ASPECTS OF PLANETARY FORMATION

by

STUART JOHN WEIDENSCHILLING

B.S., M.I.T., 1968

M.S., M.I.T., 1969

SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

December, 1975
(i.e. February 1976)

Signature of Author_____
Department of Earth and Planetary Sciences

Certified by _____
Thesis Supervisor

Accepted by _____
Chairman, Departmental Committee
on Graduate Students

ASPECTS OF PLANETARY FORMATION

by

Stuart John Weidenschilling

ABSTRACT

The masses and orbital spacings of the planets are the result of both the original structure of the solar nebula and the process of planetary formation. Reconstruction of the nebula from the equivalent solar-composiion masses of the planets shows that Venus, the earth, and the giant planets form a smooth trend of nebular surface density with heliocentric distance. Mercury, Mars, and the asteroid belt lie well below this trend, indicating mass deficiencies. Mercury's mass may be the result of incomplete condensation at high temperatures. The zones of Mars and the asteroids appear to have lost most of their orginal mass.

A self-consistent model for planetary accretion by purely gravitational forces is developed. It is shown that an initial relative velocity of planetesimals does not significantly affect the time scale for accretion. The first three terrestrial planets accrete on a time scale of 10^8 yr; Mars requires mor than 10^9 yr. This time scale for Mars cannot be ruled out, but is not supported by the lunar cratering record.

A planetesimal in an orbit which crosses that of a planet may collide with the planet, or be ejected from the solar system by a close encounter. A method is developed for computing the probabilities of these fates. The method avoids the use of Öpik's approximations, and produces significantly different results. Ejection is possible above a certain critical relative velocity, and for encounters with a massive planet is much more probable than collision. A scenario is developed in which mass was removed from the zones of Mars and the asteroids by a bombardment of planetesimals perturbed from Jupiter's

zone. The critical velocity for ejection corresponds to a minimum perihelion just outside the earth's orbit. The bombarding planetesimals are ejected from the solar system by Jupiter without reaching the zones of the other terrestrial planets. The accretion of Mars is interrupted by the bombardment. Sweep-up of the resulting debris accounts for late cratering of the moon and Mercury.

The encounter theory developed here is also applicable to the problems of origin of the comet cloud, capture of short-period comets by the giant planets, and cratering histories of the terrestrial planets.

Thesis Supervisor:

      John S. Lewis, Ph.D.
Title: Associate Professor of Geochemistry

Acknowledgments

I am grateful to many people for their assistance in the preparation of this thesis and of its author.

My advisor, Prof. John Lewis, has been of great help, not only with comments, criticisms, suggestions, and financial support, but also by providing an example of healthy skepticism.

I have benefited from comments and criticisms from Aviva Brecher, A.G.W. Cameron, Clark Chapman, Larry Cox, Edgar Everhart, W.K. Hartmann, Dave Johnston, Brian Marsden, Charles Peterson, Don Skibo, W.R. Ward, G.W. Wetherill, and several anonymous referees. Some of the most helpful disagreed with my conclusions, and I accept full responsibility for any errors in this thesis.

E.J. Öpik and G.W. Wetherill kindly provided me with useful reprints and preprints.

I thank Dr. J. Hovorka for suggesting that I should not be an engineer.

I thank Prudence Young and my wife, Susan, for their efforts in typing this manuscript.

I am especially grateful to my wife, Susan, for enduring my rare bursts of inspiration and too-frequent intervals of frustration. Her help, encouragement, and prodding were essential to the completion of this thesis.

Finally, I thank God for making this topic available by creating the solar system.

Preface

When attempting to construct and test a cosmogonical
theory, we face an overwhelming fundamental problem: we
know of, and can observe, only a single planetary system.
A given feature of it may be unique, or the improbable
outcome of a random process, but we are generally forced
to accept it as typical until proven otherwise. To be
viable, any theory must allow the possibility of the
details of our system. However, it would be unwise to
accept any that purports to make them inevitable. The
demonstration that something is possible, given some set
of necessarily restrictive assumptions, is no assurance
that it actually occurred in that manner. This thesis is
an attempt to construct a simple, self-consistent explan-
ation for certain features of our solar system, and is
offered in the spirit of "if..., then...". My own work
has led me to appreciate how little we really know about
the origin of the solar syatem, and how much less we can
agree upon. The thirty-eighth chapter of Job is still
recommended reading for the would-be cosmogonist.

Much of the material in this thesis has already been
published (Weidenschilling, 1974, 1975a-d). Since these
articles are readily available, they are not appended to

this thesis. Their material appears herein what is hoped
to be logical order, with uniform notation, deletion of
superfluous material, and some afterthoughts and correct-
ions.

ASPECTS OF PLANETARY FORMATION

TABLE OF CONTENTS

LIST OF TABLES

I. MASS DISTRIBUTION IN THE PLANETARY SYSTEM AND

SOLAR NEBULA

A. Introduction

There are two basic methods for constructing models of the solar nebula. In one case, the mass and general dimensions of the nebula are specified empirically, and the physical properties calculated by considerations of force balances, energy transport, angular momentum, etc. (Ter Haar, 1950; Cameron and Pine, 1973). Such models may correspond generally to our own particular nebula, but probably cannot reproduce it in detail. Our only possible clue to the specific parameters of such a model is its capability for producing the observed distributions of planetary masses, compositions, and orbits. However, the application of this test is limited by our ignorance of the planet-forming process. The second approach starts with the planetary system and works backward. One can add hydrogen and helium to each planet in such amounts as to restore it to solar composition, and spread its mass through some region surrounding its orbit. A nebula constructed in such manner can, by definition, reproduce any given system of planetary orbits and masses. However, the nebular structure so derived is entirely ad hoc, and will be in error if planetary formation was not simply the reverse of the "spreading" process.

The two approaches are complementary, and should converge
as we improve our understanding. Kuiper (1956) constructed
a detailed nebular model by the second method. Since
then, our knowledge of planetary compostions and solar
elemental abundances has grown considerably, suggesting
a new calculation of this type. I will attempt to re-
construct empirically the mass distribution of the solar
nebula, with some estimate of the uncertainties involved
and the type of information which might be derived.

B. Computations

   We can compute the equivalent solar composition mass
of a planet if we know the mass of a major constituent
element, and its solar abundance. Iron is ideal, since
most models of the terrestrial planets have involved
determination of their Fe content. Its high condensa-
tion temperature suggests that Fe was completely
condensed in all zones, except possibly that of Mercury.
Reynolds and Summers (1969) modeled the terrestrial
planets by varying the proportions of metal and silicate
phases; the composition of each phase was assumed to be
the same for all planets. They found Fe mass fractions
of 0.68 for Mercury, 0.35 for Venus, 0.38 for the earth
and 0.26 for Mars. However, Lewis (1972) has shown
that the bulk compositions of the metal and silicate
phases vary with condensation temperature (or helio-
centric distance). In particular, the retention of sulfur

-12-

by the earth raises its mean atomic weight; the procedure of Reynolds and Summers would overestimate the terrestrial Fe abundance. Below its condensation temperature, the weight fraction of Fe in the bulk condensate decreases monotonically with temperature for any reasonable model of nebular density (Lewis, 1975). It seems probable that the earth's Fe content is lower than that of Venus; I adopt 0.33 as a reasonable estimate. Siegfried and Solomon (1974) and Johnston et al. (1974) have constructed models of Mercury and Mars, respectively, in accordance with Lewis' chemical models. Their estimates of Fe content are 0.62 for Mercury, and 0.30 for Mars, which are adopted here. New models for Venus and the earth based on these assumptions are desirable. The range of values among the various models suggest that the Fe content for each planet is known to about $\pm 10\%$ of the total amount. Use of Reynolds and Summers' values would not affect the results significantly.

Neither the mass nor composition of the asteroid belt is known with certainty. Accepting Schubart's (1974) value of 1.96×10^{-4} earth masses for Ceres, and assuming that the entire belt is a few times more massive, I adopt the figure of 5×10^{-4} earth masses. The present belt may be only a remnant of a much larger earlier population. From a study of their collisional evolution, Chapman and Davis (1975) estimate the initial population of asteroids

was some 300 times the present population, though they do not rule out a still larger figure. That amount corresponds to about 0.15 earth masses. According to Wetherill (1975a), the present size and velocity distributions in the asteroid belt are not compatible with such a large population. The majority of asteroids appear to have carbonaceous chondritic compositions (McCord and Chapman, 1975), for which an Fe content of 0.25 by weight is appropriate.

The zones through which the planetary matter should be spread cannot be determined with certainty. Lecar and Franklin (1973) suggested that the zones might lie between the inner and outer Lagrangian points of each planet. However, this would lead to wide gaps between zones, while the solar nebula must have been continuous. Their other suggestion was that the zones filled the area between adjacent planets, with the distance to a boundary proportional to the Lagrangian distance. In the absence of any detailed model for planetary accumulation, I simply take each zone boundary to lie halfway between adjacent planets. This is certainly accurate to within a factor of two. Mercury's zone is assumed to extend as far inward from its orbit as outward; this is only twice the area between its perihelion and aphelion. Mars' zone extends to the inner edge of the asteroid belt, at 2.0 AU. Adopting Cameron's (1973a) solar abundances, with an Fe weight fraction of 0.0012, the solar composition mass and surface

density, as well as the surface density of solids, σ_s, are given in Table I.

Iron is not a significant component of the giant planets. Current models are constructed of hydrogen and helium in solar proportions, with cores of rocky and/or icy matter. Detailed models of Jupiter and Saturn have been computed by Podolak and Cameron (1974), and by Zharkov et al. (1975). Their results are in general agreement that both planets are enriched in heavy elements with respect to solar composition. The degree of enrichment is uncertain; the computed values depend rather strongly on the assumed equation of state, temperature boundary conditions. and H/He ratio. The enrichment factor for acceptable models of Jupiter lies in the range from 2 to 40; for Saturn, the range is about 10 to 60.

Podolak and Cameron also modeled Uranus and Neptune. Using their value of 0.00343 for the weight fraction of "rock" (metal and silicates in solar proportion) in solar composition, their core masses correspond to about 1000-2000 earth masses of solar material. Makalkin's (1973) models of Neptune, and unpublished calculations for both planets by Reynolds and Summers (1973) correspond to some-what smaller masses. The adopted range of values in Table II covers all models, without any choice of a "best" value. The zone boundaries were chosen in the same way

Table I

Terrestrial planet zones: masses and surface densities

	Mass (Earth=1)	Fe wt.%	Mass (solar comp.)	Zone (AU)	σ_s, g-cm^{-2} (solids)	σ, g-cm^{-2} (solar comp.)
Mercury	.053	62	28	.22-.56	1.7	900
Venus	.815	35	225	.56-.86	16	4600
Earth	1	33	270	.86-1.26	10	2700
Mars	.107	30	27	1.26-2.0	0.4	95
Asteroids (present)	.0005?	25	0.1	2.0-3.3	.0006	0.13
(original)	.15?	25	30	2.0-3.3	0.2	40

Table II

Giant planet zones: masses and surface densities

	Mass (Earth=1)	Mass (solar comp.)	Zone (AU)	σ, g-cm^{-2} (solar comp.)
Jupiter	318	600-12000	3.3-7.4	120-2400
Saturn	95	1000-6000	7.4-14.4	55-330
Uranus	14.6	700-2000	14.4-24.7	15-40
Neptune	17.2	800-2000	24.7-35.5	10-25

as for the terrestrial planets; the outer edge of the asteroid belt marks the inner boundary of Jupiter's zone, and Neptune's zone is assumed to extend as far outward from its orbit as inward. The surface density of solids is not given in Table II. In the zones of the giant planets, H_2O, NH_3, and CH_4 ices would be present in varying proportions, depending on the local temperature and pressure. Using Podolak and Cameron's abundances, the "rock" surface density is 0.00343 times the solar composition value. If H_2O is fully condensed, σ_s is 3.0 times larger; if NH_3 and CH_4 are also condensed, it is 4.6 times the value for "rock" alone. If H_2O is fully condensed in Jupiter's zone, σ_s could equal that in the earth's zone. If planetesimals form by the gravitational instability mechanism of Goldreich and Ward (1973), their masses are proportional to $\sigma_s a^6$. For comparable values of σ_s, the planetesimals in Jupiter's zone are more than 10^4 times as massive as in the earth's zone, possibly approaching diameters of 100 km. Further accretion probably proceeded more rapidly than in the earth's zone; the suggestion that Jupiter formed before the terrestrial planets (Weidenschilling, 1974) seems reasonable. Jupiter's heliocentric distance may have been determined by the inner boundary of H_2O condensation in the solar nebula.

C. Results

The data of Tables I and II are plotted in Figure 1.

The vertical "error bars" show only the uncertainty in the planetary compositions. The horizontal bars show the zone widths, and mark the mid-range of compositions. Uncertainties in the solar abundances of heavy elements do not seriously affect Fig. 1. A change in the adopted solar Fe/H ratio would shift all σ values for the terrestrial planets by the same factor. The values for the giant planets are based on the abundances of "rock-forming" elements (principally Si and Fe) in solar-composition gas. Since the solar Fe/Si ratio is known to about $\pm20\%$, the uncertainty in normalization between the two groups is small compared with the compositional uncertainties for the giant planets. Better models of Jupiter and Saturn would allow a considerable improvement in the nebular model. A diagram similar to Fig. 1, but giving only the surface densities of "heavy elements", and without error estimates, appears in Lecar and Franklin (1973).

Even at the low resolution of these results, two features are apparent in the reconstructed "nebula". There are two regions of low density, one in Mercury's zone, the other in the region of Mars and the asteroids. If the Mars-asteroids gap is smoothed over, the general trend from Venus to Neptune is roughly $\sigma \propto a^{-3/2}$. This profile resembles quite closely the nebular models of Cameron and Pine (1973), though the values of σ shown here are smaller

by nearly two orders of magnitude. Cameron has recently lowered his estimate of nebular surface densities (Cameron and Pollack, 1975), but detailed models are not yet available. The uncertainties in planetary compositions and solar abundances allow a nebular mass in the range of about 0.02 to 0.1 solar masses for our reconstructed nebula, if all the heavy elements ended up in the planets.

The low surface density in Mercury's zone is easily explainable. The temperature in at least part of the zone was probably too high to allow complete condensation of Fe. Mercury is enriched in Fe relative to silicates; its bulk composition is compatible with such a high-temperature origin (Lewis, 1972). Proximity to the sun would also favor loss of matter by the Poynting-Robertson and Yarkovsky effects, but these probably could not have caused significant mass loss in the time available. An initially higher surface density in Mercury's zone would imply rapid accretion of the planet, on a time scale of about 10^7 years (Weidenschilling, 1974). The minimum in the zones of Mars and the asteroids is probably due to the removal of matter, rather than a local minimum in the nebular density. Even for the larger original population of asteroids suggested by Chapman and Davis (1975), the surface density is still anomalously low in the asteroid belt. Smoothing over the gap, we can estimate that Mars

and the asteroids should each have contained some two or three earth masses of solid matter.

If there were nothing at all between the earth and Jupiter, there would be no local minimum in the reconstructed nebula. However, the reconstruction procedure would then implicitly assume that matter was transported across that region as part of the normal planet-forming process. The existence of bodies in stable orbits there indicates that such transport did not occur. The work of Dole (1970) is of interest in this regard. Dole simulated the formation of planetary systems by a Monte Carlo method. He assumed an empirical solar nebula model in which the surface density decreased monotonically with distance from the sun. Planetary nuclei with a range of random orbital parameters were "injected" into the nebula and allowed to accrete dust and/or gas in their vicinity, according to certain simple assumptions. The resulting planetary systems were similar to our own in numbers and masses of planets, and mean orbital spacings. However, Dole's simulations differ from our system in one respect. When σ is calculated for his simulated planetary systems in the manner described above, the values show only small deviations from the monotonic variation in the original nebular model. There are no minima comparable to the Mars-asteroids region. The more massive planets tend to

be spaced more widely, minimizing the variations in the "reconstructed" surface density values. This result follows from his model of the planet-forming process, which does not allow removal of matter from a planet's zone. Dole's simulations, therefore, could not reproduce our solar system in detail for any values of the random input parameters.

These conclusions depend on the seemingly reasonable assumption that the solar nebula's surface density decreased monotonically with distance from its center. Some types of information cannot be recovered from any such "reconstructed nebula". Any process of mass loss is not detectable if it varied smoothly and monotonically with heliocentric distance. If the sun lost an appreciable fraction of its mass during a T Tauri phase, all planetary orbits would have expanded by the same factor, which could not be inferred from Fig. 1. If Neptune originated at about 40 AU, in accordance with Bode's Law, its distance could have decreased to its present value if it ejected cometary bodies from the solar system by gravitational perturbations in close encounters. The ejected mass required is about one third of the planet's mass (Safronov, 1967). The computed value of σ for that case is shown by the bent "error bar" in Fig. 1. Since the change in σ is parallel to the general trend, such a

change in Neptune's distance could not be detected in
this way.

D. Summary

The degree of correspondence between the actual and
reconstructed solar nebulae is not certain. They will
agree in any regions which have not experienced signifi-
cant addition or removal of solid matter during the form-
ation of the planets. The observed planetary masses,
compositions, and orbital spacings are generally consist-
ent with such a model. For most of the planets, the
computed values of σ vary roughly as $a^{-3/2}$, which may
be tentatively identified with the actual structure of
the original nebula. The regions of Mercury, Mars, and
the asteroids fall well below this trend. The first can
be explained by the incomplete condensation of Fe and
silicates at the high temperatures in Mercury's zone,
and possibly the loss of some matter into the sun. The
zones of Mars and the asteroids each probably contained
several earth masses of solid matter, which was removed
before massive planets could form there. In the follow-
ing chapters, I shall examine some of the consequences of
such a removal, and demonstrate a possible mechanism for
its occurrence.

II. ACCRETION OF THE TERRESTRIAL PLANETS

A. Introduction

It is now generally accepted that the terrestrial planets were formed by the accretion of many small bodies (planetesimals). The many arguments in favor of such a hypothesis have been reviewed by Shmidt (1958) and Safronov (1972a). The mechanism(s) and time scale for this process are still disputed; published estimates for the formation time of the earth range from about 10^3 to 10^8 years. The different theories reflect different emphases and interpretations of incomplete and often contradictory evidence. Many of these estimates arise from attempts to explain a particular phenomenon; their relationships to a more complete cosmogony are often poorly defined. Turekian and Clark (1969) proposed rapid, inhomogeneous accretion of the earth in an effort to explain its structure and bulk chemical composition; the time scale was not specified, except that it is required to be short compared to the cooling time of an initially hot solar nebula. Some consequences of this theory were considered by Anderson and Hanks (1972), but there has been no attempt to make a quantitative theory of this type, or even to detail the mechanism of accretion. Among the numerous difficulties are the necessity for a high initial temperature at 1 AU, and for rapid transportation and accretion of condens-

ation products formed in a widely dispersed state.

Hanks and Anderson (1969) proposed an arbitrary accretion rate for the earth. The total accretion time was treated as a free parameter, in an attempt to produce a desired thermal history. The Hanks-Anderson accretion rate has been used for thermal history models of other bodies (Johnston et al., 1974; Siegfried and Solomon, 1974). However, there is no physical justification for the use of that particular model (Weidenschilling, 1974). Hills (1973) and Hallam and Marcus (1974) have developed elaborate mathematical models for accretion; however, they ignore or contradict explicit properties of orbits, and cannot realistically represent the conditions of planetary formation.

There are a few general cosmogonies in which accretion is considered in the context of a more inclusive theory. Alfven and Arrhenius (1970; see also Ip, 1974) proposed that electromagnetic forces affected the accreting particles. Cameron (1973b) suggested that turbulence and gas drag in a dense, massive solar nebula were important. He has recently revised his estimate of the nebular density (Cameron and Pollack, 1975); the effect on his accretional theory has not been announced. Safronov (1960, 1972a, 1972b) assumes that the accretion process was dominated by gravitational forces. Only gravitational forces will

be explicitly considered here; my results will be gener-
ally consistent with, but not necessarily limited to,
Safronov's cosmogony. The assumption that only gravit-
ational forces were important implies that gas drag was
negligible. The time scales which result force the con-
clusion that the nebula had dissipated before the terr-
estrial planets were formed. This is in apparent contra-
diction to the presence of the giant planets. They con-
tain large amounts of hydrogen and helium, and so must
have formed in the presence of the nebular gas (Cameron,
1973c; Perri and Cameron, 1974). This seeming inconsist-
ency may be explainable by the lower temperatures in the
outer part of the nebula, which allowed the condensation
of water ice in the zone of Jupiter (chapter I, above).
The orbital velocities of planetesimals, and possibly
their collision velocities, were smaller in the outer
nebula, and the icy matter may have stuck together more
easily than the rocky condensates of the inner nebula.
In any case, the presence of the giant planets must be
assumed in models of this type. We shall see that their
influence could have affected the formation of at least
some of the terrestrial planets.

One indication that the terrestrial planets accreted
after dissipation of the nebula is the abundance of noble
gases in the earth's atmosphere. Their relative abundances
differ from the solar ratios, but are similar to those of

non-radiogenic trapped gases in stony meteorites. Also,
their abundances relative to Si are similar for both
meteorites and the earth. (Wasson, 1969) If the earth had formed in
the presence of the nebula, it would have captured a
temporary solar-composition atmosphere. Escape of the
H and He, even by rapid hydrodynamic blowoff, would
leave the heavy noble gases in solar proportions (Hunten,
1973). Cameron (1973b) suggested rapid accretion, with
removal of the captured atmosphere by an intense T Tauri
solar wind, and production of a secondary atmosphere by
influx of volatile-rich meteoritic matter. However, it
has not been shown that the T Tauri solar wind could
remove an atmosphere on the required time scale of 10^6 yr.
Also, the noble gas/Si ratio of the earth implies that
the later influx must have consisted of matter extremely
enriched in noble gases (and, coincidentally, in the
amount to produce agreement with measured meteoritic
abundances), or else consisted of most of the planet's
mass. The remaining possibility is outgassing after the
solar wind removed the captured atmosphere, but the slow
outgassing required would probably be incompatible with
the high temperatures developed during rapid accretion.
I consider this to be a strong argument against inhomo-
geneous accretion or any other "rapid" accretion theory.
Knowledge of the noble gas contents of the atmosphere
of Venus and Mars could test this argument.

B. The Model

I propose the following simplified model as a
working hypothesis: A swarm of particles orbit the sun
in keplerian orbits with some range of eccentricities
and inclinations. The swarm has roughly the shape of a
disk, its total mass is equal to that of the terrestrial
planets. The conservation of mass follows from Öpik's
(1966a) demonstration that the terrestrial planets could
not eject significant amounts of matter from the swarm
(see chapter 3, below). The swarm contains a small
number of "embryos" or protoplanets, perhaps of about
lunar mass. The formation of such embryos has not been
explained in detail. Goldreich and Ward (1973) have
shown how kilometer-size bodies could form by gravita-
tional instability in a dust layer; these presumably
would be the members of the swarm. Cameron and Pollack
(1975) describe qualitatively the formation of larger
bodies from these. Lyttleton (1972) showed how bodies of
up to lunar size could form by capturing particles whose
heliocentric orbits lay between their inner and outer
Lagrange points. Unlike Lyttleton, however, I assume that
there were only a few embryos. Due to its gravitational
field, the largest body in some region of the swarm
grows more rapidly than the second largest body, and
"runs away" from it (Safronov, 1972a, ch. 9). Wetherill

-28-

(1975b, 1975c) has pointed out that a close encounter

within the Roche limit is several times more likely than

an actual collision, so the largest bodies will tend to

disrupt competing bodies in their vicinity. Large bodies

could not grow in closely spaced orbits. If they grew in

distant orbits which perturbations later caused to

intersect, they would collide at high relative velocities,

and disrupt rather than coagulate. This may have happened

many times during the growth of the planets; those embryos

which escaped destructive collisions, and accreted only

much smaller bodies, survived to become the planets.

Safronov (1966) showed that the axial tilts of the earth

and Mars could be produced if the largest impacting

bodies had masses about 10^{-3} times the planet's mass

(however, if Venus' retrograde rotation was caused by

such an impact, the mass ratio must have exceeded 10^{-2}).

In the later stages of growth, only a few massive

protoplanets survived in widely spaced orbits, and each

can be assumed to be the dominant influence on the swarm

in its vicinity.

The size distribution of the smaller bodies is not

important to this model, but deserves brief comment

here. Mass distributions have been derived by Safronov

(1966), Marcus (1967), Zvyagina and Safronov (1972), and

Hallam and Marcus (1974). These are all based on

solutions of the scalar transport equation, which assumes

only the coalescence of two colliding bodies, without
the possibility of disruption in collisions (Safronov
(1972a, ch. 8) considers some effects of fragmentation).
The various solutions which have been developed also
involve the implicit assumption that the presence of one
body in a volume of space does not influence the proba-
bility of another body being found there (Scott, 1968).
The stochastic nucleation theory of Hills (1973) also
makes these assumptions. The first is not valid for small
bodies, which have little gravitational binding energy
and are easily disrupted. The second, as we have seen,
is invalid for large bodies with appreciable gravitational
influence. The assumptions used in solving the scalar
transport equation lead to an overestimate of the
numbers and sizes of the largest bodies. Some of the
derived distributions may have been applicable at some
time, for some range of masses, but no important conclusion
can be based on them.

C. Gravitational Accretion

Let m be the mass of a planet, r its radius, and
its density. A particle of negligible size and mass
approaches the planet with a relative velocity u at
"infinity," when outside its gravitaional influence.
The impact parameter, b, is defined as the distance of
closest approach on the particle's unperturbed trajectory.
For a grazing impact, the closest approach on the

perturbed trajectory is r. Conservation of angular

momentum and energy gives

$$u^2 b^2 = r^2 (u^2 + 2Gm/r), \qquad (2.1)$$

where G is the gravitational constant, or

$$b^2 = r^2 (1 + 2Gm/ru^2)$$

$$= r^2 (1 + u_e^2/u^2) \qquad (2.2)$$

$$= r^2 (1 + 8\pi G\rho r^2/3u^2)$$

where ρ is the planet's density, and $u_e = \sqrt{2Gm/r}$ is the

escape velocity from its surface. The gravitational

capture cross-section, πb^2, is a function of u, and

formally becomes infinite when u approaches zero. This

singularity has no physical meaning in the realistic

case when the planet and particle are both in heliocentric

orbits. In that case, u can be zero only if their orbits

are identical. Any difference in orbital elements

produces some nonzero value of u. Consider the "worst

case" in which relative velocities are minimized, with

the planet and particle in coplanar circular orbits.

The circular keplerian velocity is

$$v_k = \sqrt{Gm_\odot/a}, \qquad (2.3)$$

where a is the orbital radius and $m_\odot$ the solar mass.

Differentiating,

$$\partial v_k / \partial a = -(Gm_\odot / 4a^3)^{1/2}, \qquad (2.4)$$

and the relative velocity between two orbits separated by a distance b has the magnitude

$$u = (Gm_\odot / 4a^3)^{1/2} b = kb. \qquad (2.5)$$

Inserting in Eq. (2.2), we find that

$$b^2_{max} = [r^2 + (r^4 + 4u_e^2 r^2 / k^2)^{1/2}]/2 \qquad (2.6)$$

gives the greatest value of b for which impact is possible. Let σ be the surface density of matter in the particle swarm. The accretion rate is formally

$$dm/dt = 2 \int_0^{b_{max}} \sigma \, u(b) \, db = \sigma k b^2_{max}$$

$$= \frac{\sigma k}{2} \left(\frac{3}{4\pi\rho}\right)^{2/3} m^{2/3} [1 + (1 + \frac{32\pi G\rho}{3k^2})^{1/2}] \qquad (2.7)$$

Eq. (2.7) is nonsingular for any finite value of k. A more realistic consideration of eccentric and inclined orbits would only increase the values of u; we might try a velocity distribution of the type $u = u_0 + kb$, or even a Maxwellian distribution with the mean velocity proportional to b. In either case, there is no singularity in dm/dt. The expression for the gravitational cross-

section contains two terms, proportional to r^2 and r^4. It has often been inferred from this that each term dominates in a certain size range, with the accretion rate rising sharply when the planet grows to the point that $u_e = u$ (Hartmann, 1968; Alfven and Arrhenius, 1970). However, Eq. (2.7) shows no such behavior; dm/dt is proportional to $m^{2/3}$, or r^2, for all values of m.

The value of k in Eq. (2.5) is $1.0 \times 10^{-7} a^{-3/2}$ sec^{-1}, where a is in AU. Since k is small, Eq. (2.6) is approximated by

$$b_{max} \simeq (u_e r/k)^{1/2} = (8\pi G\rho/3k^2)^{1/4} r \qquad (2.8)$$

For the earth, b_{max} is about 130 times r, or about twice the moon's distance (the figure 42r in Weidenschilling, 1974, is in error). However, this description does not correspond to physical reality, due to the earth's mass. In the three-body sun-earth-particle system, particles with very small values of u are found in Trojan or horse-shoe orbits. For values of b less than about 0.008 AU, such orbits are stable (Weissman and Wetherill, 1974), corresponding to a minimum value of u of about 0.1 km sec^{-1}, and a minimum b of about 190r. These values are for the earth's present mass, and would have been less for an embryo; this probably did not affect the earth's accretion.

A more realistic case of accretion, treatable by the two-body approximation, is that for which u is much larger than the minimum value. This is the case when the particle orbits have appreciable eccentricities and inclinations. There is then no significant correlation between u and b. The accretion rate is formally the capture cross-section, times u, times the space density of matter, δ, integrated over all velocities. In terms of the mass, the capture cross-section is

$$\pi b^2 = (\frac{3\sqrt{\pi}}{4\rho})^{2/3} m^{2/3} [1+(\frac{32\pi\rho}{3})^{1/3} G \frac{m^{2/3}}{u^2}]. \qquad (2.9)$$

The accretion rate is

$$\frac{dm}{dt} = (\frac{3\sqrt{\pi}}{4\rho})^{2/3} \delta \int_0^\infty m^{2/3}[1+(\frac{32\pi\rho}{3})^{1/3} G \frac{m^{2/3}}{u^2}] u\ g(u)\ du, \qquad (2.10)$$

where g(u) is a normalized function defining the distribution of u. Eq. (2.10) is formally correct for all cases of physical interest, but ρ, δ, and g(u) may vary with time, or with m. Any model of gravitational accretion is an implicit or explicit determination of these quantities.

The variation of ρ is small, and has little effect on the accretion rate. We can assume a constant average value, between the uncompressed and compressed planetary densities, with little error. The function g(u) is more

important, and less certainly known. Kaula and Bigeleisen
(1975) assume that it is Maxwellian, but there is little
real justification for this choice. Stellar velocities
are approximately Maxwellian (Chandrasekhar, 1960), but
there are important differences between the stars in the
galaxy and particles in the circumsolar swarm. Stellar
interactions are elastic, while the particles experience
inelastic collisions. Also, we shall see that the rate
of removal of particles from the swarm is a function of
velocity; this selective removal alters the velocity dis-
tribution. The Maxwellian distribution is strictly valid
only in a situation of thermal equilibrium; accretion is
in this sense a "thermal disequilibrium" process.

In spite of these reservations, we shall investigate
the effects of assuming a Maxwellian velocity distribu-
tion. In terms of the mean velocity, c, a three-dimen-
sional distribution is

$$g(u) \ du = 32/\pi^2 c^3 \exp(-4u^2/\pi c^2) \ u^2 \ du \qquad (2.11)$$

Inserting this expression into Eq. (2.10) and integrating,

$$\frac{dm}{dt} = (\frac{3\sqrt{\pi}}{4\rho})^{2/3} \ \delta \ c \ m^{2/3}[1+(\frac{32\pi\rho}{3})^{1/3} \frac{4G}{\pi c^2} \ m^{2/3}] . \qquad (2.12)$$

Note that Eq. (2.12) is nearly identical to that obtained
from Eq. (2.10) if all particles are assumed to have the

mean velocity c. We will assume, therefore, that g(u)

can be replaced by some effective mean velocity, whether

or not the actual distribution is Maxwellian.

The mean velocity always increases with time, prov-

ided that non-gravitational forces and inter-particle

collisions are unimportant. The principal reason for this

is the embryo's gravitational perturbations on the part-

icles. In successive close encounters (near misses),

the eccentricity of the embryo's orbit gives rise to a

statistical fluctuation in u, which results in a net

increase in the mean value (Öpik, 1966a). This change

can be described as the result of the non-existence of

the Jacobi integral in the elliptical three-body problem

(Szebehely, 1967). In the circular three-body problem,

u may be regarded as an invariant for any individual

particle. However, in an ensemble of particles with a

range of velocities, the mean velocity will increase with

time, even in this ideal case, because the accretion rate

is a function of u. If a particle has a given value of

u, but the relative velocity vector at encounter is all-

owed any orientation, then all possible particle orbits

are contained within some volume of space surrounding the

planet's orbit. The size and shape of this volume will

be discussed below. The distance from the planet's orbit

to the boundary of this volume in any direction is prop-

ortional to u, so the volume is proportional to u^2. The rate at which the planet sweeps out this volume is $\pi b^2 u$; the characteristic time for the sweeping out is the rate divided by the volume, or proportional to b^2/u. If N is the number of particles with velocity u, then

$$-\frac{1}{N} \frac{dN}{dt} \propto (\frac{1}{u} + \frac{u_e^2}{u^3}).$$

Particles with smaller values of u are depleted more rapidly, so the mean velocity of the remaining particles increases with time, even in this ideal case. In reality, the increase due to perturbations is more important.

The mutual gravitational scattering of the particles themselves also tends to increase their relative velocities. For a system of particles of equal mass, Safronov (1972a, ch. 7) derives the relation

$$u^2 = Gm/\theta r = u_e^2/2 \theta, \qquad (2.13)$$

where θ is a dimensionless parameter on the order of a few units, depending weakly on particle size. For a power law size distribution, m and r refer to the largest body, and Eq. (2.13) is obeyed by those particles which interact with it. However, elastic or inelastic encounters with other bodies which do not approach the largest body may change u. Safronov also applies Eq. (2.13) to the case where one body (planet) is much more massive than

any other. The physical basis for this step is unclear. Öpik's (1966a) theory of encounters relates the increase of u to the orbital parameters of the particle and planet. The rate of increase depends on the eccentricity and other parameters of the planet's orbit; it appears impossible to express this process in terms of a single parameter, or to relate it directly to the planet's mass. While more realistic than Safronov's model, Öpik's depends on too many unknown quantities to be usable in this case. We shall use a variation of Safronov's theory, realizing that it is only qualitatively correct, and examining the effect of different values of θ.

If we accept Eq. (2.13), we find that $b^2 = r^2(1+2\theta)$, that is, the gravitational cross-section is always $(1+2\theta)$ times the geometric cross-section. To integrate Eq. (2.12), we still need an expression for δ. In the case where the "feeding zone" (the volume surrounding the planet's orbit in which the particles move) is of constant size, we can write

$$\delta = \delta_0(1-f),\qquad\qquad(2.14)$$

where $f = m/m_p$ is the fraction of the total mass which has been accreted, with m_p the final mass of the planet. We shall see below that Eqs. (2.13) and (2.14) are not strictly consistent; however, they allow the accretion

-38-

rate to be determined analytically. With Eq. (2.12), they give

$$\frac{dm}{dt} = \delta_o \left(\frac{27\pi^3}{8\rho^4}\right)^{1/6} \sqrt{G/\theta} \ (1+4\theta/\pi) \ m \ (1-m/m_p). \qquad (2.15)$$

This can be integrated, giving

$$m(t) = \frac{m_p \ m_o \ \exp(t/\tau)}{(m_p-m_o)+m_o \ \exp(t/\tau)}, \qquad (2.16)$$

where m_o is the mass at t=0, and the characteristic time τ is

$$\tau = \frac{\sqrt{\theta/G}}{(1+4\theta/\pi)} \left(\frac{8\rho^4}{27\pi^3}\right)^{1/6} \delta_o^{-1}. \qquad (2.17)$$

For $m_o \ll m_p$, this simplifies to

$$m(t) \simeq \frac{m_o \ \exp(t/\tau)}{1+(m_o/m_p) \ \exp(t/\tau)}. \qquad (2.18)$$

A reasonable value for δ_o in the earth's zone is 10^{-11} g-cm^{-3} (see below). For this value, and $\rho=4$ g-cm^{-3}, $\theta=4$, τ is about 3×10^7 yr. Equations (2.15) and (2.18) are plotted in Figure 2. With $m_o=0.01 \ m_p$, accretion is 99% complete in an interval of 10τ, or 3×10^8 yr. Also shown in Fig. 2 are m(t) and dm/dt for the accretion rate of Hanks and Anderson (1969). That rate, which is

-39-

entirely arbitrary, is seen to add most of the mass near
the end of the accretion interval.

I have stated that Safronov's velocity relation,
Eq. (2.13), is not consistent with Eq. (2.14), which
implies a constant zone volume. Note that by (2.13),
$u = 0$ when $m = 0$; i.e., the relative velocities are
initially zero, implying that all particles lie on the
same orbit. Then δ_o would be infinite, or else (as
would actually be the case) the particles would not be
in the zone initially, but must be added later. In that
case, the factor $(1-f)$ cannot be used. We wish to
generalize Safronov's relation to allow for disturbances
in the swarm other than the embryo's own gravity.
Initially chaotic orbits may have been due to turbulence in
the solar nebula, catastrophic collisions of other
embryos, or the perturbations of other planets which
formed earlier. The formation of the planetesimals by
gravitational instability (Safronov, 1972a, ch.6;
Goldreich and Ward, 1973) could not have occurred if
velocities in the solar nebula were appreciable.
However, such velocities could have developed later,
especially if Jupiter formed before the terrestrial
planets. I propose a velocity relation of the form

$$u^2 = u_o^2 + u_e^2/2\theta, \qquad (2.19)$$

where u_o is some initial velocity. The ratio of gravi-
tational and geometrical cross-sections is $(1+u_e^2/u^2)$.
If u remained constant, this ratio would increase as the
planet grew, becoming proportional to r^2 as the second
term became dominant. For Safronov's relation, Eq. (2.13),
this ratio is constant, equal to $(1+2\theta)$. With Eq. (2.19)
the ratio approaches unity when m is small, and $(1+2\theta)$
when m is large.

It is convenient to replace u with a dimensionless
relative velocity. Let v_k be the Keplerian circular
velocity of the planet in its orbit. I define

$$U = u/v_k. \qquad (2.20)$$

Since $v_k^2 = Gm_{\oplus}/a$, Eq. (2.13) becomes

$$U^2 = m^{2/3}/C\,\theta, \qquad (2.21)$$

where $C = (3/4\pi\rho)^{1/3} m_{\oplus}/a$, and (2.19) becomes

$$U^2 = U_o^2 + m^{2/3}/C\theta. \qquad (2.22)$$

This dimensionless notation is particularly convenient,
since the range of orbital parameters of the particles
are simple functions of U (Öpik, 1951). The maximum
possible eccentricity is

$$e_{max} = U^2 + 2U, \qquad (2.23)$$

and occurs when the particle's orbit has its perihelion
at the planet's orbit, and the orbits are coplanar. The
particle's aphelion is then the largest possible value,

$$Q_{max} = (1+U)^2/(1-2U-U^2),$$ (2.24)

in units of a. If the particle's aphelion is at the
planet's orbit, then the minimum perihelion is

$$q_{min} = (1-U)^2/(1+2U-U^2),$$ (2.25)

and eccentricity is

$$e(q_{min}) = 2U-U^2.$$ (2.26)

For any U, the maximum inclination (relative to the
planet's orbital plane) is

$$i_{max} = 2 \sin^{-1}(U/2).$$ (2.27)

For small values of U, it is a sufficiently good approx-
imation to take

$$Q_{max} = 1+4U \qquad e_{max} = 2U$$

$$q_{min} = 1-4U \qquad i_{max} = U.$$ (2.28)

Eqs. (2.28) define the limits of the "feeding zone"
from which the protoplanet may accrete matter. The zone

-42-

is a flattened torus enclosing the protoplanet's orbit,
with the out of plane thickness one fourth the width in
the orbit plane. Its volume is approximately $8 \pi^2 a^3 U^2$.
If the total mass in the zone is specified, then δ is
determined by the value of U. Ip (1974) underestimates
the volume of the torus by a factor of four, thereby
overestimating the accretion rates in his model. Hartmann
and Davis (1975) do "particle in a box" calculations of
planetesimal collisional lifetimes, in which u is varied
for a "box" of constant size. The volume chosen corres-
ponds to $u = 2.3$ km sec^{-1} at the earth's orbit; they
therefore overestimate the the lifetimes for smaller u,
and underestimate them for larger values.

The simplest assumption is that of a "closed" feeding
zone: all of the mass to be accreted is present in the
zone initially, and equally available to the planet.
The volume of the zone is proportional to U^2, so

$$\delta = \delta_o \, (1-f)(U_o^2/U^2), \tag{2.29}$$

where

$$\delta_o = m_p / 8 \pi^2 a^3 U_o^2. \tag{2.30}$$

Safronov uses the surface density, rather than the
volume density of matter. His results depend on an
assumed vertical structure of the swarm. His relation is
equivalent to

$$\delta = \delta_o(1-f)(U_o/U), \qquad\qquad (2.31)$$

which implies that the entire swarm thickens uniformly.
However, in the case where a single large protoplanet
dominates a zone, Eq. (2.29) is more realistic. The
radial growth of the zone may result in matter being
added to it from adjacent regions of the swarm. The
amount will depend on the local density of the swarm,
and the mechanism, gravitational perturbations or
inelastic collisions, or both. However, in such cases
the initial mass in the zone is less than m_p, and δ_o
is less than given by Eq. (2.30), The initial stages
of accretion will be slower, and for most such models,
it appears that the total accretion time would be
lengthened. Matter added at the edge of a zone must
have a certain minimum velocity (relative to the circular
velocity) in order to reach the center of the zone and
be accreted.

Inelastic collisions at the edges of a zone would
tend to reduce relative velocities there, effectively
removing matter from the center of a zone. In the case
where two zones overlap, collisions between particles
from the different zones would lead to a buildup of
matter in the area of overlap, which would be gradually

diffused back into planet-crossing orbits by the planet's perturbations. In the later stages of accretion, when δ is small due to the large volume of the zones and the small amount of matter remaining, adjacent zones can overlap with little interaction until the edge of one zone reaches the center of the other, allowing close approaches to that zone's planet. If collisions are very effective within a zone, as well as at the edges, the eccentricities decrease, and the radial excursions of the particles are damped. A planet will then bore a "tunnel" through the swarm, and particles must diffuse into the tunnel, then increase their eccentricities to become planet-crossing. The initial accretion rate might be large, but the later stages are slowed, and the time to completion is increased. From these considerations, it appears that the "closed" feeding zone assumption represents a lower limit for the accretion time, in the absence of nongravitational forces. Note that the factor (1-f) cannot be applied to an "open" zone, to which matter is added during accretion. For the earth, with $\theta = 5$, and $f = 0.01$, Eq. (2.21) gives $U = 0.025$, with a zone width of 0.2 AU. Much of the final mass lay outside this zone initially, if there were no gaps devoid of matter between zones.

Obviously, the embryos did not know where to form,

and the value of U_o was not selected, in order to make
the "best fit" of zones. Assuming the zones to be cent-
ered on the present planetary orbits, a choice of $U_o=0.05$,
$\theta=5$, provides a reasonable fit, with adjacent zones
approximately touching originally, and finally over-
lapping about to their centers. This is arbitrarily sel-
ected as the "nominal" case for detailed consideration.
The initial and final zone limits for this case are shown
in Table III. The initial zone width is primarily set by
the value of U_o, while the final width depends mainly on
θ (for the earth, with $\theta=5$, the final value U_f is 0.12 if
$U_o=0$, and 0.13 if $U_o=0.05$). Values of U_o less than 0.05
would allow significant gaps between zones, and diminish
δ_o, unless the zones were isolated "jetstreams." Larger
values of U_o, or smaller values of θ, allow significant
overlapping of zones beyond their centers, with exchange
of matter between zones. The differences in bulk chem-
ical compositions of the terrestrial planets (Lewis, 1972)
would be blurred if this had occurred on a large scale.
Some exchange between zones must have taken place in the
late stages of accretion, but amounted to a very small
fraction of the planetary masses.

TABLE III

Feeding Zone Boundaries, Nominal Case

$$U_o = 0.05 \qquad \theta = 5$$

	Initial		Final		
	q_{min}	Q_{max}	U_f	q_{min}	Q_{max}
Mercury (a=0.39)	0.32	0.48	0.057	0.31	0.49
Venus (a=0.72)	0.59	0.89	0.11	0.48	1.13
Earth (a=1.0)	0.82	1.23	0.13	0.62	1.74
Mars (a=1.52)	1.25	1.87	0.08	1.11	1.96

D. Results

It is convenient to express the accretion rate as df/dt, rather than dm/dt. We have

$$\delta(f) = m_p(1-f)/8\pi^2 a^3 U^2(f), \qquad (2.32)$$

$$U^2(f) = U_o^2 + (4\pi\rho/3)^{1/3}(am_p^{2/3}/m_\odot\theta)f^{2/3} \qquad (2.33)$$

and

$$u_e^2/u^2 = (32\pi\rho/3)^{1/3} am_p^{2/3}f^{2/3}/m_\odot U^2(f). \qquad (2.34)$$

The accretion rate is

$$\frac{df}{dt} = (\frac{\pi v_k}{m_p})(\frac{3m_p}{4\pi\rho})^{2/3}\,\delta(f)U(f)\,f^{2/3}[1+\frac{v_e^2(f)}{v_k^2 U^2(f)}], \qquad (2.35)$$

which cannot be integrated analytically. Eq. (2.35) was integrated numerically to obtain $f(t)$, using a fourth-order Runge-Kutta method. For most cases, an initial value of $f=0.01$ was assumed. Integration was terminated at $f=0.99$. The accretion time, t_a, is arbitrarily defined as the time for f to increase from 0.01 to 0.99. Further integration would yield no useful information; it should be emphasized that this formal procedure does not account for bodies in unusually stable orbits, or scattered from other zones, which would dominate the final stage of accretion.

The results for the nominal case for the earth are shown in Fig. 3. For the earth's mass, $U_o=0.05$ in Eq. (2.30) gives $\delta_o=10^{-11}$g cm^{-3}. The accretion time is

1.56×10^{8}yr. The maximum growth rate is less than
10 cm yr^{-1}. The effect of varying U_o is shown in Fig. 4.

Since δ_o varies as U_o^{-2}, the accretion time increases

with U_o. If δ_o is artificially held constant, t_a decreases

as U_o is increased, but this is equivalent to increasing

the mass in the zone. For a closed feeding zone, the

accretion time must increase with U_o. Note that this

model does not reduce to Safronov's in the limit of

$U_o \to 0$. In reality, the closed zone assumption must break

down, and δ_o reach some limit as the zone shrinks.

However, even for $U_o = 0.02$ ($U = 0.03$ at $f = 0.01$), and the

entire mass assumed originally present in a zone 0.25 AU

wide, t_a is slightly more than 10^{8}yr. Safronov's model,

with $\theta = 5$, gives t_a about 0.7×10^{8}yr (Safronov, 1972a,

ch.9.). Eq. (2.18), using somewhat different numerical

values, gave 3×10^{8}yr. The accretion time appears to be

quite insensitive to the assumed value of U_o.

The parameter θ does not affect δ_o or U_o. Fig. 5

shows the effect of variations in θ for fixed values of

the other parameters. Smaller values of θ correspond

to larger values of U_f; for $\theta = \infty$ there is no acceleration.

Even a small acceleration drastically lowers the peak

accretion rate; t_a is affected less strongly. The late

stages of accretion are slowed, causing the maximum rate

to occur at smaller values of f. Fig. 6 shows df/dt for

the earth as a function of f, normalized to the maximum.
For $\theta = 3$, the peak is at $f = 0.33$; even for $\theta = \infty$, it occurs
only at $f = 0.57$. There is no late accretion peak at large
values of f, as suggested by Hanks and Anderson (1969),
or Hallam and Marcus (1974). There is no late peak for
Safronov's model, either.

F. Thermal Structure of an Accreting Planet

The potential and kinetic energy of a planetesimal
is released upon impact with a planet. An energy balance
at the surface gives (Benfield, 1950; Mizutani et al.,
1972)

$$\rho_s(u^2/2+Gm/r) \; dr/dt = \varepsilon\sigma(T^4-T_a^4)+\rho_s C[(T-T_b)+\lambda] \; dr/dt$$

$$+\kappa(\partial t/\partial r), \qquad (2.36)$$

where ρ_s is the planet's density at the surface, ε is the
surface emissivity, σ the Stefan-Boltzmann constant, T_a
the effective temperature of "space" seen from the surface
(including contributions by solar radiation and any atmo-
sphere), T_b the temperature of the infalling particles,
C the heat capacity, κ the thermal conductivity, and λ
the latent heat of any phase changes. It can be shown
that the first term on the right side is much larger than
the other terms (Benfield, 1950), and Eq. (2.36) can be

approximated by

$$\rho_s(u^2/2 + Gm/r)\ dr/dt \simeq \varepsilon\sigma(T^4 - T_a^4).$$ (2.37)

This equation has been used in the accretional thermal models of Hanks and Anderson (1969) and Mizutani et al. (1973). In terms of the mass, it may be written

$$[u^2/2 + (4\pi\rho/3)^{1/3}Gm^{2/3}](36\pi)^{-1/3}\ \frac{\rho^{2/3}}{\rho_s}\ \frac{dm/dt}{m^{2/3}}$$

$$\simeq \varepsilon\sigma(T^4 - T_a^4).$$ (2.38)

If the kinetic energy, $u^2/2$, is small compared with the gravitational potential energy, the power input per unit area is proportional to dm/dt. For Safronov's velocity relation, Eq. (2.13), this is also true for all values of u. In such cases, the temperature peak in accretion coincides with the peak in the accretion rate. For the velocity relation of Eq. (2.22), the kinetic energy of the particles contributes most of the accretional energy in the early stages of accretion. The temperature peak occurs at a slightly smaller mass than the peak in df/dt, but for reasonable values of U_o the displacement is very small. For all realistic combinations of U_o and θ, the temperature peak from Eq. (2.38) occurs at values of f less than 0.5. The magnitude of the temperature peak is

typically on the order of a few tens of degrees, and
never more than 100°K, for the earth. The low tempera-
ture peak and the small radial growth rates indicate that
accretional heating is negligible, under these assump-
tions.

This conclusion should be viewed with suspicion,
since some of the assumptions may be invalid. The energy
balance of Eq. (2.36) assumes that all energy of impact
is released at the planet's surface. Levin (1972a) has
suggested that seismic waves could have heated the deep
interior of the earth during accretion. The fraction of
impact energy released in this form is uncertain, but is
estimated by Schultz and Gault (1975) to be on the order
of 10^{-4}. If this figure is correct, seismic heating could
have amounted to only a few degrees. However, this does
not mean that most of the impact energy was in fact
released at the planet's surface. Eq. (2.36) implies that
dr/dt was uniform, when accretion was actually a series of
impacts which produced intense local heating and consid-
erable scattering of ejecta. If the impacting bodies were
sufficiently small, most of the energy was released at the
surface. Safronov (1972a, ch. 14) estimates that bodies
less than 100 meters in diameter would meet this condition.
Larger bodies would excavate craters in which some fraction
of the impact energy would be trapped by the fall-back of

ejecta. Heat transfer in that case is not primarily by
conduction or radiation, but by the mixing and overturn
of the surface layers by later impacts. The problem of
heating by large impacts is complex; our knowledge of the
phenomena of cratering is insufficient to analyze it in
detail. Safronov (1972a, ch. 15) obtained an approximate
solution, estimating that the earth could have been
warmed by about $1100^{\circ}K$ at a depth of 400-500 km below the
final surface. This figure should be considered only as
an order of magnitude estimate, due to the uncertainties
involved.

We may, however, consider the qualitative aspects of
Safronov's solution. The amount of heating depends prim-
arily on the sizes of the impacting bodies. Safronov
assumed a power-law size distribution in which most of
the mass was contained in the largest bodies. I have
argued that such distributions, which are based on solut-
ions of the scalar transport equation, tend to overest-
imate the numbers and sizes of the larger bodies. How-
ever, the size distribution is not important, only the
fraction of mass contained in bodies above a critical
size. Safronov's heating estimate is probably not
seriously in error from this assumption. The presence
or absence of an atmosphere during accretion has little
effect, since kilometer-sized bodies can penetrate the

-53-

earth's atmosphere without appreciable loss of energy.
Small-body impact heating is a function of the accretion
rate. Large-body heating depends primarily on the impact
energy, and so is greatest in the final stages of accretion,
with the highest temperatures developed in the outer
part of the planet. The small-body assumption, when
coupled with thermal histories which require significant
initial heating, leads to very short estimated accretion
times (Hanks and Anderson, 1969; Mizutani et al., 1972),
which cannot be achieved by the models developed here.
Accretion of larger bodies apparently can produce an
acceptable lunar thermal history with more reasonable
accretion times, on the order of 10^8yr (Wetherill, 1975d).

F. Other Planets: The Problem of Mars

The accretion time of the earth is on the order of
10^8 years, and is not very sensitive to our choices of
U_o and θ. Whatever the values of these parameters for
the other planets, their accretion times will depend on
them in the same manner. The assumption that U_o and θ
are the same for different zones does not mean they are
geometrically similar, since different masses will mean
variations in the effect of θ. The accretion times for
all the terrestrial planets were computed with initial
embryos of $f=0.01$. If we chose embryos of identical
mass, the planets other than the earth would begin at
larger values of f, and have shorter accretion times.

For $U_o=0.05$, $\theta=5$, the computed times are: Mercury, 0.69 X 10^8 yr; Venus, 0.55 X 10^8 yr; earth, 1.56 X 10^8 yr; Mars, 2.6 X 10^9 yr.

These accretion times are in general agreement with earlier results (Safronov, 1972a; Weidenschilling, 1974). The extremely long accretion time for Mars is due to its small mass and large feeding zone (the smaller mass of Mercury is offset by the small size of its zone). No reasonable values of U_o and θ can reduce the accretion time of Mars to less than 10^9 yr. This figure is not due simply to the assumption of a nonzero initial velocity; Safronov also computes accretion times in the range of 1.5-2.4 b.y. The small masses of Mars and Mercury reduce the sensitivity to the value of θ. For Mars, U_o might be even larger than the nominal case, due to Jupiter's perturbations. Even if begun with a larger embryo, f=0.1, the nominal accretion time is 2.0 b.y. If Mars originated at the same time as the other terrestrial planets, then some other mechanism, such as nongravitational forces, was required to form it, and therefore may have operated in the accretion of the other planets, as well. Before abandoning gravitational accretion, we should consider the possibility of a "young" Mars.

The age of the Martian surface cannot be determined solely by crater counts. Soderblom et al. (1974) derived

-55-

a Martian cratering history in which absolute ages of
different terrains were derived by comparison with lunar
terrains. Absolute ages of the Apollo sites are known;
crater counts in these regions can, in principle, give
the impact flux history. Soderblom et al. assumed that
the lunar and Martian fluxes were similar in time. They
concluded that Mars suffered an early intense bombardment
which declined rapidly at the end of accretion, some 4
b.y. ago. Chapman (1974) pointed out that other histories
of cratering and obliteration could produce the observed
crater distributions. Even the lunar flux histories have
been criticized by Hartmann (1975). The present-day
cratering flux at Mars is subject to considerable uncert-
ainty (Wetherill, 1974a); that due to an earlier populat-
ion of stray bodies is even more speculative. I have
proposed one mechanism by which the Martian cratering
flux may have significantly exceeded the lunar flux in
the past: asteroids stored in quasi-stable Trojan orbits
could be perturbed by Jupiter into Mars-crossing orbits;
most would be ejected from the solar system before becom-
ing earth-crossing (Weidenschilling, 1975b; ch. 4, below).
We know little about Martian endogenous erosive processes,
whether steady or episodic. Even a "young" Mars would
have a surface age of some 2 b.y.; this interval is ade-
quate to produce the observed terrains.

A more serious test of a "young" Mars is the planetary thermal hiatory. The observed moment of inertia indicates that Mars has a dense core. The division of the planet into an isostatically high, heavily cratered "continental" hemisphere and a lower, relatively smooth "ocean" basin suggests large-scale differentiation and crustal formation (Siever, 1974). The huge shield volcanoes of the Tharsis region require a source of magma, with at least some melting in the mantle. The most detailed thermal models of Mars are those of Johnston et al. (1974). They assumed that Mars formed 4.6 b.y. ago, and considered a variety of initial temperatures and uranium concentrations. For all the models tested, formation of an Fe-FeS core began within 10^9 yr. Melting of a dry silicate mantle began in from 2 to 2.75 b.y. However, Lewis (1972) has argued that Mars should have formed at least as rich in hydrous minerals as the earth. The addition of water would lower the silicate melting temperatures considerably; all of the models of Johnston et al. would reach the wet solidus in the mantle in less than 1 b.y. These models do not correspond exactly to the thermal behavior of a slowly accreting Mars, but should be generally similar. For a nominal accretion time of 2.6 b.y., Mars attains 95% of its final mass in only 1.6 b.y. There have been no thermal history calculations in

-57-

which the time scale for accretion is comparable to that for thermal evolution; they should be performed for Mars. It appears that planetary differentiation could have kept pace with accretion; indeed, the conclusion of Siever (1974) that the Martian crust formed before the end of accretion is only compatible with this thermal history if accretion lasted more than 10^9 yr.

The most serious test of the age of Mars is the completeness of accretion. By our formal procedure, the half life of planetesimals in the final stage of accretion is 2.4×10^8 yr in the nominal case. If collision with Mars were their only fate, about 10^{-6} earth masses would remain as Mars-crossing asteroids. However, the zone of Mars cannot be called "closed" on this time scale. A significant fraction of Mars-crossing bodies would become earth-crossing; the time scale for this is rather uncertain, but is on the order of a few times 10^8 yr (Wetherill, 1975c). The half-life of the Mars planetesimals would be significantly reduced by the loss of earth-crossers. However, this explanation poses a new difficulty. Since Mars would have attained most of its mass some 10^9 yr before the formal end of accretion, the amount of matter scattered into earth-crossing orbits in this interval would be comparable to that accreted by Mars, perhaps as much as a tenth of the planet's mass.

Most of this matter would be captured by the earth and Venus, but about 1% would strike the moon (Wetherill, 1975b, 1975c). The moon must have been closer to the earth at that time, but this probably would not change Wetherill's results significantly. The amount appears to be too large to be part of the post-mare cratering flux. It does match that suggested by Wetherill for a pre-mare "cataclysm" about 4 b.y. ago. The time scale is a less severe constraint if the bombardment is assoc- iated with mare formation, but special conditions must be assumed to explain the asymmetric distribution of maria, and the apparent low velocity of the Imbrium impact.

Whatever their fate, the Mars planetesimals seem to have vanished completely. Öpik (1966a) lists 34 asteroids as Mars-crossing, and identifies 11 of them as original members on the basis of their velocities relative to Mars. However, Wetherill (1974b) showed that the orbits of most of these bodies never intersect that of Mars. This scarcity of true Mars-crossers is signif- icant. Possibly, their computed lifetimes are too long. Most of the lifetimes computed by Wetherill, on the order of 10^9 yr, are for objects derived from the asteroid belt, with perihelia near Mars. Their orbits are probably quite different from those of typical planet- esimals originating near Mars. Nongravitation forces may have been effective in the past, e.g., gas drag

before dissipation of the solar nebula (Cameron, 1974), or later, on a much longer time scale, such as the Yarkovsky effect (Öpik, 1951; Peterson, 1975).

Our present knowledge cannot rule out a Martian surface significantly younger than the other terrestrial planets, but the lunar cratering record imposes severe constraints. It should be emphasized that if Mars is the same age as the other terrestrial planets, that result is by no means trivial. The age of Mars is an important clue to the processes which formed it and the other terrestrial planets. The Cameron or Alfven-Arrhenius cosmogonies could conceivably form Mars in a short time by nongravitational forces. There is yet another possibility, which involves only gravitational forces, but violates the feeding zone assumption. The next chapter will develop the theory necessary to explain this model.

III. CLOSE ENCOUNTERS OF SMALL BODIES AND PLANETS

A. Introduction

The solar system contains many small bodies in heliocentric orbits which cross the orbit of one or more planets. They include comets, meteoroids, and some asteroids; to these we may add the possibly extinct populations of planetesimals and protocometary bodies which were numerous during the formation of the solar system. The distinctions between these types of bodies are tenuous, and in this chapter I shall simply refer to them as particles, regardless of origin. Their orbits are generall unstable, since perturbations will alter the positions of the nodes and apsides, eventually leading to actual intersection with the orbit of a planet. Unless prvented by a resonance, close encounters with a planet will occur, drastically changing the particle's orbit. Its ultimate fate will generally be either collision with a planet or ejection from the solar system. The probabilities of encounter, collision, and ejection are of prime importance for establishing the fates, and inferring the origins, of these bodies. In an important series of papers, Öpik (1951, 1963, 1966a, 1966b, 1973) has developed expressions for these probabilities in terms of the particle's orbital elements. Arnold (1965) has applied these formulas in Monte Carlo simulations to invest-

igate possible sources of meteorites. Bandermann and
Wolstencroft (1970, 1971) derived an analytical express-
ion for the probability of ejection which differed from
Öpik's, but did not comment on its physical significance.
In this chapter, I present an alternative approach to
this problem, developing expressions by which the prob-
abilities of collision and ejection may be evaluated
numerically. This method avoids certain approximations
used by Öpik, and is conceptually simple.

I consider an "encounter" to be a passage within a
planet's sphere of influence, with a more precise defini-
tion to be developed later. Since each encounter changes
the particle's orbit, it is convenient to avoid formula-
tions in terms of its orbital elements wherever possible.
However, the probability that the particle's orbit inter-
sects that of the planet for a random value of the argu-
ment of perihelion depends explicitly on the values of
inclination and eccentricity (Öpik, 1951). If we assume
that the orbits do intersect, the probabilities of encoun-
ter per revolution of the particle, and of collision and
ejection per encounter, can be formulated in terms of
the relative velocity and a single angular variable.
The relative probabilities per encounter are of consid-
erable value, even if the probabilities per unit time
are poorly known.

By considering only encounters within the sphere of

influence, we overlook the effects of small perturbat-
ions by distant encounters. However, we shall see that
the angular deflections produced by distant encounters
are small, and can only cause ejection of particles in
nearly parabolic original orbits. Such cases are covered
in the extensive literature on long-period comets, which
will not be reviewed here.

B. Computations

Let the planet's mass be m_p, and the solar mass $m_\odot$.
The particle's mass is considered negligible. The
planet's orbit is considered circular, of radius a_p.
The planet's sphere of influence has radius d; the planet
has radius r. The Keplerian orbital velocity of the
planet is v_k. The particle's orbit has semimajor axis a,
inclination i with respect to the planet's orbital plane,
and eccentricity e. The particle's heliocentric velocity
has magnitude v; its velocity relative to the planet has
magnitude u. I define the following dimensionless quant-
ities:

$$U = u/v_k \qquad V = v/v_k \qquad A = a/a_p$$

$$M_p = m_p/m_\odot \qquad R = r/a_p \qquad D = d/a_p .$$

In general, upper case quantities will be considered
dimensionless. This notation is equivalent to that in

which $m_\odot$, a_p, and v_k are the fundamental units of mass,
length, and velocity, with the unit of time $\pi/2$ times
the planet's orbital period. However, such a procedure
leads to taking trigonometric functions of formally
dimensioned quantities. Also, it will be convenient at
times to use conventionally dimensioned units.

I consider the particle to be unaffected by the
planet's gravity when outside the sphere of influence,
and unaffected by the sun's when within it (two-body
approximation). With these idealizations, an encounter
changes the orientation of the particle's relative vel-
ocity vector $\underline{U}$, but its magnitude U remains unchanged.
Everhart (1973, 1974) has criticized this assumption, and
pointed out that the eccentricity of the planet's orbit,
here neglected, will result in changes of U. This effect
appeared in Arnold's (1965) simulations, and was explained
by Öpik (1966a), who showed that the variation in U could
be treated statistically. The change in a single encoun-
ter, ΔU, is small and random; the effect of many encoun-
ters is a gradual increase in U. However, ΔU is small
enough to be neglected in the derivation of the encounter
geometry. In those cases where a particle's orbit crosses
the orbits of two or more planets, U can be changed greatly
by successive close encounters (Öpik, 1966a). U may also
be changed by more distant perturbations (Everhart, 1973a).
However, we shall see that the probability of ejection is

not sensitive to the value of U, and the basic conclus-
ions of this chapter will not depend on the assumption
of a constant U. Note, however, that even in the more
realistic case of elliptical planetary orbits, the rel-
ative variation of U per encounter is much smaller than
are the changes in A, e, and i. It is certainly desir-
able to avoid formulations in terms of these orbital
elements, if possible.

The condition for ejection is easily found. When-
ever $V^2 > 2$, the particle's heliocentric velocity exceeds
the escape velocity. I define a planetocentric coord-
inate system with the x-axis directed away from the sun,
the y-axis in the direction of the planet's orbital mot-
ion, and the z-axis perpendicular to the orbital plane.
The components of U in terms of the particle's orbital
elements are (Öpik, 1963)

$$U_x^2 = 2-A(1-e^2)-1/A \qquad (3.1)$$

$$U_y = [A(1-e^2)]^{1/2}\cos i -1 \qquad (3.2)$$

$$U_z^2 = A(1-e^2)\sin^2 i \qquad (3.3)$$

The magnitude of U is given by

$$U^2 = 3-2[A(1-e^2)]^{1/2}\cos i -1/A \qquad (3.4)$$

Equations (3.1-3.4) are approximations derived from
Tisserand's criterion, but are entirely adequate for our

purposes. In these units, Tisserand's criterion is equivalent to $U^2 = 3-C$, where C is the Jacobi quantity.

Let ϕ be the angle between the negative y-axis and the vector $\underline{U}$. The heliocentric velocity is

$$V^2 = 1+U^2-2U \cos \phi, \qquad\qquad (3.5)$$

where

$$\cos \phi = (U^2-1+1/A)/2U. \qquad\qquad (3.6)$$

$V^2 = 2$ defines the condition for escape on a parabolic orbit, corresponding to $1/A = 0$, and a critical angle

$$\phi_c = \cos^{-1} [(U^2-1)/2U], \qquad\qquad (3.7)$$

where U must be between $\sqrt{2}-1$ and $\sqrt{2}+1$ for ϕ_c to be defined. For $U<\sqrt{2}-1$, ejection is not possible for any orientation of $\underline{U}$. For $U>\sqrt{2}+1$, all possible initial heliocentric orbits are hyperbolic, and will not be considered. Note that all particle orbits with $U<1$ are prograde, and all orbits with $U>\sqrt{3}$ are retrograde. Orbits between these values may be either prograde or retrograde.

The condition $\phi=\phi_c$ defines a cone symmetric about the y-axis, which we shall call the "escape cone." Any U vector which lies within the escape cone corresponds to a hyperbolic heliocentric orbit. Consider a particle, with some initial $\underline{U}$ with $\phi<\phi_c$, which encounters a planet and has its U vector turned through some angle γ (ϕ and γ

assumed known). After the encounter, the new vector $\underline{U}'$
lies on a cone with half-angle γ, symmetric about $\underline{U}$. I
assume all orientations of $\underline{U}'$ on this cone are equally
likely, therefore the width of the intersection of the
U' cone and the escape cone is a measure of the ejection
probability. The geometry of the encounter is shown in
Figure 7. From the law of cosines for spherical tri-
angles, the width of intersection of the cones is 2β,
where

$$\beta = \cos^{-1} [(\cos \gamma \cos \phi - \cos \phi_c)/\sin \gamma \sin \phi].$$

(3.8)

The probability of ejection, $P(\infty)$, is β/π. Note the
special cases: (1) If $\phi=0$, $P(\infty)=0$ for $\phi<\phi_c$; $P(\infty)=1$ for
$\phi>\phi_c$. (2) If $\phi+\gamma<\phi_c$, $P(\infty)=0$. (3) If $\phi+\gamma>2\pi-\phi_c$, $P(\infty)=0$.

I define $P(\infty|\phi,\phi_c,\gamma)$ as the probability of ejection,
given ϕ, ϕ_c, and γ. ϕ_c is determined by U, which is
assumed known. If the probability of deflection through
an angle γ is assumed independent of ϕ, we can write

$$P(\infty|U) = \int \int P(\infty|\phi,\phi_c,\gamma) \cdot P(\gamma|U) \cdot P(\phi|U) \, d\gamma \, d\phi.$$

The probability of scattering through an angle γ
can be found from the Rutherford scattering formula
(Landau and Lifshitz, 1960), by substituting the gravit-
ational potential for the Coulomb potential. The differ-

-67-

ential scattering cross-section is

$$d\sigma = \pi (Gm_p/u^2)^2 \cos (\gamma/2) \sin^{-3} (\gamma/2) \, d\gamma. \qquad (3.9)$$

Since

$$u = (Gm_\odot/a_p)^{1/2} U, \qquad (3.10)$$

Eq. (3.9) can be written

$$d\sigma = \pi a_p^2 (M_p/U^2)^2 \cos (\gamma/2) \sin^{-3} (\gamma/2) \, d\gamma. \qquad (3.11)$$

The total cross-section is

$$\sigma = \int_{\gamma_{min}}^{\gamma_{max}} d\sigma = a_p (\frac{M_p}{U^2})^2 [\sin^{-2} (\frac{\gamma_{min}}{2}) - \sin^{-2} (\frac{\gamma_{max}}{2})], \qquad (3.12)$$

where γ_{min} and γ_{max} are the smallest and largest possible
scattering angles. The largest possible value of γ is π,
corresponding to an impact parameter of zero. For γ_{min}
=0, σ is infinite. This singularity is characteristic of
any potential that varies as $1/r$, and simply means that a
deflection of zero requires an infinite impact parameter.
This definition of σ leads naturally to a definition of
an encounter as an angular deflection of the particle's
trajectory greater than some minimum value. The impact
parameter b is given by

$$b = (Gm_p/u^2) \cot (\gamma/2) = a_p (M_p/U^2) \cot (\gamma/2).$$

$$(3.13)$$

If an encounter is defined by a specified minimum defl-
ection, the maximum value of b varies as U^{-2}. However,

-68-

the two-body scattering formula is not valid where the perturbations by a third body are significant, and must be restricted to within the sphere of influence. All approaches which result in the particle's entering the sphere of influence will be considered encounters; a necessary condition is that $b < d$. For comparison with the results of Bandermann and Wolstencroft (1971), I adopt their definition of

$$d = 1.15 \ a_p(m_p/m_\odot)^{1/3}; \quad D = 1.15 \ M_p^{1/3}. \quad (3.14)$$

Setting $d = b$ in Eq.(3.13) gives

$$\gamma_d = 2 \ \tan^{-1}(M_p^{2/3}/1.15U^2). \quad (3.15)$$

The probability of encounter per revolution and the computed probabilities of collision and ejection per encounter will depend on the definition of d, but the ratio of collision and ejection probabilities is insensitive to the choice of d.

I assume that the initial orbit of the particle can intersect any point on the projected area of the sphere of influence with equal probability. Then the probability of a particle's angular deflection into an interval $d\gamma$, centered on γ, is

$$\frac{d\sigma}{\sigma} = \frac{\cos \ (\gamma/2) \ \sin^{-3}(\gamma/2)}{\sin^{-2}(\gamma_d/2) - 1} \ d\gamma = P(\gamma|U) \ d\gamma. \quad (3.16)$$

Note that the cross-section for any deflection greater

than some given value of γ is given by substituting that value for γ_{min} in Eq.(3.12). It follows that if γ_g is the deflection produced by a grazing collision with a planet, then the collision cross-section is

$$\sigma_c = a_p^2 (M_p/U^2)^2 [\sin^{-2}(\gamma_g/2)-1], \qquad (3.17)$$

and the collision probability per encounter is

$$P_c = \frac{\sin^{-2}(\gamma_g/2)-1}{\sin^{-2}(\gamma_d/2)-1}, \qquad (3.18)$$

and is independent of ϕ. This is acually an upper limit on the collision probability, since there are classes of orbits for which encounters are possible, but the minimum impact parameter is too large for collision. The condition for a grazing impact is

$$b^2 = r^2(1+u_e^2/u^2), \qquad (2.2)$$

Combining Eqs.(2.2), (3.10), and (3.13),

$$\gamma_g = 2 \tan^{-1}[M_p/U^2 R(1+2M_p/U^2 R)^{1/2}]. \qquad (3.19)$$

We now have the limits on γ needed for evaluation of $P(\infty | U, \phi)$. In accordance with the special cases mentioned above, we take γ_{min} to be the larger of γ_d and $(\phi_c - \phi)$, and γ_{max} to be the smaller of γ_g and $(2\pi - \phi - \phi_c)$. Then

-70-

$$P(\infty \,|\, U, \phi) = \int_{\gamma_{min}}^{\gamma_{max}} (\beta/\pi) \; d\sigma/\sigma \qquad\qquad (3.20)$$

$$= \int_{\gamma_{min}}^{\gamma_{max}} \frac{\cos^{-1}[(\cos\gamma \; \cos\phi - \cos\phi_c)/\sin\gamma \; \sin\phi]\cos(\gamma/2)}{\pi \; \sin^3(\gamma/2) \; [\sin^{-2}(\gamma_d/2) - 1]} \; d\gamma.$$

For some problems, it is desirable to know the mean and root-mean-squared deflections per encounter. Considering only encounters without collisions, the cross-section is

$$\sigma_{nc} = \pi a^2 (M_p/U^2)^2 [\sin^{-2}(\gamma_d/2) - \sin^{-2}(\gamma_g/2)], \qquad (3.21)$$

and the mean deflection is

$$\overline{\gamma} = \frac{1}{\sigma_{nc}} \int_{\gamma_d}^{\gamma_g} \gamma \; d\sigma(\gamma), \qquad\qquad (3.22)$$

With $d\sigma$ from Eq. (3.9), this can be integrated by parts to give

$$\overline{\gamma} = [\gamma_d \; \sin^{-2}(\gamma_d/2) + 2 \; \cot(\gamma_d/2) - \gamma_g \sin^{-2}(\gamma_g/2)$$
$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (3.23)$$
$$- 2 \; \cot(\gamma_g/2)]/[\sin^{-2}(\gamma_d/2) - \sin^{-2}(\gamma_g/2)].$$

Similarly, the mean squared deflection is

$$\overline{\gamma^2} = \frac{1}{\sigma_{nc}} \int_{\gamma_d}^{\gamma_g} \gamma^2 \; d\sigma(\gamma) = \{\gamma_d^2 \; \sin^{-2}(\gamma_d/2) - \gamma_g \; \sin^{-2}(\gamma_g/2)$$

$$+ 4[\gamma_d \; \cot(\gamma_d/2) - \gamma_g \; \cot(\gamma_g/2)] + 8 \; \ln[\sin(\gamma_g/2)/$$

$$\sin(\gamma_d/2)]\}/[\sin^{-2}(\gamma_d/2) - \sin^{-2}(\gamma_g/2)]. \qquad (3.24)$$

The mean and rms values of γ are typically about two and three times γ_d.

Equation (3.20) is suitable for evaluating the probability of ejection for a particular object for which ϕ is known. For encounters by a large population of objects, or repeated encounters by a single body, some distribution of ϕ must be assumed. The simplest assumption is that of "equipartition" in the orientation of the relative velocity vector, i.e., that all directions of $\underline{U}$ are equally probable, with the exclusion of the escape cone. This assumption was made explicitly by Öpik (1966a, 1966b), and implicitly by Bandermann and Wolstencroft (1971). The resulting distribution for ϕ is then

$$P(\phi) \, d\phi = \sin \phi \, d\phi / (1 - \cos \phi_c). \qquad (3.25)$$

This expression overestimates the number of particles in nearly parabolic orbits. The loss of particles into the escape cone by both close and distant encounters tends to deplete the population with ϕ near ϕ_c. A realistic distribution requires $P(\phi)$ to go to zero at ϕ_c, with $P(\phi)$ linear in $\cos \phi$(or $1/A$) near ϕ_c (Everhart, 1973a). In most of the calculations presented here, I have used Eq. (3.25), in order to compare my results with those of Öpik and Bandermann and Wolstencroft. The calculated ejection probabilities can be considered upper limits. The effect

of other distributions of ϕ on the calculated ejection

probabilities will be discussed below. One advantage of

this approach is the ease with which different distribut-

ions of ϕ may be used for different classes of particles.

We shall see that the probabilities of collision and

ejection per encounter are both small for all orbits of

interest. Since an "old" population of particles will

have undergone many encounters, with the rms deflection

small compared to the allowed domain of ϕ, we can expect

approximate equipartition except near ϕ_c. Comets are an

exception, since their observable lifetimes are not

limited by collisions or ejection, but by their disint-

egration. Lowrey (1973) classified short-period comets

by their velocities relative to Jupiter, and their values

of ϕ. High-velocity comets (U>1) were not found with

small values of ϕ, though low-velocity comets had appar-

ently reached equipartition. The average deflection per

encounter decreases with increasing U, so the number of

encounters needed to reach equipartition increases with

U. Apparently, when U>1 the time required exceeds the

visible lifetime of the comet. The numbers of comets

visible as functions of U and ϕ could provide an estimate

of their visible lifetimes, if one could compensate for

observational selection effects. Note that this result

is consistent with the conclusion of Everhart (1973a)

that long-period comets ($\phi \approx \phi_c$) evolve into short-period comets, but not vice versa.

With the distribution of ϕ given by Eq. (3.25), the average ejection probability per encounter for a given value of U is

$$P(\infty|U) = \int_{\gamma_{min}}^{\gamma_{max}} \int_0^{\phi_c} \frac{\cos^{-1}\{\frac{\cos\gamma\,\cos\phi - \cos\phi_c}{\sin\quad\sin}\}\cos(\gamma/2)\sin\phi\;d\phi\;d\gamma}{\pi\,\sin^3(\gamma/2)\,(1-\cos\phi_c)\,(\sin^{-2}(\gamma_d/2)-1)}$$

(3.26)

The value of $P(\infty|U)$ must be evaluated numerically. The limits of integration have been defined above. Only the limits on γ depend on the properties of the different planets. The region of integration on the ϕ,γ plane ~~are~~ is shown schematically in Figure 8. The heavily outlined region is the range for ejection. For $\gamma < (\phi_c - \phi)$, the U vector cannot reach the escape cone. For $(2\pi - \phi_c - \phi) < \gamma < \gamma_g$, the U vector is turned completely through the escape cone, and ejection does not occur (this outcome is significant for small values of U, when the escape cone is narrow, and turning angles are relatively large).

The lifetime of a particle is determined by the rate of encounters, as well as the probabilities of ejection and collision. Öpik (1951, Eq. 18) calculated the probability of encounter per revolution of the particle, for orbits assumed to intersect. In our notation, Öpik's result becomes

-74-

$$P_e \simeq D/(4 \sin \phi). \qquad (3.27)$$

This approximate formula is not valid for very small values of $\sin \phi$ (P_e obviously cannot exceed unity). If we assume the distribution of Eq. (3.25), then the mean probability of encounter per revolution for a population of particles which has attained equipartition would be

$$\overline{P_e} \simeq \int_0^{\phi_c} P_e(\phi) \; P(\phi) \; d\phi = D\phi_c/4(1-\cos \phi_c). \qquad (3.28)$$

The average time between encounters can be found if the mean orbital period is known. The average semi-major axis is found from Eq. (3.6):

$$\overline{1/A} = \int_0^{\phi_c} (1-U^2+2U \cos \phi) \; P(\phi) \; d\phi. \qquad (3.29)$$

Since $\cos \phi_c = (U^2-1)/2U$, we find

$$\overline{(1/A)} = U+(1-U^2)/2, \qquad (3.30)$$

and the mean time between encounters (in revolutions of the planet) is

$$\overline{T} = \int_0^{\phi_c} (1/A)^{-3/2} \; P_e^{-1} \; P(\phi) \; d\phi$$

$$= \int_0^{\phi_c} (1-U^2+2U \cos \phi)^{-3/2} \sin \phi \; P(\phi) \; d\phi. \qquad (3.31)$$

If $P(\phi)$ is given by Eq. (3.26), this integral is not finite. $P(\phi)$ must go to zero at ϕ_c at least as rapidly as $1/A$ for convergence. If we assume equipartition for

-75-

$\phi < \phi_c - 2\gamma_{rms}$, and $P(\phi)$ approaching zero linearly in $\cos \phi$

for larger values of ϕ, this integral can be evaluated

numerically. For $U=0.45$, 1.0, and 2.0, $\overline{T}$ is about 120,

130, and 280 revolutions, respectively, for Jupiter. Eq.

(3.31) can be considered a lower limit for the mean time

between encounters, since perturbations will keep the

particle orbits from intersecting the planet's orbit

for much of the time.

C. Results

Equation (3.26) was integrated numerically, using

Simpson's rule in a computer program. The probabilities

of collision and ejection as functions of U are shown for

the giant planets in Figure 9. The ejection probability

curves have nearly identical shapes for all of the planets.

This result was unexpected, since the limits of integration

are quite different for the different planets. The eject-

ion curve for Jupiter matches quite closely that given by

Bandermann and Wolstencroft (1971), though my values are

about 20% higher. In view of the different methods used,

I do not consider this difference to be significant.

The ejection probability curves are remarkably flat.

Over most of the possible range of U, the ejection prob-

ability varies by less than a factor of three. As

expected, it goes to zero at $U=0.414$, and to unity at

$U=2.414$, but for $0.6 < U < 1.8$, the ejection probability act-

ually decreases with increasing U. One might intuitively

expect it to increase monotonically with U, since the volume of the escape cone increases, and the required deflection decreases. However, Eq. (3.13) shows that the cross-section for a given deflection varies as U^{-4}. Over most of the range of U, this effect dominates.

Öpik (1961, 1963) has stated that the ejection probability is proportional to the volume of the escape cone. Bandermann and Wolstencroft (1971) erroneously interpreted this statement as a claim that the escape probability was simply the fractional solid angle of the escape cone, equal to $(U^2+2U-1)/4U$, and shown in Fig. 9. Actually, Öpik stated that the ejection probability was proportional to this solid angle after randomization of the U vector. This randomization requires a sufficient number of encounters so that the rms total deflection equals $\pi/2$. Öpik's expression for the average cross-section for this deflection was $\pi B \ln[(D^2+B)/(S^2+B)]$, where $B=16M_p^2/\pi^2 U^4$, and $S^2 = R^2(1+2M_p/RU^2)$. Dividing this cross-section by πD^2 and multiplying by the fractional solid angle of the escape cone evidently gives an average ejection probability per encounter of

$$P(\infty) = \frac{(U^2+2U-1)}{4U} \frac{B}{D^2} \ln[(D^2+B)/(S^2+B)], \qquad (3.32)$$

though I am unaware of any explicit statement of this
equation by Öpik. Eq. (3.32) was evaluated for Jupiter,
using the definition of D in Eq. (3.14) (Öpik's definit-
ion is somewhat different). The resulting curve is lab-
eled "Jupiter (Öpik)" in Fig. 9. It is similar in form
to my result for low values of U, but deviates widely at
large values, and does not turn upward. The large differ-
ence in computed probabilities of ejection apparently
rsults from Öpik's (1961) assumption that all deflections
are small. Large deflections, though rare, are very
effective for ejection. In the case of the terrestrial
planets, the small-deflection assumption should be more
realistic. However, the agreement between the two app-
roaches is even poorer for those planets. The results
for the terrestrial planets are shown in Figure 10. The
curve labeled "Mars (Öpik)" was evaluated using Eq. (3.32).
In this case, the solid angle of the escape cone is not
a good measure of the escape probability. When most
deflections are small, escape is possible only through a
narrow annulus with $\phi \simeq \phi_c$, rather than through the entire
cone. The assumption that an accumulated deflection of
$\pi/2$ is necessary for ejection underestimates the ejection
probability for a population of particles which has ach-
ieved equipartition and has some members with $\phi \simeq \phi_c$. How-
ever, equipartition is unlikely for encounters with the

terrestrial planets, as we shall see. The results in Fig.
10 are probably a considerable overestimate of the eject-
ion probability.

The collision probability curves for the Jovian and
terrestrial planets are of different types, due to diff-
erences in both their masses and orbits. The inner
planets have higher orbital velocities, hence a given
value of U corresponds to a higher absolute velocity.
For relative velocities much greater than the escape vel-
ocity, the collision cross-section is essentially equal
to the geometric area. The terrestrial planets have
orbital velocities several times their escape velocities,
so for U greater than about 0.5, the collision probabil-
ities are nearly constant. The giant planets all have
escape velocities several times larger than their orbital
velocities, so collision probabilities vary rather strongly
with all allowed values of U.

The giant planets have ejection cross-sections more
than three orders of magnitude greater than their collis-
ion cross-sections. For the terrestrial planets, the
formal collision probability shown in Fig.10 is only about
one order of magnitude less than that for ejection. For
Mercury, the two probabilities are about equal; the coll-
ision probability is large because the sphere of influence
is so small. This difference is qualitative, as well as

quantitative. Figure 11 shows γ_d and γ_g for the earth
and Jupiter as functions of U, as well as ϕ_c. For vir-
tually all values of U, γ_g for Jupiter is larger than ϕ_c.
Therefore, Jupiter is capable of ejecting particles with
small initial values of ϕ. The other giant planets also
have this ability. The earth and the other terrestrial
planets all have $\gamma_g << \phi_c$; they can therefore eject only
particles with nearly parabolic initial orbits. A part-
icle with an initially small ϕ must have its orbit evolve
through the random walk described by Öpik before it can
be ejected. During this process, the particle will usu-
ally evolve into a Jupiter-crossing orbit, if it does not
first collide with a terrestrial planet. The probability
of ejection by Jupiter is so large that direct ejection
by terrestrial planets is of no real significance. None
of the known Apollo asteroids can be ejected by a single
encounter with a terrestrial planet; their lifetimes are
determined primarily by the probability of collision with
the terrestrial planets. Ejection from near-parabolic
orbits accounts for a significant part of the computed
ejection probability for all planets, but for the giant
planets, ejection is more probable than collision, even
for small values of ϕ.

As stated above, the assumption of complete equipart-
ition overestimates the number of particles in nearly

parabolic orbits, and therefore Figures 9 and 10 represent upper limits for the ejection probabilities. I have also evaluated $P(\infty)$ using the more realistic assumption that equipartition prevails except when ϕ is within a few times γ_{rms} of ϕ_c, with $P(\infty)$ going to zero at ϕ_c, linearly in $\cos \phi$. The result is not sensitive to the value of the transition point. The curve of $P(\infty)$ is lowered by about one order of magnitude, but the shape is not changed significantly. This assumption brings the calculated ejection probability curves into better agreement with Öpik's results, but the agreement is fortuitous. The flattened shape of the ejection probability curves does not depend strongly on the assumed distribution of ϕ. If we evaluate $P(\infty|U,\phi)$ by Eq. (3.20) as a function of U with ϕ fixed, we find a curve of similar shape, but rising to unity at the value of U for which the chosen value of ϕ is equal to ϕ_c. The curve of $P(\infty|U)$ will be of this same general shape for any reasonably smooth distribution of ϕ, and will be rather flat for most of the allowed range of U.

From Eq. (3.4), we have U as a function of A, e, and i. Figure 12 shows the contours of constant U on an a-e diagram for $i=0$. For nonzero values of i, the U contours are shifted to the left. A deflection in an encounter can be considered as a displacement on a surface of constant U in A-e-i space.

D. Discussion

The methods developed here occupy a place between the analytic approach of Öpik and the Monte Carlo techniques and numerical integration of orbits used by Arnold (1965) and Everhart (1973a, 1973b). The latter are more useful for determining the absolute lifetimes of particles, since for that we must know the probability that the orbtis will intersect in the first place. That probability is determined mostly by distant encounters and the perturbations of other planets. Analytic approximations of sufficient accuracy for many purposes should be possible (Zimmerman and Wetherill, 1973). The direct integration of orbits used by Everhart may be the most useful in this respect, but such methods are not practical for determining collision probabilities, due to the extremely low rate of collisions. I caution against applying these formulas to individual objects, since in many cases exact or approximate commensurabilities exist, invalidating the assumption of random encounters. Many of the small bodies in the solar system are protected from close encounters with planets; it is for precisely this reason that they have survived to the present day.

Some applications of these methods have been mentioned in passing. The visible lifetimes of short-period comets may be estimated from their distribution of ϕ, if

allowance can be made for observational selection, and the possible tendency for successive encounters with Jupiter to be correlated. Öpik (1963) concluded that the time scale for dynamical elimination of comets was three to five orders of magnitude greater than the time scale for their disintegration. The fact that thousands of dead comet nuclei are not observed led him to conclude that most comets disintegrate completely, leaving no solid bodies of asteroidal size. My computed ejection probabilities are greater than Öpik's, easing this problem to some extent. However, the Apollo asteroids, which may be dead comet nuclei (Öpik, 1963), are eliminated chiefly by collision with the terrestrial planets. Their calculated lifetimes are not changed by my results, except that deflection to Jupiter-crossing orbits may be somewhat more probable that calculated by Öpik.

If comets originated within the solar system, and were ejected into distant orbits by encounters with the giant planets (Oort, 1950; Öpik, 1973), we have in principle a means of calculating the efficiency of this process and limits on the mass of the comet cloud. Some estimates already exist (Safronov, 1970; Öpik, 1973), but a better estimate should be possible. The total mass and angular momentum lost from the solar system by the ejection of comets may have been large enough to be of cosmogonical significance (Levin, 1972b).

The methods developed in this chapter are directly applicable to the problem of possible origin of short-period comets by capture of parabolic comets. Everhart (1969) has investigated such captures be close encounters with giant planets, using conic matching and some direct integration of orbits for large numbers of random parabolic initial orbits. He also followed the orbital evolution of comets captured in this way by integration of their orbits (Everhart, 1972). While this approach offers greater potential accuracy, considerable insight may be gained from two-body scattering models, with a considerable saving in computation time. A quantitative treatment will not be attempted here, but some useful qualitative results can be developed immediately. It is known that most short-period comets have rather low values of U with respect to Jupiter, typically less than 0.6 (Lowrey, 1973). Strömgren (1947) suspected that such low-velocity comets were captured from parabolic orbits of low inclination, with perihelia near Jupiter's orbit. Such orbits have U near the minimum value of 0.414. This was confirmed by Everhart (1969). Strömgren suggested that their preferential capture was due to their spending more time near Jupiter's orbit, making an encounter more likely. However, there is another important effect. A parabolic comet has $\phi = \phi_c$, orginally. In the limit of

small deflections, the capture probability in a single

encounter is 0.5. When the deflection is comparable in

magnitude to the size of the escape cone, the capture

likelihood is larger. Whenever $\gamma \gtrsim 2(\pi - \phi_c)$, the capture

probability is unity. For Jupiter, the mean deflection

is comparable to the size of the escape cone for $U \simeq 0.5$.

Considerably more than half of all low-velocity encounters

result in capture. Such large-deflection captures also

result in more stable orbits; another close encounter

is necessary to eject the comet. Those captured by small

deflections can be ejected again by small perturbations,

without re-entering the planet's sphere of influence.

Note that most such captures and re-ejections are not

observable from the earth; ϕ must be fairly small for a

comet to become visible. Some of these considerations

will be mentioned in the next chapter. We shall see that

the results developed here can offer an explanation for

the anomalously low masses of Mars and the asteroid belt.

IV. MASS LOSS FROM THE REGION OF MARS AND THE ASTEROID BELT.

We have seen in chapter 1 that the total mass of
Mars and the asteroids is much lower than the amount
which was probably condensed in that region. An ad hoc
local minimum in the nebular density would be dynamically
stable (Kuiper, 1956), but the origin of such a feature
during formation of the nebula is unexplained. Ter Haar
(1972) actually predicts a local maximum in nebular
density at 2AU; the results of chapter 1 suggest a
monotonic variation. We have seen in chapter 2 that the
small mass of Mars implies an unreasonably long accretion
time for that planet for a closed feeding zone model.
The removal of mass from its zone appears necessary.
The mass ratios: earth/ Mars/ asteroids suggest that
such a process was much more effective in the zone of
Mars than in that of the earth, and was nearly complete
in the asteroid belt. The magnitude of the deficiency
of mass in that region is seldom appreciated; perhaps
Mars seems large in our consciousness because we possess
a relatively large amount of data about that planet.
Actually, 95% of the mass contained in the terrestrial
planets lies within 1AU of the sun; the much larger area
between the earth and Jupiter contains only 5%.

Jupiter has long been suspected of causing this state (Kuiper, 1951), but the actual mechanism has not been adequately explained. Since the orbits of Mars and the existing asteroids are obviously stable over very long time scales, direct gravitation perturbations by Jupiter could not have removed the excess mass from these regions. The relative velocities of planetesimals would have been increased by these perturbations. This would have slowed accretion in the asteroid belt, but might not have prevented it, particularly if much more mass was originally present. Such a process would not have been effective in the zone of Mars. Safronov (1972, chs. 9,13) has suggested that matter was "swept away" from the zones of Mars and the asteroids by bodies which originated in the zone of Jupiter and were perturbed into eccentric orbits by that planet as it grew. A more detailed examination of this process indicates that this mechanism would have the required properties.

Orbits in the present asteroid belt, with semi-major axes less than 3.3 A.U. (the resonance at 1/2 of Jupiter's period) are stable. Except for certain commensurable orbits, those beyond 4.0 A.U. are unstable, and subject to close approaches to Jupiter (Birn, 1973; Lecar and Franklin, 1973). Orbits between 3.3 and 4.0 A.U. are probably unstable on a longer time

scale. Many rocky (and possibly icy) bodies of asteroidal size would have condensed from the solar nebula in these regions, as evidenced by the numerous observed Trojan asteroids (Van Houten et al., 1970). To avoid confusion with presently existing asteroids and comets, I refer to these bodies as "projectiles". The formation of Jupiter caused most of these bodies to be scattered into other parts of the solar system. An intense bombardment before planetary accretion could have disrupted planetesimals in those regions. Such a bombardment could have been intense; one terrestrial mass, if divided into kilometer-sized objects, could produce some 10^{13} projectiles, plus secondary collisional fragments. The disruption of the planetesimals, with the resulting decrease in the mean size, could have cause preferential removal of matter from that zone. Depending on their sizes, fragments could be removed by radiation pressure, the Poynting-Robertson effect, nebular gas drag, or the Yarkovsky effect (Öpik, 1951; Peterson. 1975). The difference in effects on Mars and the earth is much greater than that which can be attributed to the geometrical factors of the increased distance from Jupiter and smaller size of the earth's zone. However, there is an effect which would cause the total bombardment flux to vary by about two orders of magnitude between Mars and the earth.

Consider a projectile in a Jupiter-crossing orbit.
Its heliocentric orbit is determined by the magnitude
and direction of its velocity relative to Jupiter at the
point of intersection (Öpik, 1951, 1963; Lowrey, 1973;
ch. 3, above). The minimum perihelion distance for any
value of U is attained when the U-vector is opposite in
direction to Jupiter's orbital motion. The projectile
then has its aphelion at Jupiter's orbit, and perihelion
at

$$q_{min} = (1-U)^2/(1+2U-U^2), \qquad (2.25)$$

in units of Jupiter's orbital radius. Whenever the pro-
jectile encounters Jupiter, the direction of the U-vector
is changed; the process can be described in terms of a
"random walk" (Öpik, 1963). When $U < 0.414$, the projectile
can be eliminated only by collision with Jupiter or
another body. For U greater than this critical value, an
encounter with Jupiter can put the projectile into a
hyperbolic heliocentric orbit, ejecting it from the solar
system. Figure 13 shows the probabilities of collision
and ejection as functions of U. Whenever ejection is
possible, it is more probable than collision with Jupiter,
by two to three orders of magnitude.

Figure 14 shows q_{min} as a function of U, and the
positions of the asteroid belt and terrestrial planets.

The critical value of U corresponds to q_{min}=1.1 AU. In order to reach the earth, a projectile with aphelion at or beyond Jupiter's orbit must have U>0.42 at Jupiter. To reach Mars, U needs only to be about 0.3. The lifetime of a projectile is determined chiefly by encounters with Jupiter. If U remained constant, then Fig. 12 implies that the lifetime of a potentially Mars-crossing project- ile would be about 100 times that of a potential earth- crosser. Only a small fraction of those projectiles with sufficiently large values of U actually cross the orbits of the earth at any instant, but encounters with Jupiter can allow any of them to do so eventually.

Actually, U is not constant; the eccentricity of Jupiter's orbit, and its inclination to the invariable plane, cause U to vary slightly between successive encoun- ters. After many encounters, the average effect is an increase in U (Arnold, 1965; Öpik, 1966a). According to Öpik's statistical theory, several thousand encounters, on a time scale of 10^5 to 10^6 years, are required to increase U from 0.1 to 0.4. The rate of acceleration decreases as U increases, so U > 0.3 for most of that time. Öpik's estimate of the time scale is probably too small, since he uses the present value of Jupiter's eccentricity, near the maximum, and too large a value of the inclination $(0.027 \leqslant e_J \leqslant 0.062; 0.004 \leqslant \sin i_J \leqslant 0.008$ (Brouwer and

Clemence, 1961); Öpik used $e_J = 0.05$, $\sin i_J = 0.02$).

Most of the projectiles will escape collision with Jupiter

during this acceleration. When U exceeds 0.42, the time

scale for elimination of potential earth-crossers will

be about 100 encounters, or 10^4 years. The absolute time

scale depends on the rate of encounters with Jupiter,

which will not differ greatly between Mars-crossers and

earth-crossers. The ratio of their lifetimes is insen-

sitive to the encounter rate, since the rates of accel-

eration, collision with Jupiter, and ejection are all

proportional to the encounter rate. A hypothetical

proto-Jovian core of smaller mass (Perri and Cameron,

1973; Kaula and Bigeleisen, 1975) would produce the same

effects on a longer time scale, with a larger fraction

of projectiles lost by collision with Jupiter.

The projectiles would have original values of U near

0.1. The acceleration by Jupiter would cause the region

of the present asteroid belt to be bombarded first, with

impact velocities of several km sec^{-1}. The zone of Mars

would receive a less severe bombardment, since for

$U > 0.15$, some projectiles could have aphelia at Saturn's

orbit and be eliminated there. Most projectiles with

$U > 0.414$ would be ejected by Jupiter, with little addit-

ional acceleration. Öpik (1966a) estimates the fraction

surviving to $U=0.6$ at 10^{-8}. This is considerable over-

estimate; the rate of acceleration is smaller than Öpik's calculation, as stated above, and the ejection probability is greater (Weidenschilling, 1975a; ch. 3, this thesis). The bombardment of Mercury and Venus from this source would be negligible; the craters of Mercury must have had a different source, possibly bodies scattered by multiple emcounters with other planets. There may have been significant bombardment of the earth's zone, with some effect upon the earth (Wetherill, 1972) or the moon (Kaula and Bigeleisen, 1975), but of much lower intensity than in the zone of Mars. The lunar highlands appear too young to show evidence of this bombardment (Hartmann, 1975), which must have occurred some 4.5×10^9 years ago. The heavily cratered areas of Mars are probably more recent, also.

This bombardment mechanism does not affect the presently observed meteorite flux at the earth or Mars, since meteoroids originating in the present asteroid belt can reach earth-crossing orbits without encountering Jupiter (Wetherill, 1969, 1974a, 1974b). There is, however, one possibly significant effect upon the Martian cratering rate. The main bombardment was a sudden event, lasting only about 10^6 years. It was probably triggered by the dissipation of the solar nebula, in which drag would have inhibited the acceleration of projectiles.

Possibly, the formation of Jupiter by hydrodynamic
collapse (Perri and Cameron, 1973) was responsible.
However, some projectiles probably had original orbits
which were only unstable on a much longer time scale,
such as horseshoe or Trojan type orbits (Everhart, 1973b).
Bodies "leaking" from such nearly stable orbits would
have been much more likely to encounter Mars than the
other terrestrial planets before ejection. If such
bodies were sufficiently numerous, the cratering flux
at Mars could have been much higher than the lunar flux
for a significant fraction of the age of the solar
system (I am indebted to W. K. Hartmann for this sugges-
tion). Martian chronology based on comparison of lunar
and Martian crater counts may be in error from this
cause.

Rabe (1971) has suggested such a Trojan origin for
some of Jupiter's family of short-period comets. While
there are probably other adequate sources of periodic
comets, his conclusion points out the dynamical simi-
larity between the "projectiles" (after significant
acceleration) and the short-period comets. The observed
distribution of cometary perihelia supports the general
concept of a "barrier" near 1AU for Jupiter-influenced
objects. Fig. 15 shows the distribution of perihelia
for observed parabolic and periodic comets (Marsden,

1975). According to the theory of Oort (1950), the histogram for parabolic comets should be flat. The peak near 1 AU is therefore an indication of discovery selection effects. Apparently, the discovery probability for q=1.5 AU is only about one fourth that for q=1 AU. The periodic comets show an entirely different distribution of q. Even without correction for discovery selection, there is an excess of periodic comets with q > 1 AU. If the selection factor for parabolic comets is applicable, the cometary flux at 1.5 AU is more than an order of magnitude grater than that inside the earth's orbit. Actually, the selection factor may be even greater for periodic comets, since their brightness tends to vary more rapidly with heliocentric distance (Oort and Schmidt, 1951). Note also that the majority of periodic comets with q < 1 AU have Q > 10 AU, suggesting that their perihelia were reduced by the effects of the giant planets other than Jupiter. It is impossible to demonstrate or disprove this suggestion conclusively, since the orbits of individual comets cannot be integrated backward with sufficient accuracy. However, the limited lifetimes against disintegration suggest that relatively few comets with Q < 10 have been strongly influenced by planets other than Jupiter.

The time scale for the main bombardment is much

shorter than that for planetary accretion by gravitational forces (Safronov, 1972, ch. 9; Weidenschilling, 1974, 1975c; this thesis, ch. 2). Two scenarios are possible: the bombardment might have removed the excess mass from the zone of Mars before accretion began, or might have terminated accretion which was already in progress. We have seen that the first case leads to extremely long accretion times, in excess of 2 b.y. While such a time scale cannot be definitively ruled out on the basis of evidence from Mars alone, the lack of a heavy lunar cratering flux in this interval is a strong argument against it. In the second, more plausible case, Mars might have attained, say, 90% of its present mass before the bombardment. This scenario suggests that some significant accretion in Mars' zone took place before dissipation of the solar nebula, but not in the asteroid region. The post-bombardment fragments in its zone probably exceeded the mass left in the asteroid belt. If some 20% of the mass of Mars remained, roughly half would be accreted by Mars, on a time scale of a few times 10^8 yr. The rest would be scattered into earth-crossing orbits after a delay of 10^8 yr. or more. Encounters with the earth would lead to Venus- and Mercury-crossing orbits. This scenario is compatible with the intense lunar cratering of 4 b.y. ago, in both

timing and amount. It would also result in the delayed
post-accretion bombardment of Mercury suggested by
Murray et al. (1975). Wetherill (1975b, 1975c) suggested
a bombardment by bodies derived from Mars-crossing orbits
are a possible source of pre-mare lunar cratering; a
necessary consequence of his model is a much heavier
bombardment of Mars during the same period.

The total mass scattered among the terrestrial
planets during this late bombardment period was not
large. Wetherill (1975b, 1975c) suggests a mass on the
order of 10^{-4} earth masses, not including that which
impacted Mars. Actually, this figure is merely a lower
limit. The early primary bombardment probably produced
a much larger mass of small fragments which have left
no visible cratering history. Some of this pulverized
matter probably was captured by the earth. Lewis (1972a)
suggested that the earth's H_2O content was due to its
accretion zone extending slightly beyond the inner edge
of the stability field of the hydrous mineral tremolite.
However, it seems possible that the earth's water and
volatiles were derived from the zone of Mars. Life on
our own planet may be the result of this inheritance.

FURTHER SPECULATION

From the author's admittedly biased vantage point,
the picture presented here - a low-mass nebula, "slow"
gravitational accretion, and bombardment - appears to be
generally self-consistent, without excessively blunting
Occam's razor. At our present state of knowledge, this
may be as much as one can expect from any cosmogony.
Certainly, a more quantitative treatment is desirable,
particularly for the behavior of the "projectiles" as they
are accelerated and ejected. If, for the moment, we do
accept this scenario for the formation of our own solar
system, what can we infer about the origin of other
systems?

The mechanism of formation of an "embryo" is
unknown, but the large numbers of planetesimals available
suggests that the origin and growth of a planet was not
governed by the statistics of small numbers. If the
formation of Jupiter was triggered by the presence of
solid H_2O among the condensates in the nebula, then the
same should have happened in other systems. In any nebula
with sufficient oxygen abundance, a massive planet
should form near the limit of ice formation. If the
nebula is massive enough, the planet would become a gas
giant. The resulting bombardment should cause a gap or

mass deficiency in the region beyond about one fifth of the innermost gas giant's distance from the star. This gap may be a common feature of planetary systems.

The location of the earth just inside the boundary of the "safe" region (Fig. 14) is probably coincidental. One may speculate that a possible inward drift of pulverized matter from the bombarded region might raise the local surface density and favor formation of a planet at that point, if one was not already present. Again, I lament the lack of a statistically significant number of observable solar systems.

Undoubtedly, Mars would be a more congenial abode for life, were it as massive as the earth. We would expect it to possess a massive atmosphere with a significant greenhouse effect, milder temperature, and liquid water. Since the earth was not severely affected by, and might even have benefited from, the bombardment, this particular "Jupiter effect" probably does not influence the possible abundance of life in the universe. However, it may have deprived us of the chance to have nearby neighbors.

REFERENCES

Alfven, H., and Arrhenius, G. (1970), Astrophys. Space
Sci. 9,3.

Anderson, D., and Hanks, T. (1972), Nature 237, 387.

Arnold, J.R. (1965), Astrophys. J. 141, 1536.

Bandermann, L., and Wolstencroft, R. (1970), Mon. Not.
Roy. Astron. Soc. 150, 173.

_____ (1971), Mon. Not. Roy. Astron. Soc. 152, 377.

Benfield, A.E. (1950), Trans. AGU 31, 53.

Birn, J. (1973), Astron. Astrophys. 24, 283.

Brouwer, D., and Clemence, G. (1961), in Planets and
Satellites (G. Kuiper and B. Middlehurst, eds.)
Univ. of Chicago Press, pp. 31-94.

Cameron, A.G.W. (1973a), Space Sci. Rev. 15, 121.

_____ (1973b), Icarus 18, 407.

_____ (1973c), Space Sci. Rev. 14, 383.

Cameron, A.G.W., and Pine, M. (1973), Icarus 18, 377.

Cameron, A.G.W., and Pollack, J. (1975), paper presented
at Jupiter conference, Tucson.

Chandrasekhar, S. (1960), Principles of Stellar Dynamics,
Dover Publications, New York, ch. 1.

Chapman, C.R. (1974), Icarus 22, 272.

Chapman, C.R., and Davis, D.R. (1975), Science, in press.

Dole, S.H. (1970), Icarus 13, 494.

Everhart, E. (1969), _Astron. J._ 74, 735.

_____ (1972), _Astrophys. Lett._ 10, 131.

_____ (1973a), _Astron. J._ 78, 329.

_____ (1973b), _Astron. J._ 78, 316.

_____ (1974), private communication.

Goldreich, P., and Ward, W.R. (1973), _Astrophys. J._ 183, 1051.

Hallam, M., and Marcus, A. (1974), _Icarus_ 21, 66.

Hanks, T., and Anderson, D. (1969), _Phys. Earth Planet. Int._ 2, 19.

Hartmann, W.K. (1968), _Astrophys. J._ 152, 337.

_____ (1975), _Icarus_ 24, 181.

Hartmann, W.K., and Davis, D.R. (1975), _Icarus_ 24, 504.

Hills, J.G. (1973), _Icarus_ 18, 505.

Hunten, D. (1973), _J. Atmo. Sci._ 30, 1481.

Ip, W.-H. (1974), _Astrophys. Space Sci._ 31, 57.

Johnston, D., McGetchin, T., and Toksöz, M.N. (1974), _J. Geophys. Res._ 79, 3959.

Kaula, W., and Bigeleisen, P. (1975), _Icarus_ 25, 18.

Kuiper, G.P. (1951), in _Astrophysics_ (J. Hynek, ed.), McGraw-Hill, New York, p. 357.

_____ (1956), _J. Roy. Astron. Soc. Canada_ 50, 158.

Landau, L., and Lifshitz, E. (1960), _Mechanics_, Pergamon, London, ch. 4.

Lecar, M., and Franklin, F. (1973), _Icarus_ 20, 422.

Levin, B. (1972a), Tectonophys. 13, 7.

_____ (1972b), in The Origin of the Solar System
(H. Reeves, ed.), C.N.R.S., Paris, p. 341.

Lewis, J.S. (1972), Earth Planet. Sci. Lett. 15, 286.

_____ (1975), private communication.

Lowrey, B. (1973), Astron. J. 78, 428.

Lyttleton, R.A. (1972), Mon. Not. Roy. Astron. Soc. 158,
463.

Makalkin, A.B. (1973), Solar Syst. Res. 6, 153.

Marcus, A. (1967), Icarus 7, 283.

Marsden, B. (1975), Catalogue of Cometary Orbits, Central
Bureau for Astronomical Telegrams, Cambridge.

Mc Cord, T.B., and Chapman, C.R. (1975), Astrophys. J.
197, 781.

Mizutani, H., Matsui, T., and Takeuchi, H. (1972), The
Moon 4, 476.

Murray, B., Strom, R., Trask, N., and Gault, D. (1975),
J. Geophys. Res. 80, 2508.

Oort, J. (1950), B. A. N. 11, 91.

Oort, J., and Schmidt, M. (1951), B. A. N. 11, 259.

Öpik, E.J. (1951), Proc. Roy. Irish Acad. 54A, 165.

_____ (1961), Ann. Acad. Sci. Fenn. A3 61, 185.

_____ (1963), Adv. Astron. Astrophys. 2, 219.

_____ (1966a), Mem. Roy. Soc. Liege, 5 ser. 12, 523.

_____ (1966b), <u>Adv. Astron. Astrophys.</u> <u>4</u>, 301.

_____ (1973), <u>Astrophys. Space Sci.</u> <u>21</u>, 307.

Perri, F., and Cameron, A.G.W. (1974), <u>Icarus</u> <u>22</u>, 416.

Peterson, C. (1975), <u>Bull. Amer. Astron. Soc.</u> <u>7</u>, 376.

Podolak, M., and Cameron, A.G.W. (1974), <u>Icarus</u> <u>22</u>, 123.

Rabe, E. (1971), in <u>Physical Studies of Minor Planets</u>,
    NASA SP-267, p. 407.

Reynolds, R., and Summers, A. (1969), <u>J. Geophys. Res.</u>
    <u>74</u>, 2494.

Safronov, V.S. (1960), <u>Voprosy Kosmogonii</u> <u>7</u>, 59.

_____ (1966), <u>Sov. Astron.</u> <u>9</u>, 987.

_____ (1967), <u>Sov. Astron.</u> <u>10</u>, 65C.

_____ (1970), in <u>The Motion, Evolution of Orbits,</u>
    <u>and Origin of Comets</u> (G. Chebotarev et al., eds.),
    D. Reidel, Dordrecht, p. 329.

_____ (1972a), <u>Evolution of the Protoplanetary Cloud</u>
    <u>and Formation of the Earth and Planets</u>, Israel Pro-
    gram for Scientific Translation, Jerusalem.

_____ (1972b), in.<u>The Origin of the Solar System</u>
    (H. Reeves, ed.), C.N.R.S., paris, p. 89.

Schubart, J. (1974), <u>Astron. Astrophys.</u> <u>30</u>, 289.

Schultz, P., and Gault, D. (1975), <u>The Moon</u> <u>12</u>, 159.

Scott, W. (1968), <u>J. Atmo. Sci.</u> <u>25</u>, 54.

Shmidt, O. Yu. (1958), _A Theory of Earth's Origin - Four_
Lectures (3rd ed.), Foriegn Languages Publishing
House, Moscow.

Siegfried, R., and Solomon, S.C. (1974), _Icarus_ 23, 192.

Siever, R. (1974), _Icarus_ 22, 312.

Soderblom, L., Condit, C., West, R., Herman, B., and
Kreidler, T. (1974), _Icarus_ 22, 239.

Strömgren, E. (1947), _Kgl. Danske Videnskab. Selskab,_
Mat-fys. Medd. 24, no. 5.

Szebehely, V. (1967). _Theory of Orbits_, Academic Press,
ch. 10.

ter Haar, D. (1950), _Astrophys. J._ 111, 179.

_____ (1972), in _The Origin of the Solar System_
(H. Reeves, ed.), C.N.R.S., Paris, p. 71.

Turekian, K., and Clark, S. (1969), _Earth Planet. Sci._
Lett. 6, 346.

van Houten, C., van Houten-Groeneveld, I., and Gehrels, T.
(1970), _Astron. J._ 75, 659.

Wasson, J.T. (1969), _Nature_ 223, 163.

Weidenschilling, S.J. (1974), _Icarus_ 22, 426.

_____ (1975a), _Astron. J._ 80, 145.

_____ (1975b), _Icarus_, in press.

_____ (1975c), _Icarus_, in press.

_____ (1975d), _Icarus_, submitted.

Weissman, P., and Wetherill, G.W. (1974), *Astron. J.* $\underline{79}$, 404.

Wetherill, G.W. (1969), in *Meteorite Research* (P. Millman, ed.), D. Reidel, Dordrecht, p. 573.

_____ (1972), *Tectonophys.* $\underline{13}$, 31.

_____ (1974a), *The Moon* $\underline{9}$, 227.

_____ (1974b), *Ann. Rev. Earth Planet. Sci.* $\underline{2}$, 303.

_____ (1975a), private communication.

_____ (1975b), in *Proc. Sov.-Amer. Conf. on Cosmochem. of Moon and Planets*, Moscow, in press.

_____ (1975c), *Proc. Sixth Lunar Sci. Conf.*, in press.

_____ (1975d), paper presented at Mare Basalt Conference, Houston.

Zharkov, V., Makalkin, A., and Trubitsyn, V. (1975), *Sov. Astron.* $\underline{18}$, 768.

Zimmerman, P., and Wetherill, G.W. (1973), *Science* $\underline{182}$, 51.

Zvyagina, E., and Safronov, V.S. (1972), *Sov. Astron.* $\underline{15}$, 810.

FIGURE CAPTIONS

Figure 1. Reconstructed nebular surface densities
obtained by adding H and He to restore each planet to
solar composition, and spreading the resulting masses
into contiguous zones centered on their orbits. The
meaning of the "error bars" is discussed in the text,
chapter 1.

Figure 2. Normalized mass and accretion rate vs. time
for model of Eq. (2.15) (solid lines). Accretion is 99%
complete in an interval of 10τ. Also shown is the Hanks-
Anderson accretion rate, $dr/dt \propto t^2 \sin \gamma t$, for a simi-
lar total time for accretion (dashed line).

Figure 3. Accretion of the earth for the nominal case
($U_o=0.05$, $\theta=5$), for the model of Eq. (2.35). Shown are
fraction of mass accreted (f), accretion rate (df/dt),
radial growth rate (dr/dt), and dimensionless relative
velocity of particles (U). From f=0.01 at t=0, f reaches
0.99 in 1.56×10^8 yr.

Figure 4. Effect of U_o on accretion of the earth, with
$\theta=5$. With total mass fixed, the initial space density of
matter, δ_o, varies as U_o^{-2}.

Figure 5. Effect of varying θ on accretion of the earth, for model of Eq. (2.35), with $U_o=0.05$. For $\theta=\infty$, particle velocities are constant. Finite values of θ drastically reduce the peak value of df/dt.

Figure 6. Normalized accretion rates for the earth according to Eq. (2.35), as function of f, for different values of θ. The increase of particle velocities for finite values of θ slows the later stages of accretion, causing the peak value of df/dt to occur at smaller values of f.

Figure 7. Geometry of encounter for escape. The heavily outlined spherical triangle defines the angle β, according to Eq. (3.8). The probability of escape is β/π.

Figure 8. Region of integration of Eq. (3.26) in the ϕ,γ plane (schematic) for a giant planet. For a terrestrial planet, the collision region extends to small values of γ.

Figure 9. Probabilities of collision (dashed lines) and ejection (solid lines) for the giant planets, from Eqs. (3.18) and (3.26). Also shown are fractional solid angle of the escape cone, $(U^2+2U-1)/4U$, and Öpik's ejection probability for Jupiter from Eq. (3.32).

Figure 10. Probabilities of collision (dashed lines) and ejection (solid lines) for the terrestrial planets, if equipartition is assumed. This greatly overestimates the ejection probability. Also shown is Öpik's ejection probability for Mars from Eq. (3.32).

Figure 11. Maximum (γ_g) and minimum (γ_d) deflection angles for Jupiter and the earth. All terrestrial planets have $\gamma_g << \phi_c$.

Figure 12. A-e diagram for the case i=0, showing contours of U and ϕ.

Figure 13. Probabilities of collision with Jupiter (dashed line) and ejection from the solar system (solid line) per encounter, as a function of U. When U exceeds the critical value for ejection, the ejection probability is much greater than the probability of collision.

Figure 14. Minimum possible perihelion as a function of U, for a projectile with aphelion at Jupiter's orbit. The position of the asteroid belt is shown, as are the ranges of heliocentric distances of the terrestrial planets at their greatest orbital eccentricities (Brouwer and Clemence, 1961). The vertical dashed line marks the critical velocity for ejection from the solar system.

Figure 15. Perihelion distributions for near-parabolic and short-period comets, after Marsden (1975). By Oort's theory (Oort, 1950), the histogram for near-parabolic comets should be flat; the apparent peak near 1 AU is due to discovery selection effects.
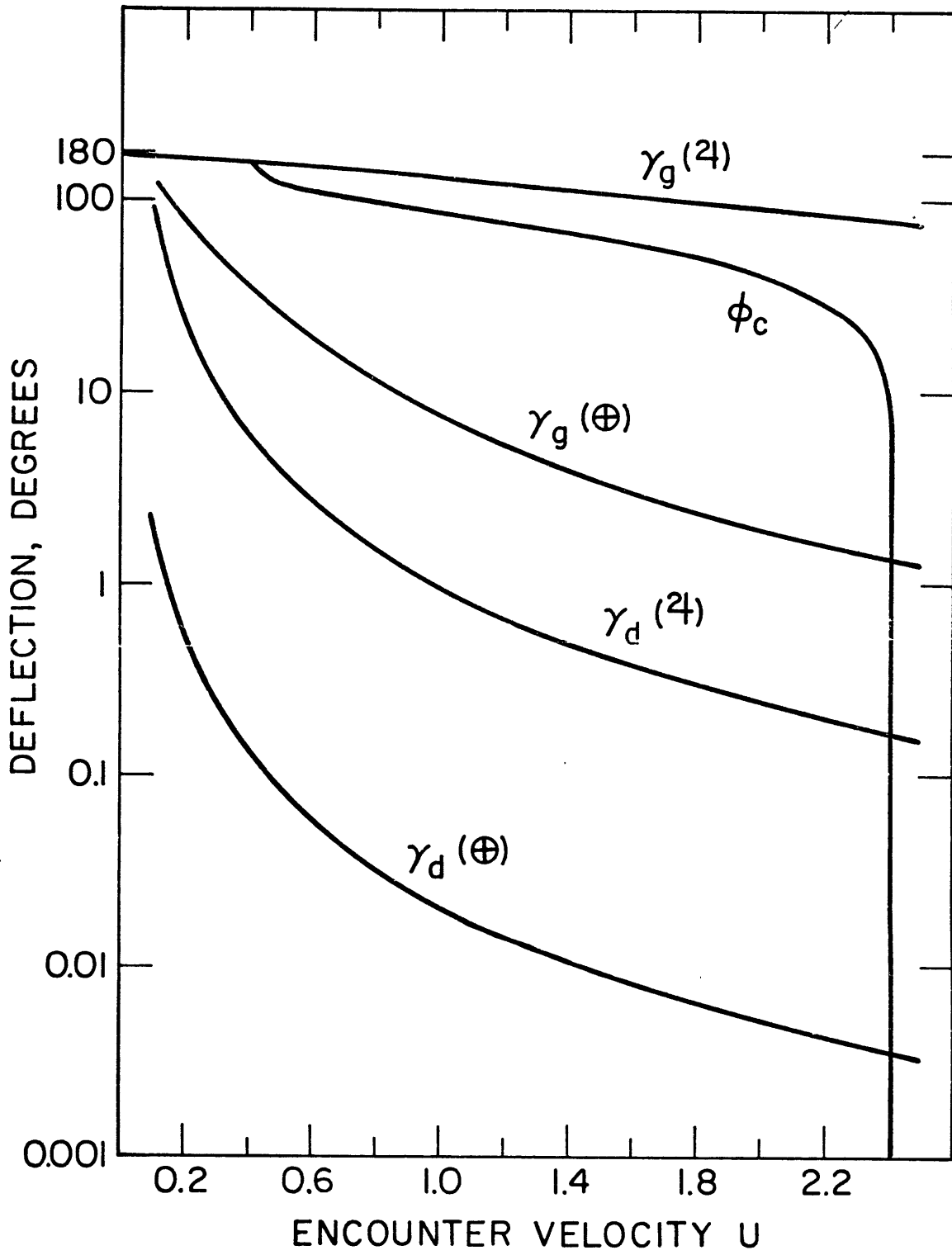
Figure 1

Figure 2

Figure 3

EARTH

$\theta = 5$

$\delta_0 = \dfrac{2.5 \times 10^{-14}}{U_0^2}$ g-cm^{-3}

Figure 4

EARTH
$U_o = .05$
$\delta_o = 10^{-11} g-cm^{-3}$

Figure 5

-113-
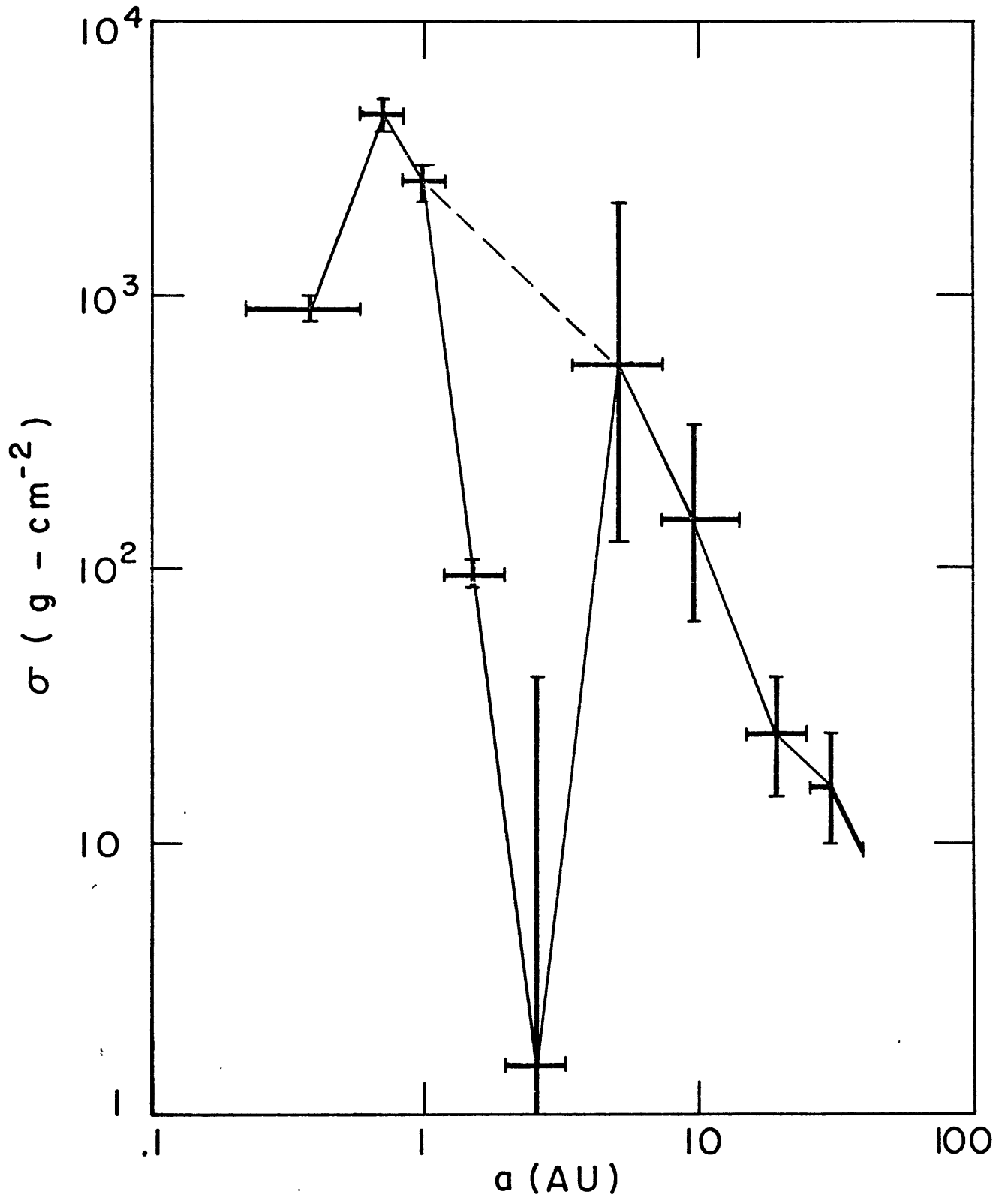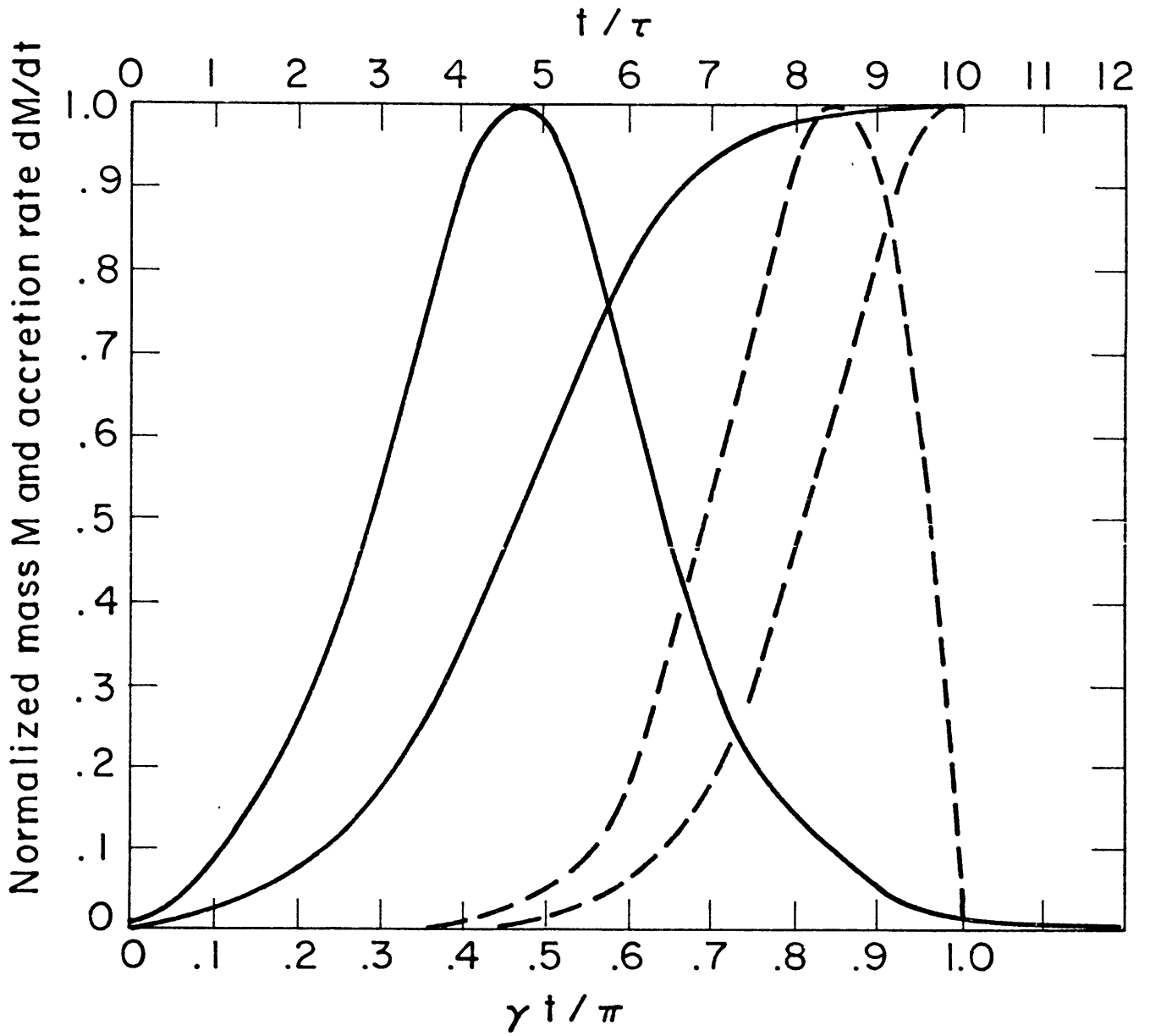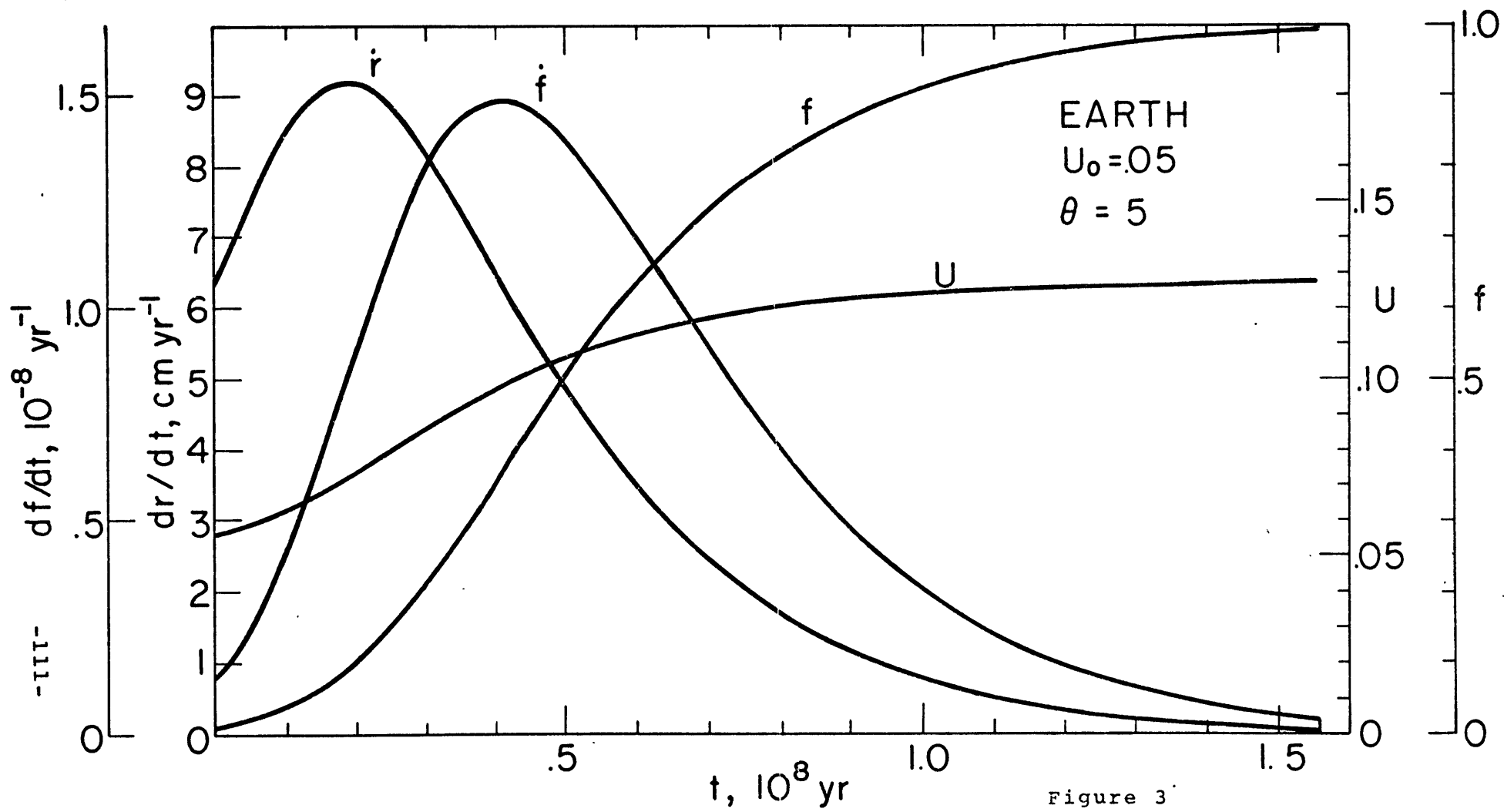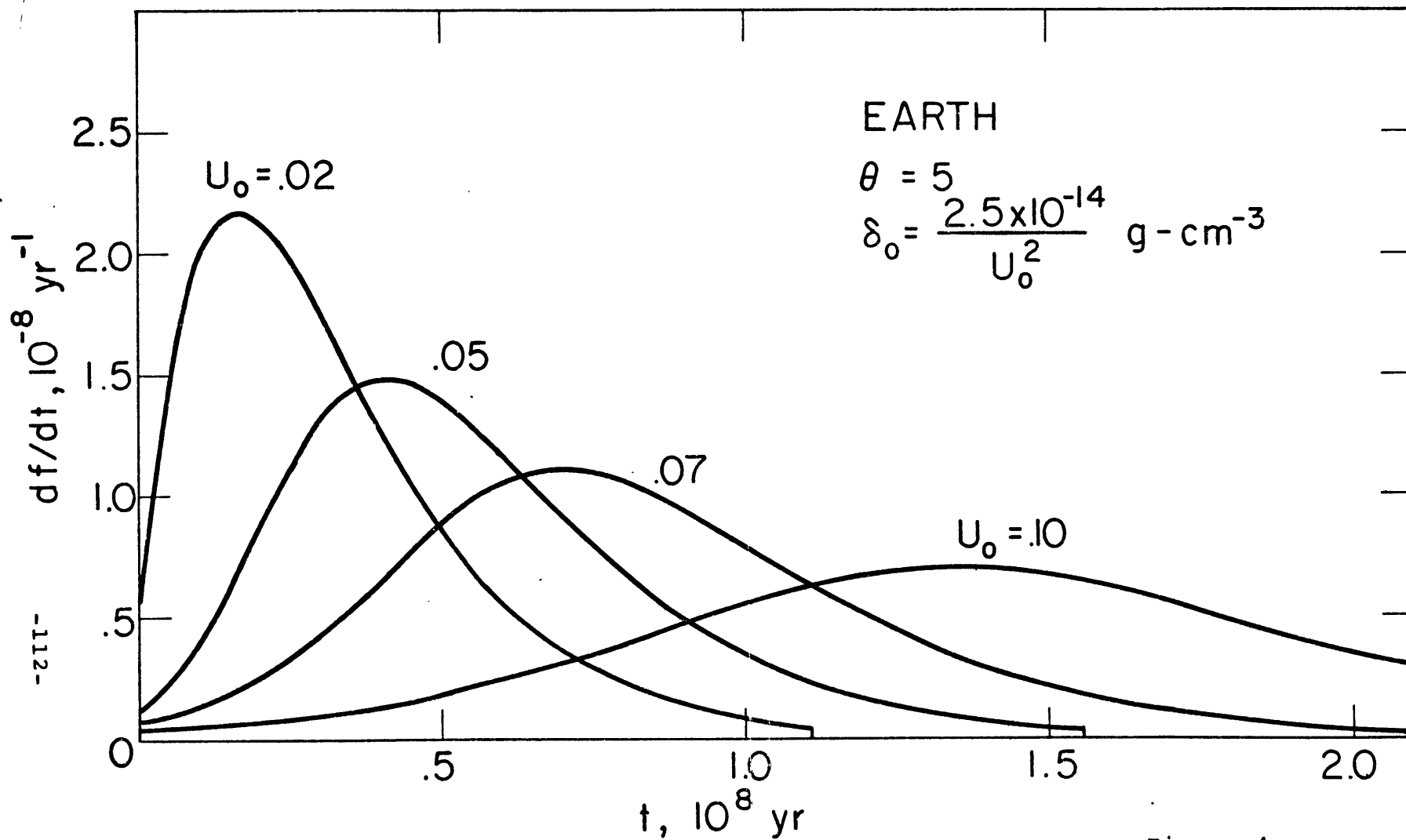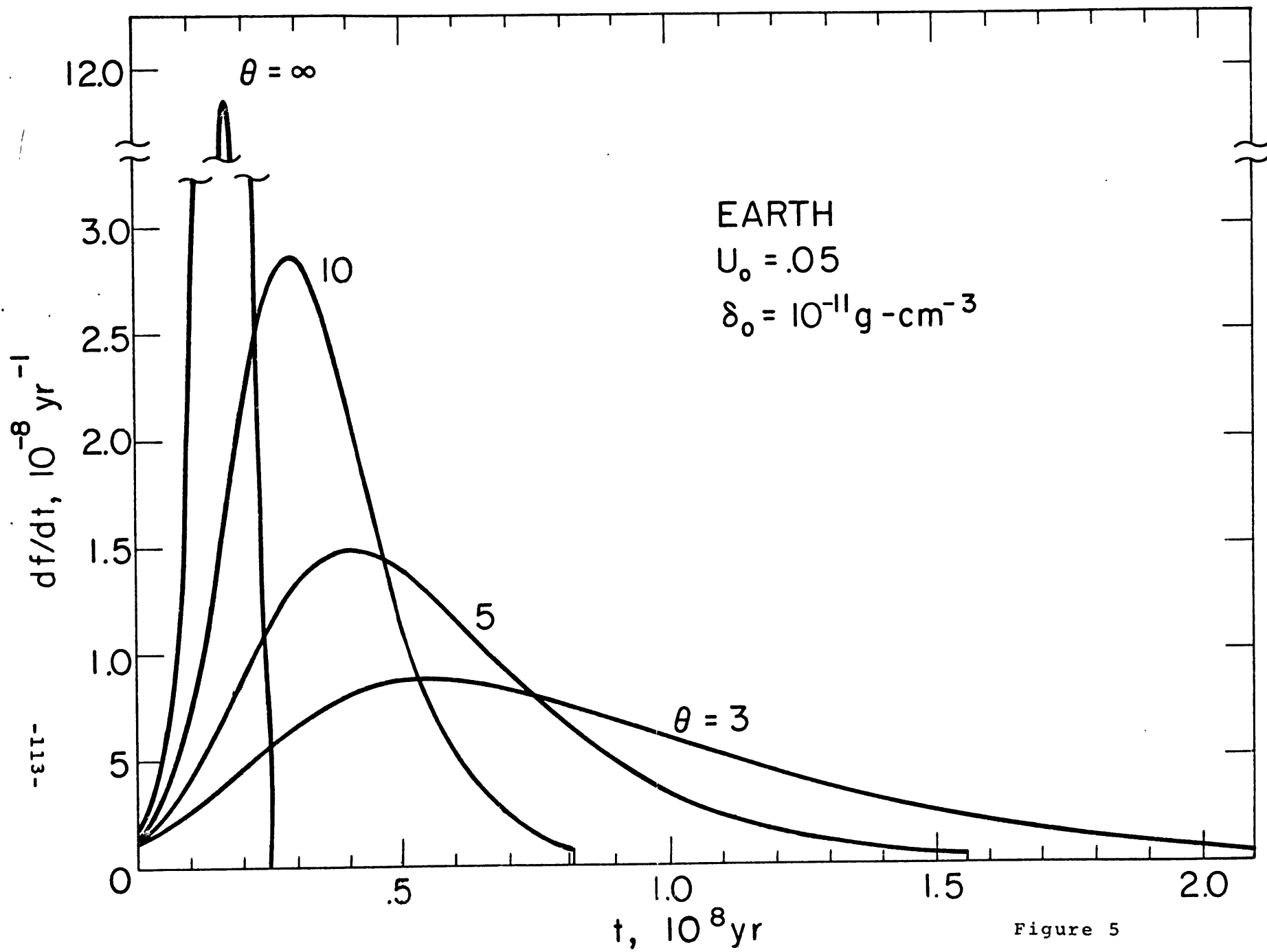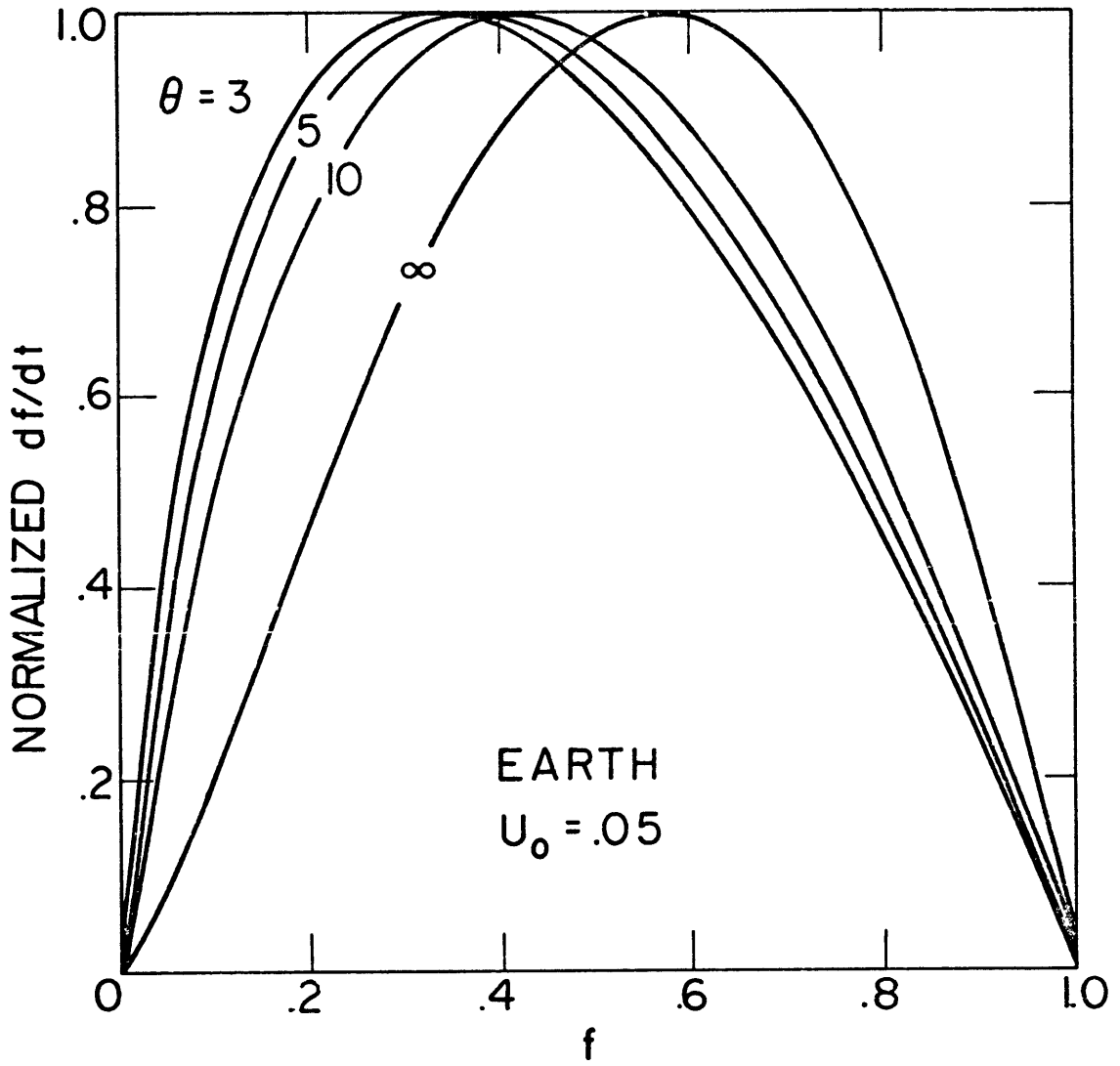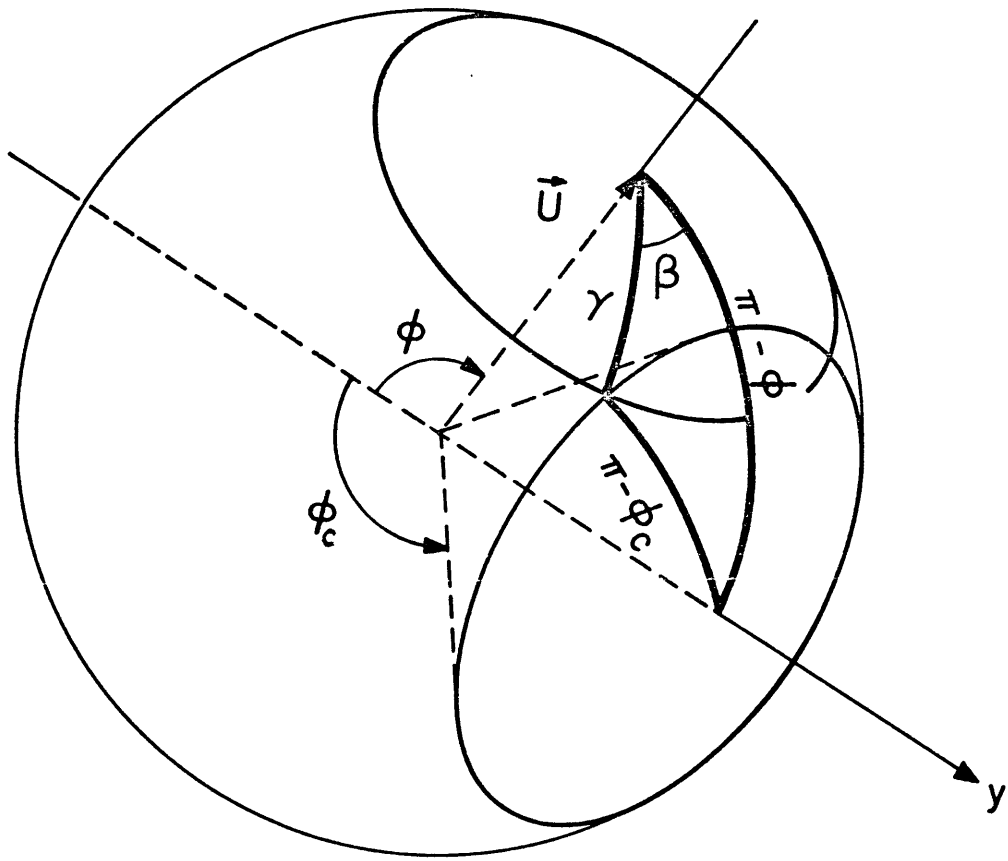
Figure 6

Figure 7

Figure 8
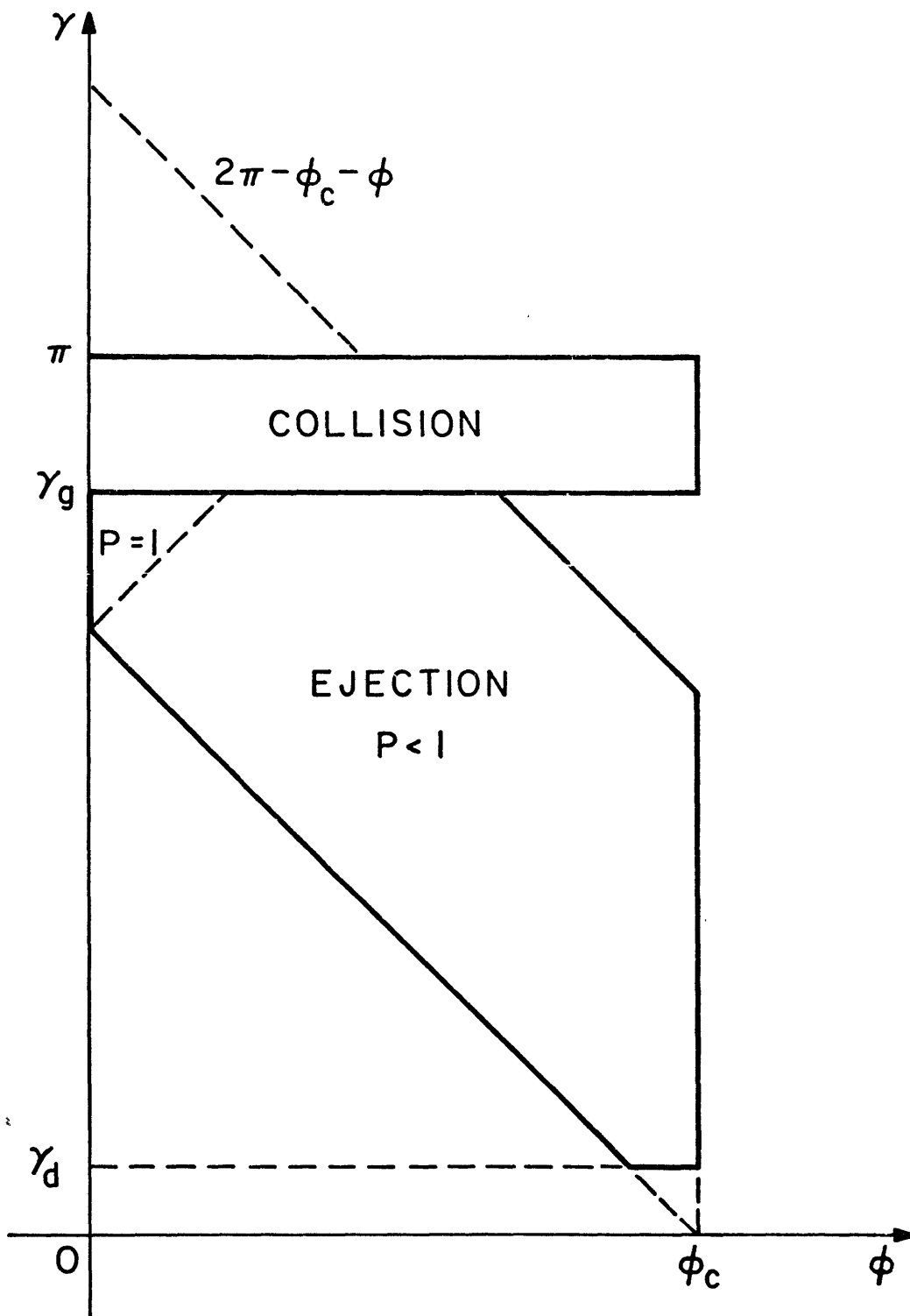
Figure 9
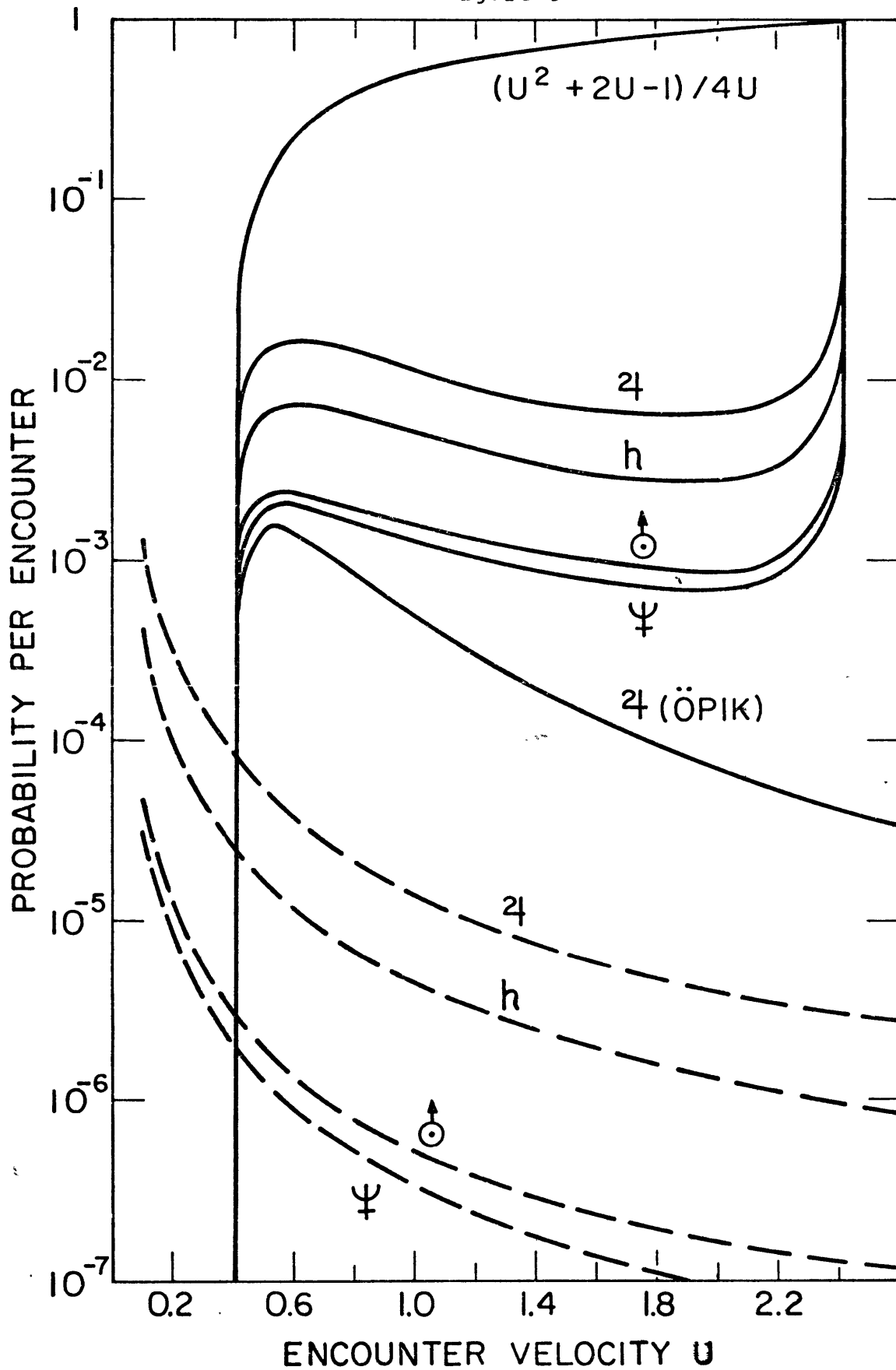
Figure 10
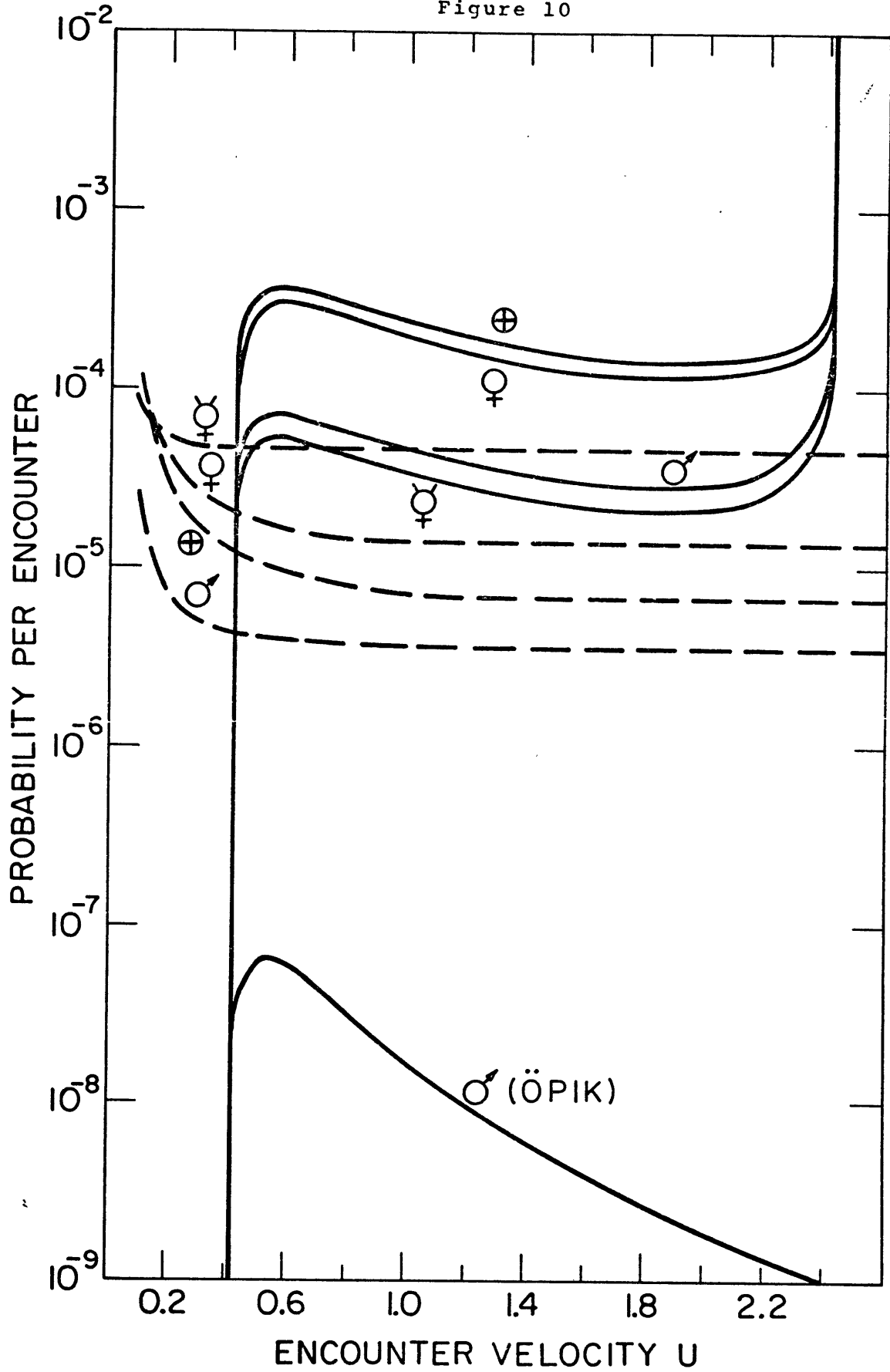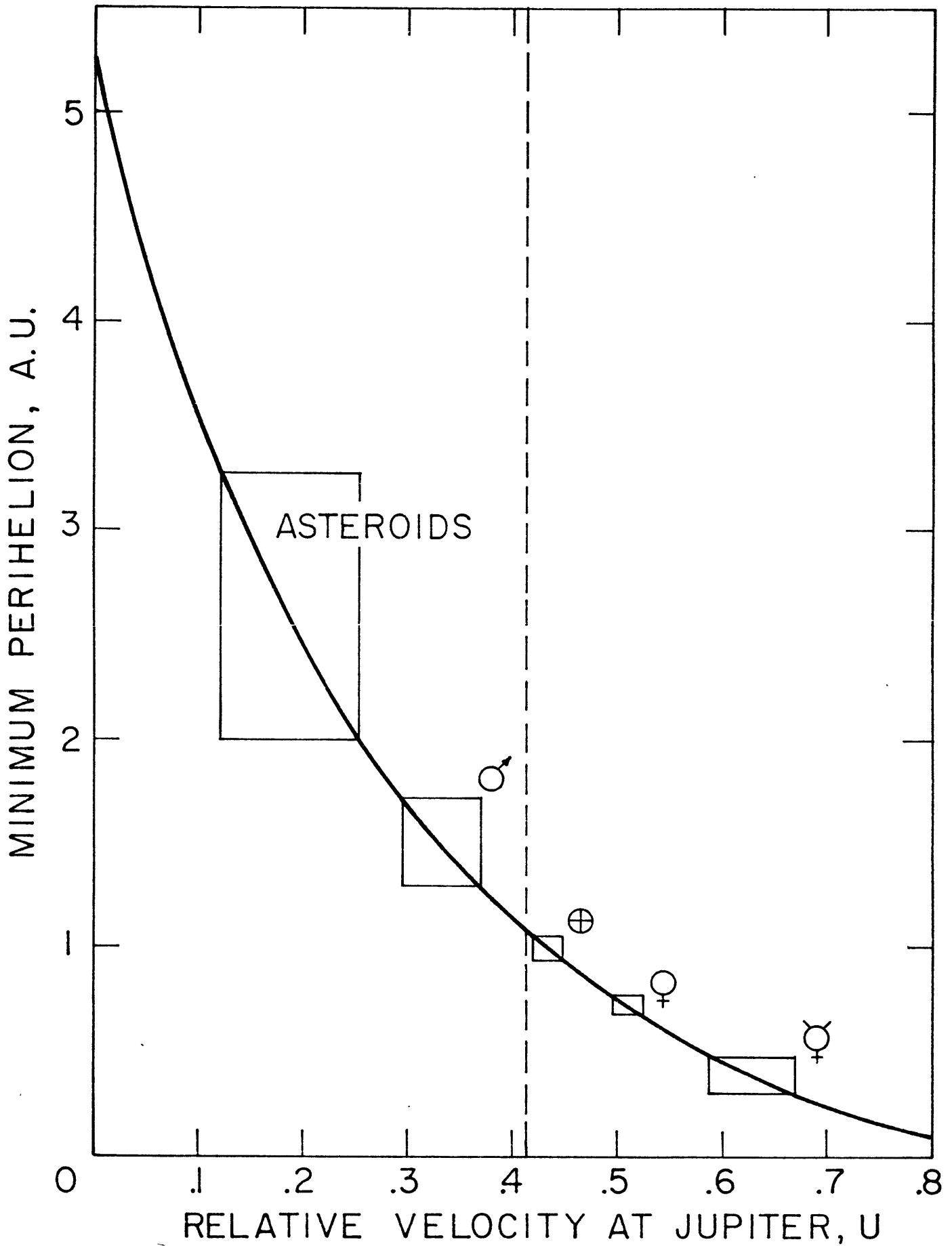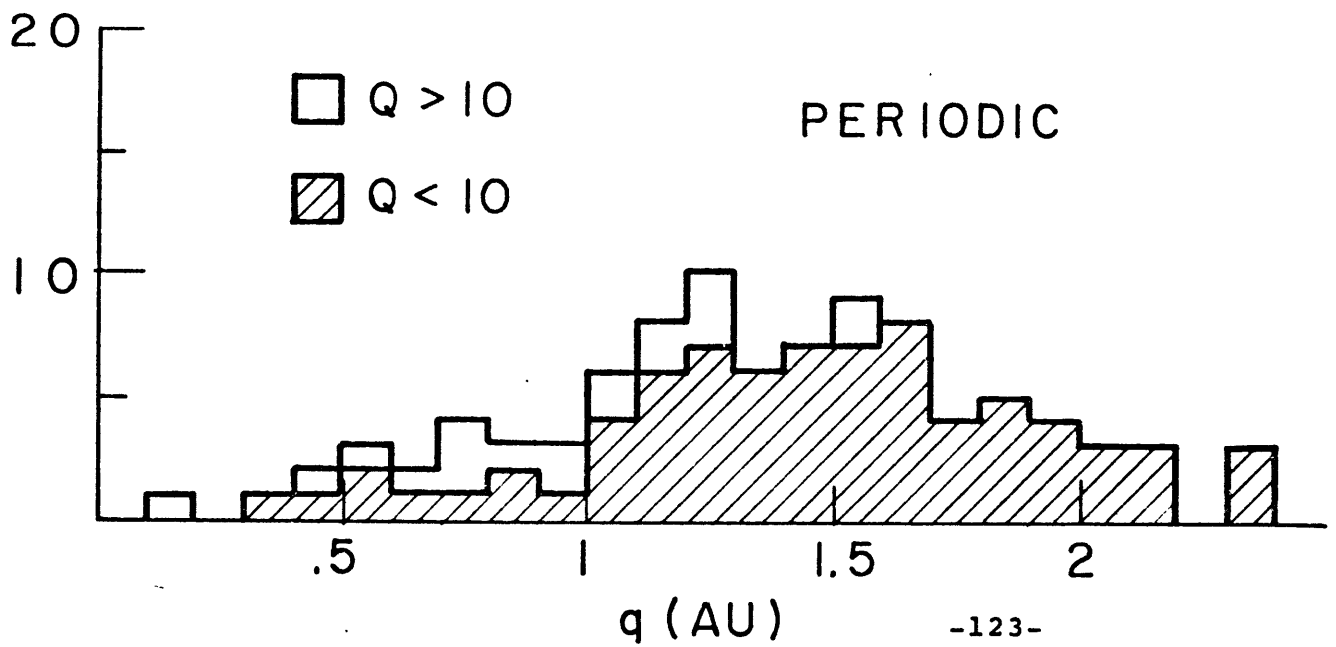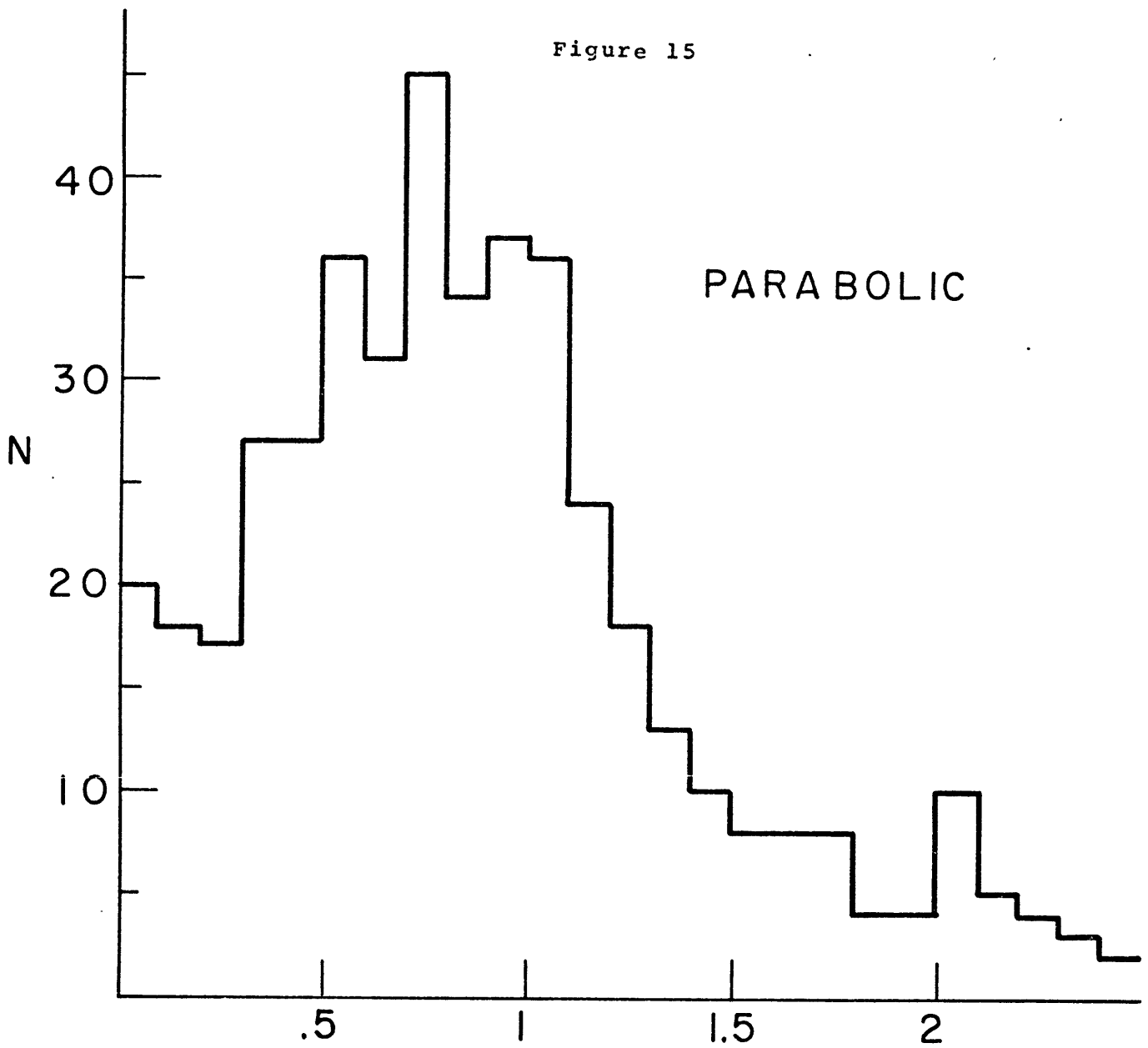
Figure 11

Figure 12
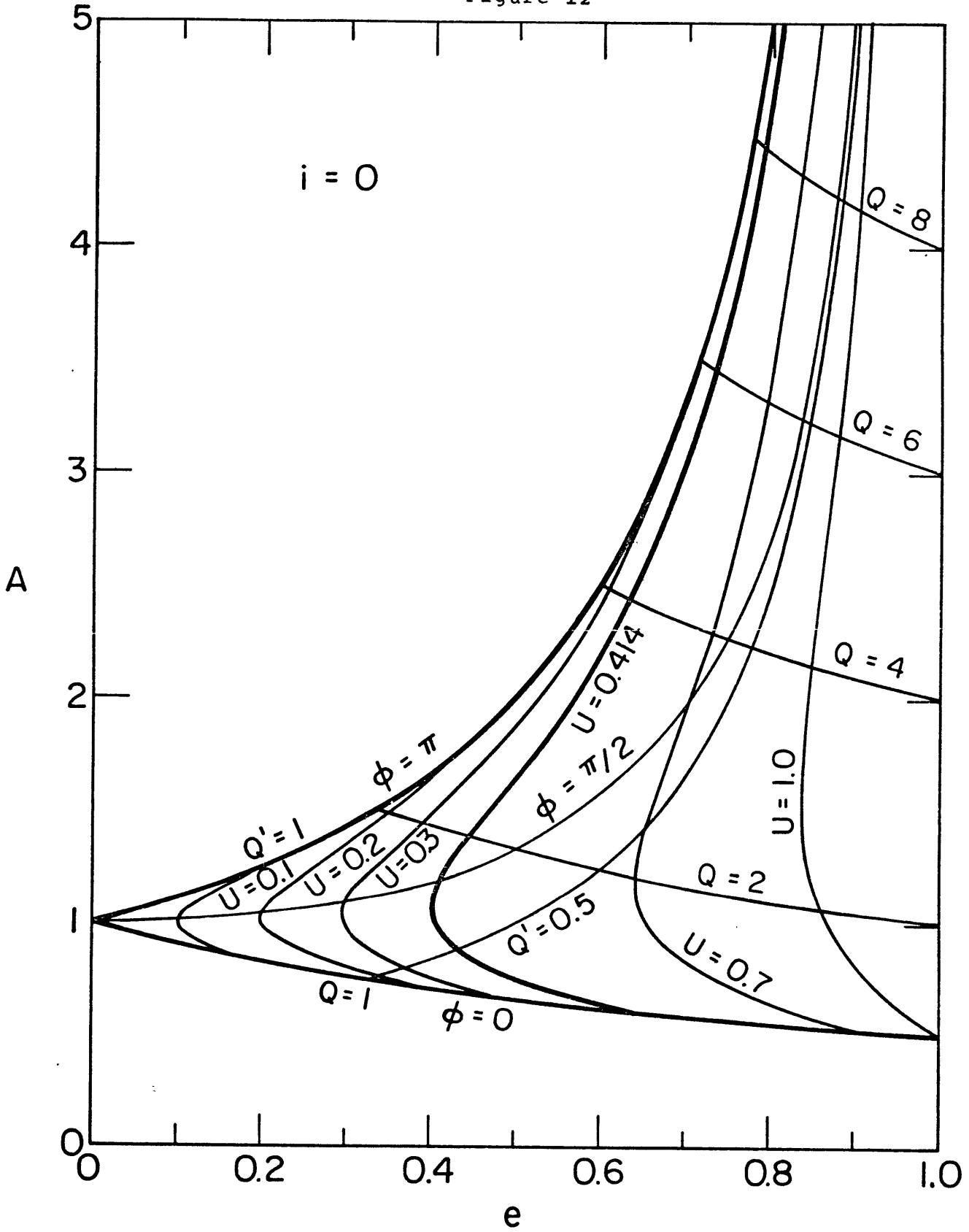
Figure 13
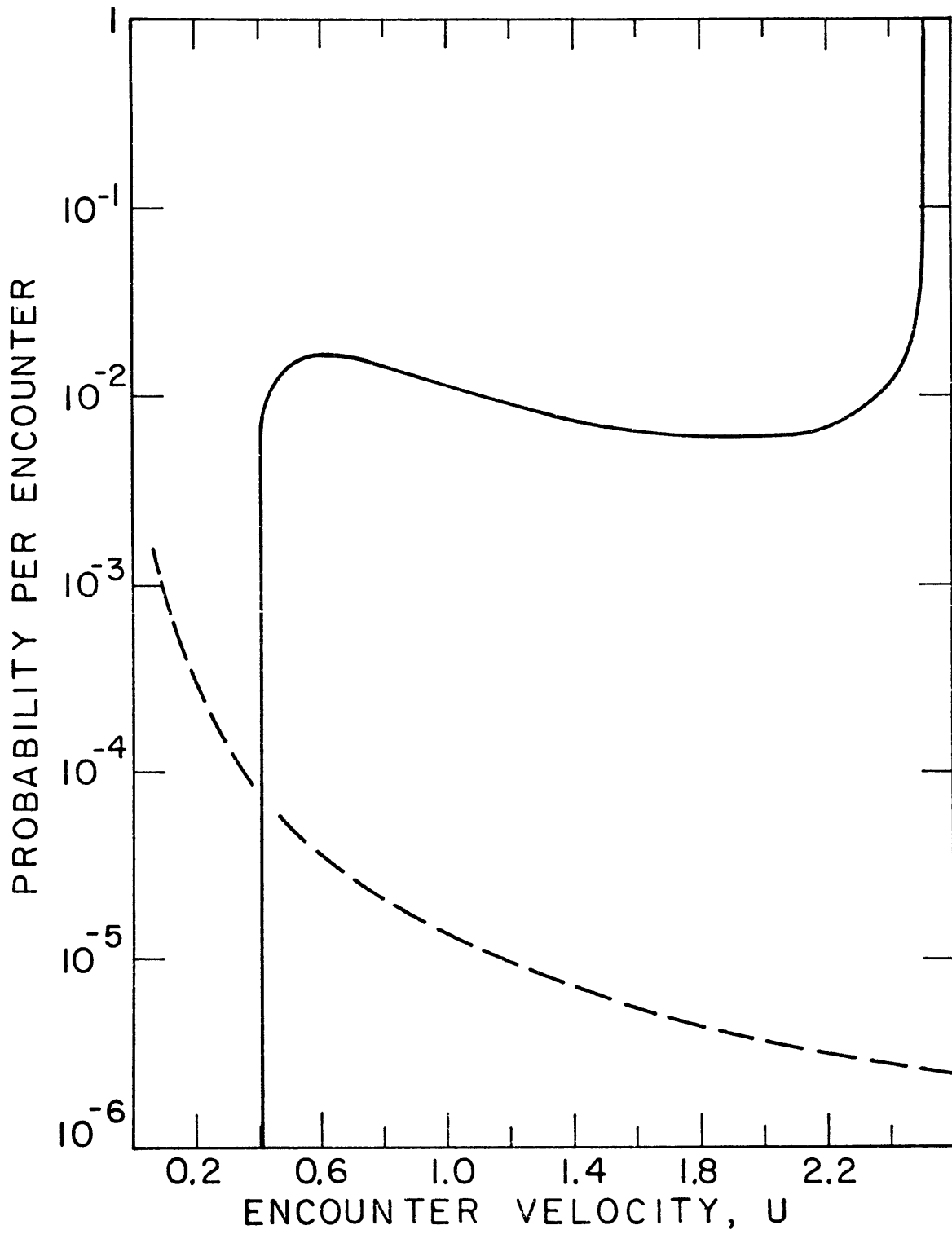
Figure 14

Figure 15

-123-

Biographical Note

Stuart John Weidenschilling was raised in the town of Lincoln Park, New Jersey. He was graduated from Boonton High School, Boonton, N.J. in 1964. He attended the Massachusetts Institute of Technology, receiving a B.S. in 1968 and a M.S. in 1969, both from the Department of Aeronautics and Astronautics. After a sabbatical in the U.S. Army, he returned to M.I.T. in 1971, and formally entered the Department of Earth and Planetary Sciences in February, 1973.

He is married to the former Susan Ann Olsen. The Weidenschillings have two cats, named Mitten and Woozle.