root-mean-square perturbation extends from the upper broken curve near q = 0 to the solid curve near q = 5 AU. The right-hand scale gives the equivalent revolution period P.*

$(1/a) - (1/a)_{orig}$ for all the long-period comets observed between 1800 and 1970; values of $u_b$ for more recent comets are included in the annual comet reports in *Quart. J. Roy. Astron. Soc.* Brady (1970) has in fact derived original orbits for 143 comets, but his compilation makes indiscriminate use of some very accurate orbit determinations and some extremely uncertain ones.

Fig. 1 is a plot of $(1/a)_{orig}$ versus perihelion distance q for the 99 long-period comets (*i.e.*, period > 200 years) of highest quality. The criterion used for defining the orbit quality depends on the mean error of 1/a, the length of the arc covered by the observations and the number of planets whose perturbations were taken into account. This criterion tends to favor large-q comets, and it also favors our own orbit determinations, which are made according to a uniform procedure that utilizes individual observations rather than the normal places so commonly used in the generally rather haphazard and heterogeneous collection of orbits by various earlier authors.

The "Oort effect" is readily apparent in Fig. 1, for fully 58 percent of the $(1/a)_{orig}$ values are less than $+0.000100$ $AU^{-1}$. The spread in $(1/a)_{orig}$ is much greater for q < 1 AU than for q > 3 AU, but there are still as many as 44 percent of the former group that have $(1/a)_{orig} < +0.000100$ $AU^{-1}$. Examination of the Everhart-Raghavan results suggests that the dispersion in 1/a during a single perihelion passage decreases from a mean value (without regard to sign) of $|\bar{\Delta}| \approx 0.000550$ $AU^{-1}$ at q = 0 to $|\bar{\Delta}| \approx 0.000150$ $AU^{-1}$ at q = 5 AU. (It is necessary to consider $|\Delta|$, rather than $\Delta$, because positive and negative changes are equally probable, and the rate of dispersion is significantly modified by the negative values of $\Delta$ when $1/a \geq |\Delta|$). One concludes from Fig. 1 that it is extremely unlikely that any of the comets having q > 3 AU can have made more than two passages inside the orbit of Jupiter, although it is possible that a *few* of the comets having q < 1 AU have made 100 or more such passages.

As already stated, our accuracy criterion favors the large-q comets. Owing to the so-called nongravitational forces that are certainly observed to affect the motions of the regularly-returning short-period comets (Marsden 1968), the large-q $(1/a)_{orig}$ values should be the ones that are most reliably determined (Marsden and Sekanina 1973). Although one can rarely detect the influence of nongravitational forces on the motion of a single-apparition, long-period comet, there is no reason to believe that such forces, which are apparently due to the vaporization of a comet's water-ice content by solar radiation (Whipple 1950, Marsden *et al.* 1973, Delsemme 1974), are not acting. Comets of q ≥ 3 AU appear to be immune from such forces, for their water ice would remain perpetually frozen. On the other hand, the $(1/a)_{orig}$ values derived for comets of small q will be consistently too small, typically by, say, 0.000020 $AU^{-1}$ (Marsden *et al.* 1973), but sometimes by indeterminately large amounts. There is thus certainly no reason to believe that the negative $(1/a)_{orig}$ values in Fig. 1 refer to interstellar comets (Whipple 1975, Sekanina 1976a). Indeed, one of the negative values, that at $10^6$ $(1/a)_{orig} = -99 \pm 6$, q = 0.316 AU, belongs to comet 1957 III (Arend-Roland), a rare case in which it can be shown that a purely gravitational orbit determination gives an unsatisfactory fit to the observations: computations made under various assumptions for the nongravitational effects cause $10^6$ $(1/a)_{orig}$ to increase to as much as $+58 \pm 14$ (Marsden *et al.* 1973).

Oort (1950) explained the existence of large numbers of comets of small $(1/a)_{orig}$ by supposing that these comets had been thrown from orbits that were originally circular into orbits of small q by the gravitational attractions of passing stars. The general problem of stellar perturbations on comets (and meteors) had first been tackled by Öpik (1932). It was in the celebrated 1950 paper that the concept of the "Oort cloud" was introduced. In putting the outer edge of the cloud at 200,000 AU from the sun [corresponding to $(1/a)_{orig} = +0.000010$ $AU^{-1}$], Oort recognized that comets that were more distant than that would amost certainly have escaped; Nezhinskij (1972) has estimated that a cometary cloud of this radius would have a halflife of at least $1.1 \times 10^9$ years. Oort



https://doi.org/10.1017/S0252921100069943 Published online by Cambridge University Press

(1963) later reduced the presumed radius of the cloud to some 100,000 AU, and Sekanina (1968), in an attempt to examine the stability of the cloud under the action of the known nearby stars, concluded that in some directions the radius could scarcely be larger than this. Our Fig. 1 confirms that 100,000 AU is undoubtedly a very reasonable upper limit for the original aphelion distances Q in all directions, although there really seem to be relatively few comets that have come in from $Q \geq 50{,}000$ AU.

Oort (1963) has also made the rather unfortunate statement that the cloud will "not contain many comets" with $Q \leq 30{,}000$ AU. This has prompted Lyttleton (1968, 1974) to interpret Oort's model as referring to a cometary "shell," rather than to a cloud, and he goes on to claim that the shell does not exist. While it is true that the word "shell" was used by some nineteenth-century writers (*e.g.*, Young 1900), the word has never in fact been used by Oort -- LePoole and Katgert (1968) have also remarked on Lyttleton's propensity to attribute to Oort statements that he never made -- and whether the comets form a shell or exist in some rather more substantial configuration is irrelevant. In his fundamental paper Oort (1950) made very little mention of comets of smaller Q. He did not say that such comets did not exist -- only that stellar perturbations would not particularly influence them. There might be comets with $Q \leq 20{,}000$ - $30{,}000$ AU and $q \geq 100$ AU, but there is no way whereby we can detect them: there would be no planetary encounters to cause Q to increase to distances where stellar perturbations could then decrease q to values low enough for the comets to be observed! The existence or nonexistence of such comets is of very little concern to us. The matter becomes of interest only in speculations on the origins of comets. Although it is not basic to his theory, Oort favored the idea that comets actually originated together with the asteroids in the asteroid belt. In such an instance, and also if one takes the slightly different view that comets originated in the vicinity of the outer planets (Kuiper 1951), it would be difficult to produce large numbers of comets with both $Q \leq 20{,}000$ - $30{,}000$ AU and $q \geq 100$ AU. More recent hypotheses (*e.g.*, Cameron 1973) consider the possibility that comets did form at vast distances from the sun, and if that is the case, there could exist very large numbers of comets in essentially stable orbits in this range: it can be estimated that there may be more than $10^{14}$ such unobservable, irrelevant comets (Whipple 1975), a number that is three orders of magnitude greater than the entire cometary population discussed by Oort!

Lyttleton's criticisms also make much of other irrelevancies. He suggests that one might want to consider the "volume density of aphelion points" and claims that one must therefore introduce the factor $a^{3/2}$ to take into account the varying frequency with which comets pass through perihelion. He questions the use of $1/a$ in frequency distributions and seems to be advocating that one ought instead to use $a$ itself. Elementary celestial mechanics yields for the perturbations in the semimajor axis $a$ of the orbit of a comet a differential equation of the form $da/dt = a^2H$, where H is a well-defined quantity that depends largely on the changing positions of the comet and the objects that perturb it. In the inner part of the solar system these perturbing bodies are the major planets, and H can be accurately calculated as a function of the time t. Integration of the differential equation requires the introduction of the multiplicative factor $a^2$: for a long-period comet this is not only an almost infinite quantity but in practice a very ill-defined almost infinite quantity. The differential equation $d(1/a)/dt = -H$ is mathematically equivalent to that for $da/dt$, but it is far easier to integrate in practice. If one insists on using the semimajor axis, the process of reduction to the barycenter also requires the use of a quantity that is proportional to $a^2$, whereas use of the reciprocal form easily enables one to calculate with high accuracy the difference between the osculating and the original orbits, as well as the difference between the "future" (also referred to the barycenter of the solar system) and the original



https://doi.org/10.1017/S0252921100069943 Published online by Cambridge University Press

orbits, even when the observations are insufficient for one to detect any deviation of the comet's osculating orbit from a parabola.

Some researchers, notably Fayet (1906) and Brady (1970), have considered the perturbations in the eccentricity e, rather than in 1/a. There is no *a priori* reason why they shouldn't, and if stellar perturbations can indeed be ignored, the relevant differential equation is $de/dt \approx qH$. Use of an e-distribution has the effect of diminishing, not only the dependence on q of the expected dispersion of the perturbations, but also some of the difference between the spread of original orbits observed for comets of small and large q. On the other hand, it might seem that there is some significance in the fact that large-q comets show a preponderance of *osculating* eccentricities that are substantially greater than unity, and analysis of the effects of possible stellar perturbations would become a more complex problem. On the whole, it does seem that 1/a is the most convenient quantity to use, but not just because "[it] is the energy" (Lyttleton 1974).

Lyttleton (1974) has criticized our point (Marsden and Sekanina 1973) that if we *do* humor him by making a conversion of our results from 1/a to *a*, there is a tendency for a peak in the distribution to appear in the range a = 20,000 - 25,000 AU. In "correcting" our figures for an "average mean error...[of] ± $13 \times 10^{-6}$ $AU^{-1}$" he made the strange assumption that all points within this range from the specified value of $(1/a)_{orig}$ are equally probable -- and found the peak to disappear. If one adopts the more usual convention that the probability diminishes away from the specified value according to a Gaussian error curve, the conclusion (which is in any case irrelevant) is very close to what we simply derived by ignoring the errors entirely.

The whole point of this type of cometary research is to try to explain why the distribution of original values of 1/a (or e) is the way it is. If Lyttleton would simply accept Fig. 1 as an observed fact and join us in our quest for an explanation of it, worthwhile progress might be made. Oort has rather qualitatively described a plausible mechanism for producing an influx of comets having 1/a only slightly greater than zero, but the problem of considering the interplay of stellar and planetary perturbations has not yet been tackled with the help of the powerful computing techniques that are now available.

Furthermore, although we may have a plausible source for the distribution in Fig. 1, we have not yet discussed a sink, and in this respect Oort's (1950) hypothesis is perhaps his most important contribution to the whole discussion. [It is just possible, of course, that comets have been prevalent in the solar system (or at least in its inner part) for less than 0.1 percent of the lifetime of the solar system -- a comet that comes in from an aphelion distance of 100,000 AU does so in less than 6 million years].

The general approach to the problem demands that we consider it to be one of maintaining an essentially steady state. Oort (1950) therefore suggested that the reason there are so few comets of larger $(1/a)_{orig}$ is simply that comets disintegrate: they fade with each successive passage near the sun, thus become less likely to be discovered, and eventually they completely disperse into meteoroids. The nongravitational forces we have discussed are certainly consistent with the idea that comets are constantly losing material. Kendall (1961), Shtejns (1961) and Whipple (1962) have attempted to refine the diffusion calculations to include the effects of cometary dissipation, this last author concluding that although as many as 50 percent of the comets of smallest 1/a fail to survive more than one approach to the sun, some comets have lifetimes on the order of $10^4$ revolutions.

Oort and Schmidt (1951) showed that there do seem to be physical differences between "new" comets that may be making their first approach to the sun and those that have passed near the sun many times before: these suggested differences were mainly that the brightness of a first-timer varies according to a lower inverse power of heliocentric distance and that its spectrum is more likely to show a



https://doi.org/10.1017/S0252921100069943 Published online by Cambridge University Press

strong continuum, particularly if $q \geq 1$ AU. There are many exceptions, however, and the problem requires careful reexamination. Would comet 1973 XII (Kohoutek), located in Fig. 1 at $10^6$ $(1/a) = +20 \pm 3$ AU$^{-1}$, $q = 0.142$ AU and initially detected while still out near the orbit of Jupiter, have failed us at and after its perihelion passage if this had *not* been its first approach to the sun? Were comets 1957 III (Arend-Roland) and 1941 I [Cunningham; $10^6$ $(1/a) = \pm 1 \pm 10$ AU$^{-1}$; $q = 0.368$ AU] more satisfactory affairs (in spite of complaints made at the time about the latter comet) simply because they had larger perihelion distances -- and, if so, why did comet 1962 III [Seki-Lines; $10^6$ $(1/a) = +25 \pm 12$ AU$^{-1}$; $q = 0.031$ AU] become a reasonably bright object? Did the early recognition that comets 1970 II [Bennett; $10^6$ $(1/a) = +7334$ AU$^{-1}$; $q = 0.538$ AU] and 1975n (West; not shown in Fig. 1) were *not* "new" provide a reason to expect these comets to become spectacular? Is it an accident that the largest values of $(1/a)_{orig}$ for the comets of $q > 2$ AU and $q > 3$ AU, respectively, belong to comets 1962 VIII (Humason) and 1927 IV (Stearns), which are among the three or four intrinsically brightest and spectroscopically most unusual comets of the present century?

Since water-ice vaporization is the presumed cometary dissipation mechanism, some comment should be made on the fact that it is the comets of $q > 3$ AU, where this mechanism is not applicable, that show the greatest propensity to disappear after only one approach to the sun. It is important to note that comets are recognizable at large heliocentric distances only because of the vaporization of some substance that is more volatile than water. On its first approach to the sun, some of this volatile material perhaps exists in the exterior regions of the nucleus. Since the water-ice content of a large-q comet would not vaporize, further supplies of volatile material would be trapped inside the ice and unable to migrate to the surface of the nucleus. On its second and subsequent returns to the sun a large-q comet could therefore simply be an inert frozen water-ice ball that would almost certainly escape detection (Marsden and Sekanina 1973). Although not readily observable, these objects, their reciprocal semi-major axes gradually dispersing toward larger values, would continue to exist more or less indefinitely, many of those of low orbital inclination eventually being captured by Jupiter into short-period orbits (Everhart 1972) and, if their perihelion distances then decrease, the objects would become active and observable short-period comets.

Finally, it is instructive to inspect Fig. 2. This is a plot of the quantities shown in Fig. 1, but it applies to the 96 cometary orbits we have judged as of second-class quality. The error bars on $(1/a)_{orig}$ are still not excessively large, but the distribution is seen to be very different from that of Fig. 1. The "Oort effect" is still noticeable, but the spread in $(1/a)_{orig}$ is much more substantial than for the first-class orbits and much more uniform in q -- although only five orbits of $q > 2$ AU qualify for inclusion. We have already suggested that some of the highly dispersed orbits in Fig. 1 refer to unusually spectacular comets -- to enormous comets that must have put on stupendous displays at their first approaches to the sun. Fig. 2 presumably contains the more run-of-the-mill comets that were too faint to be observed long enough to qualify among the comets considered for Fig. 1.

Fig. 2. also includes some comets with apparently negative $(1/a)_{orig}$ that must have been influenced strongly by nongravitational forces: comets 1960 II [at $10^6$ $(1/a) = -402$ AU$^{-1}$; $q = 0.504$ AU] and 1971 V [$10^6$ $(1/a) = -293$ AU$^{-1}$, $q = 1.233$ AU] are two more cases where gravitational solutions are known to give unsatisfactory representations of the observations, and where nongravitational solutions are preferable and, furthermore, yield original orbits that are elliptical. The pair of even more negative $(1/a)_{orig}$ values in Fig. 2 refer to 1975q and 1955 V, small comets that perhaps experienced even larger nongravitational forces (comet 1955 V was also observed to split). Comet 1944 I (Marsden *et al*. 1973) has been mentioned as a faint comet that possibly experienced very large nongravitational effects. Large differential nongravitational perturbations



https://doi.org/10.1017/S0252921100069943 Published online by Cambridge University Press

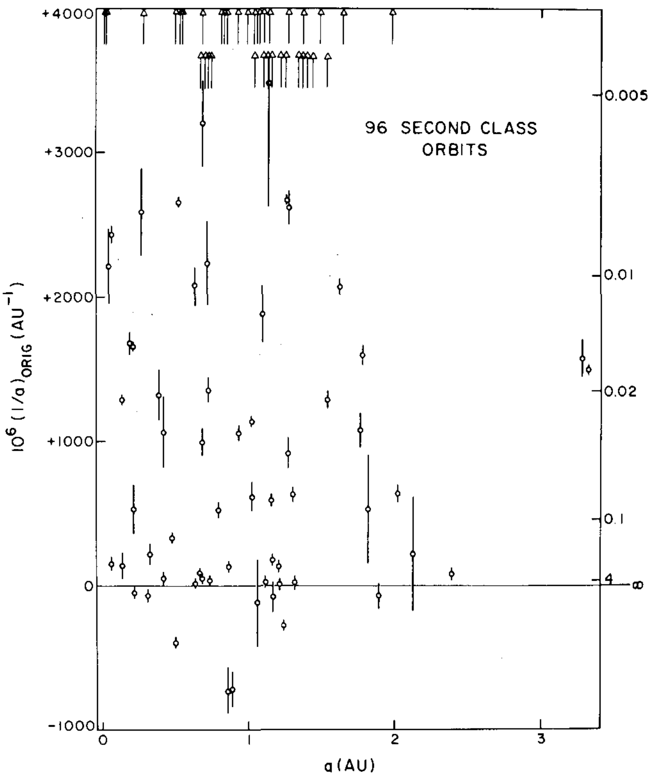


*Figure 2. Plot of $10^6$ $(1/a)_{orig}$ (and $10^{-6}$ P) versus q for 96 comet orbits of second -class quality. The bars represent mean errors. The upper set of arrows refers to $10^6$ $(1/a)_{orig}$ values > 9000 $AU^{-1}$ (the limit of Fig. 1), the lower set to values > 4000 $AU^{-1}$.*

act on the components of comets that have split (Sekanina 1976b), but the fact that each of these components (*e.g.*, in the case of comet 1975n) sometimes sports its own coma and tail is a strong indication that such comets can exist, at least for short intervals of time. These comets with very large nongravitational forces would be intrinsically faint and rarely observed well (unless they made a close approach to the earth), but their probable existence introduces a new dimension into the analysis and interpretation of the Oort-cloud phenomenon.

## ACKNOWLEDGEMENTS

I thank Zdenek Sekanina for making several of the new orbit determinations required for this work and Edgar Everhart for providing improved values of the quantity necessary to derive the "original" orbits.

## REFERENCES

Brady, J. L. 1970, *Astron. J.*, *75*, 1052.
Cameron, A.G.W. 1973, *Icarus*, *18*, 407.

Delsemme, A. H. 1974, *IAU Colloq. No. 22,* p. 313.
Everhart, E. 1972, *Astrophys. Lett., 10,* 131.
Everhart, E., and Raghavan, N. 1970, *Astron. J., 75,* 258.
Fayet, G. 1966, *Ann. Obs. Paris Mem., 26,* A.1.
Kendall, D. G. 1961, *Proc. Fourth Berkeley Symp. Math. Stat. Probl., 3,* 99, 121.
Kuiper, G. P. 1951, in *Astrophysics,* edited by J. A. Hynek (McGraw-Hill, New York), p. 457.
LePoole, R. S., and Katgert, P. 1968, *Observatory, 88,* 164.
Lyttleton, R. A. 1963, in *The Nature of Comets,* by N. B. Richter (Methuen, London), p. xvii.
Lyttleton, R. A. 1968, *Monthly Notices Roy. Astron. Soc., 139,* 225.
Lyttleton, R. A. 1974, *Astrophys. Space Sci., 31,* 385.
Marsden, B. G. 1968, *Astron. J., 73,* 367.
Marsden, B. G., and Sekanina, Z. 1973, *Astron. J., 78,* 1118.
Marsden, B. G., Sekanina, Z., and Yeomans, D. K. 1973, *Astron. J., 78,* 211.
Nezhinskij, E. M. 1972, *IAU Symp., No. 45,* p. 335.
Oort, J. H. 1950, *Bull. Astron. Inst. Neth. 11,* 91.
Oort, J. H. 1963, in *The Moon, Meteorites and Comets; The Solar System,* edited by B. M. Middlehurst and G. P. Kuiper (U. Chicago Press, Chicago), vol. 4, p. 665.
Oort, J. H., and Schmidt, M. 1951, *Bull. Astron. Inst. Neth., 11,* 259.
Öpik, E. 1932, *Harvard Repr.,* No. 79.
Russell, H. N. 1920, *Astron. J., 33,* 49.
Sekanina, Z. 1968, *Bull. Astron. Inst. Czech., 19,* 223.
Sekanina, Z. 1976a, *Icarus, 27,* 123.
Sekanina, Z. 1976b, *Icarus,* (to be published).
Shtejns, K. A. 1961, *Astron. Zh., 38,* 107.
Strömgren, E. 1914, *Publ. Købenbavns Obs.,* No. 19.
van Woerkom, A.J.J. 1948, *Bull. Astron. Inst. Neth., 10,* 445.
Whipple, F. L. 1950, *Astrophys. J., 111,* 375.
Whipple, F. L. 1962, *Astron. J., 67,* 1.
Whipple, F. L. 1975, *Astron. J., 80,* 525.
Young, C. A. 1900, *General Astronomy,* (Ginn, Boston), revised edition, p. 455.

*DISCUSSION*

*GOLD: If stellar perturbations are sufficient to supply visible comets at the rate now seen, one can estimate the explusion to infinity that would equally be caused by the same perturbations. This would seem to be a large proportion and thus one needs to increase the necessary number in the cloud by another large factor. If this factor is very large it will make a steeply decreasing rate of cometary appearances, and the rate at earlier times would have been very much higher than now. A limit can be placed on this from the study of impact cratering on the Earth and Moon.*

*SINGER: Could you comment on the possible exhaustion of comets in the Oort cloud. In other words, could the incidence of comets into the solar system have been much greater some eons ago?*

*MARSDEN: It is certainly possible that there were many more comets in the past. The available orbital data just cannot provide an answer to the question one way or the other.*



https://doi.org/10.1017/S0252921100069943 Published online by Cambridge University Press