SOME PROBLEMS IN THE THEORY OF COMETS, I

DAVID G. KENDALL
MAGDALEN COLLEGE, OXFORD UNIVERSITY

1. Introduction

This paper is concerned with the integral equations which arise in *any* theory of comets when account is taken of the perturbation in energy state which occurs at each penetration of and passage through the planetary zone. The comet can be thought of as executing a random walk along the scale of energy states, where we define the energy state of a comet as the negative of its total energy per unit mass (so that the energy state is zero when the comet is at rest "at infinity"). It will be convenient to measure the energy state z by the reciprocal $1/a$ of the semimajor axis of the instantaneous orbit, in a.u.$^{-1}$ (so that $z = 1$ for the earth). For a parabolic comet $z = 0$, and when $z < 0$, as is sometimes the case at perihelion, it will mean that a particle having the comet's position and velocity would, if not further perturbed, leave the solar system on a hyperbolic orbit. However, the ultimate fate of a comet is not entirely determined by its instantaneous orbit at perihelion; a further passage through the planetary zone must take place before it will be even approximately correct to think of the comet as moving in a pure inverse-square field, and elaborate perturbation calculations are necessary before what I shall call the *postorbit* (the orbit after emerging from the planetary zone) can be determined. Similar calculations directed backward in time enable the *preorbit* to be found (this is the orbit before entry into the planetary zone). Some 24 preorbits have been calculated by E. Strömgren [17] and others, and recently I. V. Galibina [4] (employing a new method due to S. G. Makover [10]) has calculated 20 postorbits; practically all the comets studied in this way had hyperbolic or at least near-parabolic orbits at perihelion. Two decisive results have emerged from this work:

(1) *the observations are consistent with the statement that hyperbolic preorbits do not exist;*

(2) *some (perhaps about* 15 *per cent) of the comets having parabolic perihelion orbits leave the solar system on hyperbolic postorbits.*

This estimate is based on the fact that the energy state in the postorbit minus that in the perihelion orbit is found to be distributed about a mean value of about +0.000 50 with a standard deviation of about 0.000 50.

It is evident that the sun's family of comets is suffering continual depletion, the risks of "falling into the sun," of disintegration, and of diminution of brilliancy being further sources of loss (or apparent loss). The rate of formation of new observable comets will depend on the physical origin of these objects, and this is at present a matter of some controversy. Whatever view one may take of the origin of comets, however, essentially the same mathematical problem must be solved; namely, what mode of equilibrium, if any, is possible for a family of comets generated in some unspecified manner, suffering successive energy perturbations, and facing the constant risk of disintegration at perihelion as well as the risk of total loss if ever the barrier $z = 0$ is crossed to the negative side during the perturbations following a perihelion passage? In general terms this problem was considered by H. N. Russell in an early paper [16]; more recent and very much more detailed contributions have been made by A. J. J. van Woerkom [18] and by J. H. Oort [11]. The question has now been taken up again by J. M. Hammersley [5] and R. A. Lyttleton [9]; a significant new feature of their work is that the probabilistic character of the processes of perturbation and eventual loss is fully taken into account. The present investigation is also probabilistic in character, but the line taken here is rather different from that followed in [5]; it arose in the course of informal discussions with Dr. Hammersley, to whom I am much indebted both for stimulating my interest in the problem and for many valuable technical comments.

The plan of the present work is as follows. The general problem with which we shall be concerned is formulated in section 2, and it is solved in section 3 by assuming a double-exponential form for the perturbation distribution; the evidence for this assumption will be discussed in section 3, the relevant data having been collected and analyzed elsewhere (Kendall [6]). In section 4 the preceding results will be examined in relation to theories which have been proposed for the origin of comets. Finally, in part II of the present paper some aspects of the problem will be studied in the general case (without the double-exponential assumption).

In pursuing this work I have been very conscious of the debt which I owe to the late Professor E. A. Milne for constant encouragement and advice during my early efforts at mathematical astronomy. To his memory this paper is respectfully and affectionately dedicated.

2. Formulation of the problem

Consider a "new" comet in the initial (positive) energy state x (measured before it enters the planetary zone); if we have in mind Oort's theory [11] of the origin of comets, then a "new" comet will here mean a member of the swarm of precomets which has just been directed by stellar perturbations into an orbit with a small perihelion distance (of the order of one astronomical unit). We shall speak of the epoch of formation of a new comet, and so on, with reference to the epoch at which the object ultimately to be an observable comet first approaches the sun along an orbit which will bring it into the observable region. Let the perturbations in energy state during its first, second, third, $\cdots$ passage into and out of the planetary zone be $y_1, y_2, y_3, \cdots$, until the comet is lost or has disintegrated. We can adequately think of these perturbations as occurring at the successive perihelion epochs, so that the comet passes successively through the energy states

$$(2.1) \quad z_0 = x, \quad z_1 = x + y_1, \quad z_2 = x + y_1 + y_2, \quad z_3 = x + y_1 + y_2 + y_3, \cdots,$$

and describes the greater part of a complete orbit in each of these states except the first (in which it spends only half an orbit from aphelion to perihelion). Thus we may take the times spent in the successive states to be

$$(2.2) \qquad \frac{1}{2} V(x), \qquad V(z_1), \qquad V(z_2), \qquad V(z_3), \cdots,$$

respectively, where $V(z) = z^{-3/2} = a^{3/2}$ is the periodic time (in years) for a comet in the energy state z. It will be assumed that x and the y are statistically independent and that the y are identically distributed. The available evidence concerning the y-distribution $dG(y)$ was collected in [6]; here we shall only assume that *it is symmetrical about* $y = 0$ and that *there is a positive probability that* $y \neq 0$. These assumptions are dynamically plausible and they are consistent with the observations. They suffice for the truth of the following assertion: *with probability one there will exist a, say first, positive integer N such that $z_N \leqq 0$.* The sufficiency of the assumptions follows from the theorems of Frank Spitzer, which form the basis of part II of the present paper. In an earlier draft of this work the (almost) certainty of ultimate loss was derived from theorem 4 of [1] and the additional assumption, now redundant, that $E(|y|) < \infty$. Thus we can be sure that the comet will be perturbed out of the solar system (unless it disintegrates at some earlier perihelion passage).

Hammersley [5] and Lyttleton [9] are not especially interested in the disintegration effect, and so for them the total time of residence as an effective member of the solar system is

$$(2.3) \quad \frac{1}{2} V(x) + V(x + y_1) + V(x + y_1 + y_2) + \cdots + V(x + y_1 + y_2 + \cdots + y_{N-1});$$

the distribution of this random variable is one of the main topics discussed by them. We shall follow Oort [11] in writing k for the chance of loss by disintegration (or some similar mechanism) at or near perihelion on each circuit. Oort estimated $k = 0.019$ (from 11 disintegrations in 576 apparitions), and repeated his calculations with the smaller value $k = 0.003$. We shall merely assume that $0 \leqq k < 1$.

Without committing ourselves to any specific theory of the origin of comets we shall suppose that new comets are formed at epochs which constitute a Poisson process with parameter ρ, so that $\rho(t_2 - t_1)$ is the expected number of new comets formed during a time interval (t_1, t_2), the actual number being

Poisson-distributed and the numbers corresponding to nonoverlapping intervals being independent. To begin with we shall merely assume that this process has continued since (about) the epoch of formation of the solar system. Once we are through with the initial discussions, however, we shall treat this as an infinitely remote event; that is, we shall suppose that the system has existed long enough for statistical equilibrium to have been attained. We shall avoid making any special assumptions about the distribution $dF(x)$ of the initial energy state beyond requiring that $F(0) = 0$ (because certainly $x > 0$). Thus a complete formulation of the probabilistic model requires a sequence of negative-exponential random variables with mean value $1/\rho$, a sequence of random variables having the distribution $dF(x)$, a doubly indexed sequence of random variables having the distribution $dG(y)$, and a doubly indexed sequence of 0-1 variables having the mean value k, all these random variables being statistically independent.

We now turn our attention away from the history of an individual comet and consider instead the statistical properties of the ensemble which is the sun's whole family of (observable) comets; this is "un élément aléatoire de nature quelconque" in the sense of M. Fréchet [3], at least insofar as any specific mathematical object can be described as "quelconque." Let $t = 0$ denote the epoch of observation (that is, now) and let $M(Z)$ denote the number of comets in energy states $z > Z \geqq 0$, so that $M(0)$ is the (perhaps infinite) size of the whole family. A comet formed at the epoch $t = -\tau$ in the energy state x will contribute one unit to $M(Z)$ if and only if it is still in the system and has not been destroyed by the disintegration mechanism at $t = 0$ and is then in an energy state $z > Z$. Let the probability of this event be denoted by $P(\tau, Z|x)$, so that

$$Q(\tau, Z) = \int_0^\infty P(\tau, Z|x)\, dF(x) \tag{2.4}$$

is the probability that a comet known to have been formed at $t = -\tau$ will contribute to $M(Z)$. We can now find the probability-generating function

$$\phi(Z; w) = E[w^{M(Z)}], \qquad 0 \leqq w \leqq 1, \tag{2.5}$$

as follows. For fixed Z and w let ϕ_T denote the corresponding generating function when only comets formed in the interval $(-T, 0)$ are to be counted. Then, by considering the first comet to be formed after $t = -T$, we find that

$$\phi_T = e^{-\rho T} + \int_0^T [1 + (w - 1)Q(T - u, Z)]\phi_{T-u}\, e^{-\rho u}\, \rho\, du, \tag{2.6}$$

and on writing $T - u = v$ and differentiating with respect to T, we then easily find that

$$\phi_T = \exp\left[\rho(w - 1) \int_0^T Q(\tau, Z)\, d\tau\right]. \tag{2.7}$$

We obtain $\phi(Z; w)$ from ϕ_T by setting T equal to the effective age of the solar system, which we shall denote by (∞). Thus we conclude that $M(Z)$ is a Poisson variable whose mean value is

$$E[M(Z)] = \rho \int_0^{(\infty)} Q(\tau, Z)\, d\tau. \tag{2.8}$$

The integral at (2.8) may diverge; if so, then $M(Z) = \infty$ almost certainly.

Now the absolute statistical fluctuations in $M(Z)$ are of no significance for us because we cannot (yet) examine a number of independent replicates of the solar system; we could not even make use of the joint distribution of the $M(Z)$ for the same system but relating to different epochs, because the period of time covered by astronomical observations is negligible on this scale. We can, however, make use of the joint distribution of the variables $M(Z_1), M(Z_2), \cdots$ relating to different energy levels for the same system at the same epoch; that is, we can study the way in which the observed z-values are distributed along the z-axis; we shall call this the *z-spectrum.*

On extending the preceding argument we find that the numbers $M(Z_j, Z_j')$ of comets with energy states in the nonoverlapping half-open intervals $(Z_j, Z_j']$ are independent Poisson variables with mean values

$$\rho \int_0^{(\infty)} [Q(\tau, Z_j) - Q(\tau, Z_j')]\, d\tau, \qquad j = 1, 2, \cdots . \tag{2.9}$$

It follows that, if $Z \geqq 0$ is fixed, if the integral

$$\int_0^{(\infty)} Q(\tau, Z)\, d\tau \tag{2.10}$$

is finite, and if the (finite!) number $M(Z)$ of comets with energy states $z > Z$ is *known*, then the individual energy states of these M comets will be distributed as if they constituted a sample of size M from the distribution

$$P\{z > \zeta\} = \frac{\int_0^{(\infty)} Q(\tau, \zeta)\, d\tau}{\int_0^{(\infty)} Q(\tau, Z)\, d\tau}, \qquad Z \leqq \zeta < \infty. \tag{2.11}$$

This makes it possible to apply statistical tests when comparing a section of an empirical z-spectrum with that predicted by theory, provided that the function Q can be calculated. However, the function Q involves the distribution $dF(x)$ of the initial energy state, and the latter distribution is closely connected with one's choice of theory of the origin of comets. We shall therefore eliminate $dF(x)$ from the analysis as far as possible by supposing that x is fixed and is the same for all comets; the function $Q(\tau, Z)$ will then be replaced by $P(\tau, Z|x)$. The question of how one should average over the x-values, or, in case a single value of x is appropriate, of what single value of x to choose, will be left open for the moment. We therefore turn to the consideration of the quantities

$$R(Z|x) = \int_0^{\infty} P(\tau, Z|x)\, d\tau, \tag{2.12}$$

noting that when $R(Z|x)$ is finite then

$$P\{z > \zeta | z > Z; x\} = \frac{R(\zeta|x)}{R(Z|x)}, \qquad 0 \leqq Z \leqq \zeta < \infty. \tag{2.13}$$

We have here replaced (∞) by ∞, indicating that from now on we shall treat the epoch of formation of the solar system as an infinitely remote event. The extent to which this is a justifiable assumption can be assessed by comparing the age of the solar system with the distribution of residence times for comets found by Hammersley in [5].

We have used N for the ordinal number of the perihelion passage at which the comet (on emergence from the planetary zone) would first be thrown into a parabolic or hyperbolic postorbit, in the absence of the disintegration mechanism. We shall now use N^* to denote either this integer N or the ordinal number of the perihelion passage at which disintegration would occur, whichever is the smaller; thus the last complete circuit, or half circuit, if $N^* = 1$, will be described in the energy state $z_{N^*-1} = x + y_1 + \cdots + y_{N^*-1}$, and the total time spent as a bound member of the system in energy states $z > Z$ will be

$$S = \frac{1}{2} V(x) + \sum_{\substack{1 \leq j < N^* \\ z_j > Z}} V(z_j), \tag{2.14}$$

where the first term $V(x)/2$ is to be omitted if $x \leqq Z$.

We can think of S as the area below the "curve" constructed as follows. Consecutive segments of length $V(x)/2$, $V(z_1)$, $V(z_2)$, $\cdots$, $V(z_{N^*-1})$ are laid out along a horizontal axis, starting at zero, and the ordinate is required to be $+1$ if $z_j > Z$ (respectively $x > Z$) and zero if $z_j \leqq Z$ (respectively $x \leqq Z$) in the interior of the segment of length $V(z_j)$ (respectively $V(x)/2$). Everywhere else the ordinate is to be equal to zero. The function so constructed is nonnegative, and is jointly measurable with respect to its two arguments, which are the real variable labeling the horizontal axis (carrying Borel sets and Lebesgue measure) and the probability parameter. Thus Fubini's theorem applies and gives

$$E[S] = R(Z|x). \tag{2.15}$$

We shall obtain an explicit formula for $R(Z|x)$ by calculating $E[S]$ in another way. First let $A(z|x)$ be the number of *complete* circuits (*excluding the initial half circuit*) performed in energy states $> z$, and let $B(z|x)$ be the number of complete circuits performed in energy states $\leqq z$, so that $A(z|x) + B(z|x) = N^* - 1$. Then

$$S = \begin{cases} \dfrac{1}{2} V(x)H(x - Z) - \displaystyle\int_Z^\infty V(z)d_zA(z|x), \\ \dfrac{1}{2} V(x)H(x - Z) + \displaystyle\int_{Z+0}^\infty Vzd_zB(z|x), \end{cases} \tag{2.16}$$

because $B(\cdot|x)$ is continuous to the right. Here $H(u) = 1$ if u is positive, and otherwise $H(u)$ is zero. Let $C(z|x)$ denote the *expected* number of *complete* circuits performed at an energy state *less than or equal to* z, so that $C(z|x) = E[B(z|x)]$. If $J = \int_{z_1}^{z_2} V(z)d_zB(z|x)$, where $0 < z_1 < z_2 < \infty$, then

$$J = [V(z)B(z|x)]_{z_1}^{z_2} - \int_{z_1}^{z_2} B(z|x)\,dV(z), \tag{2.17}$$

and so

$$E[J] = [V(z)C(z|x)]_{z_1}^{z_2} - \int_{z_1}^{z_2} C(z|x)\,dV(z) = \int_{z_1}^{z_2} V(z)d_zC(z|x), \tag{2.18}$$

provided that $C(z|x)$ is finite for $z = z_2$. Alternatively, the taking of expectations can be justified by an appeal to H. Robbins [15] or W. Feller [2]. Thus, letting $z_1 \downarrow Z$ and $z_2 \uparrow \infty$, we find that

$$R(Z|x) = \frac{1}{2} V(x)H(x - Z) + \int_{Z+0}^{\infty} V(z)d_zC(z|x), \qquad Z \geqq 0, \tag{2.19}$$

provided that $C(z|x)$ *is finite for all finite* z. This last formula, together with (2.13), shows that the statistical discussion of the z-spectrum hinges on the determination of the function $C(z|x)$ for all positive z and x. Plainly $C(0|x) = 0$.

We shall now set up an integral equation which is satisfied by $C(x) = C(z|x)$ for fixed positive z, and from which in certain circumstances it is possible to determine $C(x)$ when $x > 0$. We start with the explicit formula

$$C(x) = \sum_{n=1}^{\infty} (1 - k)^n \iint \cdots \int_{E_n} dG(y_1) \cdots dG(y_n), \tag{2.20}$$

where the region E_n of integration is that determined by the inequalities

$$\begin{aligned} &0 < x + y_1 + y_2 + \cdots + y_s, \qquad s = 1, 2, \cdots, n - 1, \\ &0 < x + y_1 + y_2 + \cdots + y_n \leqq z. \end{aligned} \tag{2.21}$$

It is now clear that $C(x)$ must be a solution to the integral equation

$$C(x) = (1 - k) \int_{0<x+y_1\leqq z} dG(y_1) + (1 - k) \int_{0<x+y_1} C(x + y_1)\,dG(y_1), \qquad x > 0. \tag{2.22}$$

Let $C_r(x)$ denote the sum of the first r terms of the series (2.20), and let $K(x)$ be *any* nonnegative solution to (2.22). Then iteration of (2.22) gives

$$K(x) = C_r(x) + (1 - k)^r \iint \cdots \int_{F_r} K(x + y_1 + \cdots + y_r)\,dG(y_1) \cdots dG(y_r), \tag{2.23}$$

where F_r denotes the region of integration determined by the inequalities $0 < x + y_1 + y_2 + \cdots + y_s$, with $s = 1, 2, \cdots, r$. If now $K(x)$ is finite and bounded for all $x > 0$, then the last term on the right side of (2.23) will tend to zero when r tends to infinity, because we know that with probability one the barrier at $z = 0$ will ultimately be crossed, and then it will follow that $K(x) = C(x)$. Thus *if the integral equation* (2.22) *possesses a finite nonnegative measurable solution* $C(x)$ *which is bounded for* $x > 0$, *then*

$$C(x) = C(z|x), \tag{2.24}$$

and $R(Z|x)$ *can be calculated from* (2.19) *for all* $x > 0$ *and* $Z \geqq 0$.

3. Double-exponential perturbations

In physics and astronomy unknown symmetrical distributions with finite variances are commonly taken to be Gaussian, and this course has been followed in the present problem by Oort [11] and, in a portion of their work, by Hammersley and Lyttleton. We have so far only assumed that the perturbation distribution $dG(y)$ is symmetrical, but we are of course quite ready to grant also the finiteness of the second moment. On this occasion there is even a sound theoretical reason for the Gaussian assumption; we are concerned essentially with a problem about the cumulative sums of independent random variables having the distribution $dG(y)$, and if the second moment is finite then all but the first few such sums will have nearly Gaussian distributions in virtue of the central limit theorem. It is easy to overlook the fact that precisely the same argument can be employed to justify the choice of any other distribution which lies in what is called the domain of attraction of the Gaussian law (that is, which is such that its convolutions ultimately approach the Gaussian law). Thus if we are dealing with cumulative sums of identical independent random variables having a finite variance, but of otherwise unknown distribution, and if some other distribution within the Gaussian domain of attraction is analytically more convenient, there is every reason to choose it, as a working approximation, rather than the Gaussian law itself. This is the case in the present problem. We shall find that there are great analytical advantages in approximating to $dG(y)$ by the *double-exponential* law

$$\frac{1}{2} e^{-|y|/b} \frac{dy}{b}, \qquad -\infty < y < \infty. \tag{3.1}$$

The empirical evidence on the form of $dG(y)$ is collected in my paper [6]; of the various estimates there obtained the one most free from possible selection errors is that based on Galibina's computations of 20 postorbits. This empirical distribution is shown again here in figure 1 as a convolved histogram, the polygonal line in the figure, together with the density curve for the double-exponential law having the same (zero) mean and variance. A Gaussian curve of the same mean and variance is also shown, but does not give a markedly better fit; the differences would be small after a few convolutions. In this section of the paper we shall adopt the double-exponential assumption and explore its consequences.

The fundamental integral equation (2.22) now takes the form

(3.2)

$$C(x) = (1-k)\int_0^z \gamma(w-x)\,dw + (1-k)\int_0^\infty C(w)\gamma(w-x)\,dw, \qquad x > 0,$$

for any fixed $z > 0$, where

$$\gamma(u) = \frac{1}{2b} e^{-|u|/b}, \qquad -\infty < u < \infty. \tag{3.3}$$

It is slightly easier to handle (3.2) if we first put $D(x) = Z(x) + C(x)$, where

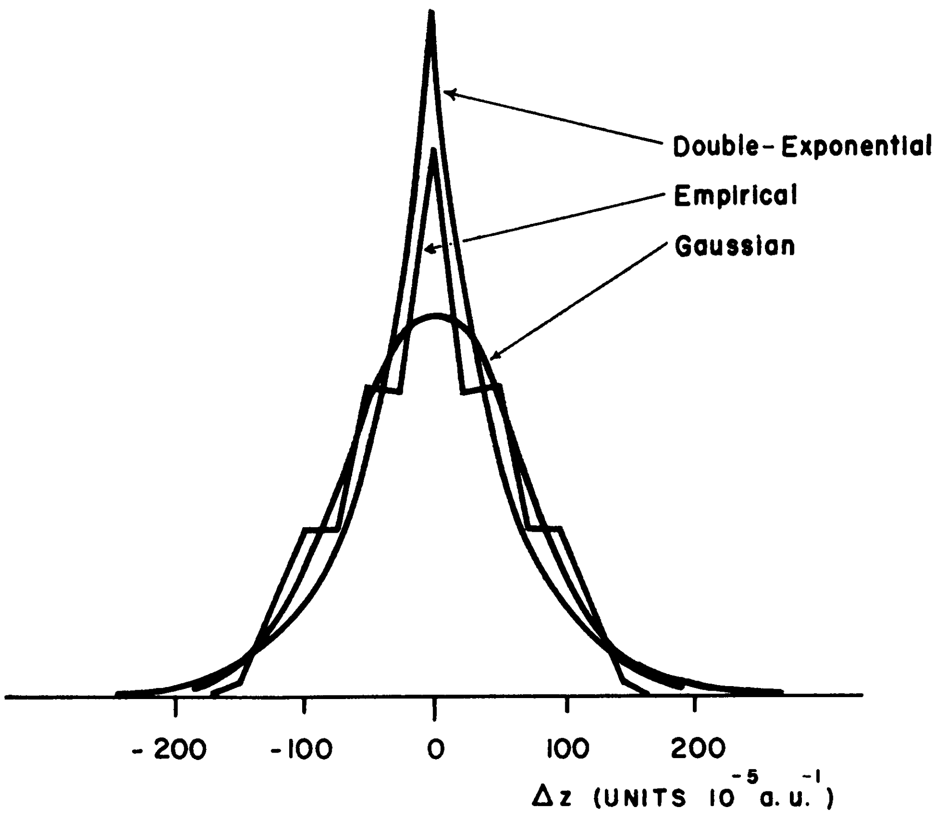

FIGURE 1

Comparison of Gaussian, double-exponential, and empirical distribution of Δz.

$Z(x) = 1$ for $0 < x \leqq z$, and $Z(x) = 0$ for $z < x$. The function $D(x)$ then satisfies the equation

(3.4)

$$D(x) = Z(x) + \frac{1-k}{2b} e^{-x/b} \int_0^x D(w) e^{w/b}\, dw + \frac{1-k}{2b} e^{x/b} \int_x^\infty D(w) e^{-w/b}\, dw,$$

and we try to find a solution $D(\cdot)$ which is bounded for $x > 0$. Such a function will have to be continuous save for a jump of amount -1 at $x = z$, and it will have to satisfy

$$D(0+) = 1 + \frac{1-k}{2b} \int_0^\infty D(w) e^{-w/b}\, dw, \tag{3.5}$$

the integral being finite. It will be differentiable except when $x = z$, and for all other positive values of x we shall have

$$D'(x) = -\frac{1-k}{2b^2} e^{-x/b} \int_0^x D(w) e^{w/b}\, dw + \frac{1-k}{2b^2} e^{x/b} \int_x^\infty D(w) e^{-w/b}\, dw, \tag{3.6}$$

from which we see that $D'(\cdot)$ must be continuous (that is, it can be defined at $x = z$ so as to be continuous) throughout $(0, \infty)$, that

$$D'(0+) = \frac{1-k}{2b^2} \int_0^\infty D(w) e^{-w/b}\, dw, \tag{3.7}$$

and that

$$D(0+) = 1 + bD'(0+). \tag{3.8}$$

When $0 < x \neq z$, then $D''(x)$ exists and is given by

$$D''(x) = \frac{D(x) - Z(x)}{b^2} - \frac{1-k}{b^2} D(x), \tag{3.9}$$

so that $D(\cdot)$ must satisfy the differential equation

$$D''(x) - kb^{-2}D(x) = -b^{-2}Z(x), \qquad 0 < x \neq z. \tag{3.10}$$

On integrating (3.10) we obtain

$$D(x) = \begin{cases} k^{-1} + Ae^{\lambda x} + Be^{-\lambda x}, & 0 < x \leqq z, \\ Ce^{-\lambda x}, & z < x; \end{cases} \tag{3.11}$$

here λ denotes the positive square root of k/b^2, and the term in $\exp(\lambda x)$ is omitted when $x > z$ because we want $D(x)$ to be bounded. When $k = 0$, the last two equations have to be replaced by

$$D(x) = \begin{cases} A + Bx - \dfrac{x^2}{2b^2}, & 0 < x \leqq z, \\ C, & z < x. \end{cases} \tag{3.12}$$

We now have to find the constants A, B, and C in each case. This is easily done by using the continuity of $D'(\cdot)$, the known jump of $D(\cdot)$ at $x = z$, and the relation between $D(0+)$ and $D'(0+)$. The final results are as follows.

$$C(z|x) = \begin{cases} \dfrac{1 - \sqrt{k}}{k} \{1 - e^{-\lambda x} - e^{-\lambda z} \sinh \lambda x \\ \qquad\qquad + k^{1/2}[1 - e^{-\lambda z} \cosh \lambda x]\}, & 0 < x \leqq z, \\ \dfrac{1 - \sqrt{k}}{k} \{\cosh \lambda z - 1 + k^{1/2} \sinh \lambda z\} e^{-\lambda x}, & z < x. \end{cases} \tag{3.13}$$

When there is no disintegration ($k = 0$) a separate analysis is needed, and we then find that

$$C(z|x) = \begin{cases} \dfrac{z}{b} + \dfrac{zx}{b^2} - \dfrac{x^2}{2b^2}, & 0 < x \leqq z, \\ \dfrac{z}{b} + \dfrac{z^2}{2b^2}, & z < x. \end{cases} \tag{3.14}$$

That the functions defined at (3.13) and (3.14) are identical with $C(z|x)$ is a consequence of the facts that they are nonnegative and bounded, and can be shown, by substitution, to satisfy the integral equation (3.2); the identification with $C(z|x)$ then follows immediately in virtue of the result established at the end of section 2. In (3.13) and (3.14) we have all that is needed for a study of the z-spectrum when the perturbation distribution $dG(y)$ has the double-exponential form. It will be recalled that $\lambda = b^{-1}\sqrt{k}$; the constant b can be estimated by noting that the root-mean-square perturbation $\sigma = b\sqrt{2}$. Notice that in each case $C(\cdot|x)$ is continuous at $z = x$.

The technique which we have employed for the calculation of $C(z|x)$ is a very powerful one, and it can be used to obtain the expectations or more generally the distributions of many random variables connected with the motion of a single comet, when the perturbation distribution has the double-exponential form. To illustrate this we shall now find the joint probability-generating function for the random variables $A(z|x)$ and $B(z|x)$ which were introduced in section 2; it will be recalled that the function $C(z|x)$ with which we have just now been concerned is the mathematical expectation of $B(z|x)$, and that $A(z|x) + B(z|x)$ is the total number of *complete* circuits performed by the comet before it is lost by one mechanism or another. In order to simplify the rather complicated formulas we shall suppose for the remainder of this section of the paper that $k = 0$; the results now to be obtained should be useful in graduating the numerical (Monte Carlo) studies by Hammersley and Lyttleton, and can be so employed because these writers did not include a disintegration effect in that part of their work. The reader who wishes to try out the method for himself may like to obtain the analogous formulas which hold when $k > 0$.

Let u and v be fixed real numbers in the half-closed interval $[0, 1)$, let z be a fixed positive number, and let us consider

$$\theta(x) = \theta(u, v; z|x) = E[u^{A(z|x)}v^{B(z|x)}] \tag{3.15}$$

as a function of the positive real variable x. It is easily seen that it must satisfy the integral equation

$$\theta(x) = \frac{1}{2}e^{-x/b} + \frac{1}{2}e^{-x/b}\int_0^x \pi(w)\theta(w)\,e^{w/b}\frac{dw}{b} + \frac{1}{2}e^{x/b}\int_x^\infty \pi(w)\theta(w)\,e^{-w/b}\frac{dw}{b}, \tag{3.16}$$

where $\pi(w) = v$ when $0 < w \leqq z$, and $\pi(w) = u$ when $z < w$. Now $\theta(\cdot)$ is a bounded nonnegative measurable function, and from the fact that it satisfies the integral equation we see that it is continuous for $x > 0$ and differentiable for $x \neq z$. We find that

$$2\theta(0+) = 1 + \int_0^\infty \pi(w)\theta(w)e^{-w/b}\frac{dw}{b}, \tag{3.17}$$

where the integral is convergent. Let us write P, Q, and R for the three terms which occur on the right side of the integral equation; then when $x \neq z$ we shall have

$$\theta'(x) = \frac{-P - Q + R}{b}, \tag{3.18}$$

so that $\theta'(z)$ can be defined in such a way that $\theta'(\cdot)$ is continuous and

$$2b\theta'(0+) = -1 + \int_0^\infty \pi(w)\theta(w)e^{-w/b}\frac{dw}{b}; \tag{3.19}$$

thus

$$\theta(0+) = 1 + b\theta'(0+). \tag{3.20}$$

When $x \neq z$ we can differentiate again, and this time we find that

$$\theta''(x) = \frac{+P + Q + R}{b^2} - \frac{\pi(x)\theta(x)}{b^2}; \tag{3.21}$$

accordingly $\theta(\cdot)$ satisfies the differential equations

$$\theta''(x) = \begin{cases} \dfrac{1 - v}{b^2}\,\theta(x), & 0 < x < z, \\[2ex] \dfrac{1 - u}{b^2}\,\theta(x), & z < x. \end{cases} \tag{3.22}$$

All we have to do now is to solve these equations and use the continuity and boundedness of $\theta(\cdot)$ and $\theta'(\cdot)$ and the relation between $\theta(0+)$ and $\theta'(0+)$ to determine the four constants. We find that

$$\theta(u, v; z|x) = \begin{cases} \dfrac{2\lambda \sinh \mu(z - x) + 2\mu \cosh \mu(z - x)}{(1 + \mu b)(\lambda + \mu)e^{\mu z} + (1 - \mu b)(\mu - \lambda)e^{-\mu z}}, & 0 < x \leqq z, \\[2ex] \dfrac{2\mu\, e^{-\lambda(x - z)}}{(1 + \mu b)(\lambda + \mu)e^{\mu z} + (1 - \mu b)(\mu - \lambda)e^{-\mu z}}, & z \leqq x. \end{cases} \tag{3.23}$$

Here $\lambda b = +(1 - u)^{1/2}$ and $\mu b = +(1 - v)^{1/2}$; this λ has no connection with the λ in (3.13).

From the formulas (3.23) a large number of particular results of some interest can be obtained. Thus if we want the distribution of the number $A(z|x)$ of complete circuits at energy levels above z, we must let v approach one from the left; we get

$$E[u^{A(z|x)}] = \begin{cases} \dfrac{b + (z - x)(1 - u)^{1/2}}{b + (z + b)(1 - u)^{1/2}}, & 0 < x \leqq z, \\ \dfrac{e^{-(x-z)(1-u)^{1/2}/b}}{1 + \dfrac{(z + b)}{b}(1 - u)^{1/2}}, & z \leqq x. \end{cases} \tag{3.24}$$

If in (3.24) we put $u = 0$, we obtain

$$P\{A(z|x) = 0\} = \begin{cases} \dfrac{b + z - x}{2b + z}, & 0 < x \leqq z, \\ \dfrac{be^{-(x-z)/b}}{2b + z}, & z \leqq x. \end{cases} \tag{3.25}$$

But the probability given at (3.25) is just the chance that the highest energy state $Z = z_{\max}$ in which complete circuits are made will not exceed z, and so the distribution of the random variable $z_{\max}$ has an atom of mass$(1/2)\exp(-x/b)$ at zero (this is just the chance that *no* complete circuits are made, since we conventionally define $z_{\max} = 0$ in this case) and the remaining probability mass is distributed as follows:

$$\begin{aligned} &\frac{(Z + b)e^{-(x-Z)/b}}{(2b + Z)^2}\, dZ, && 0 < Z < x, \\ &\frac{x + b}{(2b + Z)^2}\, dZ, && x < Z < \infty. \end{aligned} \tag{3.26}$$

Normally x will be much smaller than b (at least if we follow Oort), and then we can easily obtain the upper five and one per cent points in the $z_{\max}$-distribution; they are

$$\begin{aligned} Z_{95} &= 18b + \ \ 20x, \\ Z_{99} &= 98b + 100x. \end{aligned} \tag{3.27}$$

These formulas are of particular interest in connection with the Monte Carlo runs of Hammersley and Lyttleton; they may also, when amended to allow $k > 0$, be of value in deciding how frequently a comet will be lost (from the long- to the short-period family) by entering too small an orbit.

If we keep v fixed and let u approach one from the left, we obtain in exactly the same way the probability-generating function for $B(z|x)$. We already know the expectation $C(z|x)$ for this random variable, and now we can supplement this by calculating, for example, the variance.

4. The z-spectrum, theoretical and empirical

According to the formulas of section 2 the expected number of comets with energy states greater than z can be calculated by multiplying $R(z|x)$ by the constant ρ, where x is the energy state of a comet during its initial approach to perihelion. If we knew the distribution of x, then $R(z|x)$ would have to be weighted accordingly. But we do not know this; we therefore retain x as a variable and leave the question of the correct weighting open for the time being.

We want to compute a theoretical z-spectrum which can be compared with the empirical distribution of *z as measured in the preorbit,* and we shall have to be careful to make an ascertainment correction to allow for the fact that we can observe, not *all* the comets in a given energy band, but only those which happen to have passed through perihelion during the comparatively short period of time T covered by astronomical observations.

On examining formulas (2.19), (3.13), and (3.14) we see that $\rho R(z|x)$ corresponds to a concentration of $\rho V(x)/2$ "new" comets in the energy state x plus a continuous distribution of "old" comets with density

$$\rho V(z) \frac{\partial}{\partial z} C(z|x)\, dz, \qquad 0 < z < \infty. \tag{4.1}$$

It will be convenient to think of these as being the numbers of comets in the various energy ranges at the *commencement* of the period T. The chance that one of the "new" comets will pass through perihelion during the period T is equal to $2T/V(x)$. Consider those comets at the commencement of the period T which are in the band $(z, z + dz)$ of energy states. We shall only be concerned with the long-period end of the z-spectrum, so that $T \ll V(z)$. These comets will approach perihelion (in preorbits having energy states in the stated band) during a period starting at this epoch and of length $V(z)$, and because of the stationarity of the situation a fraction $T/V(z)$ of them will come to perihelion within the period of length T covered by the observations. Thus the ascertainment factor for the continuous distribution of "old" comets is equal to $T/V(z)$, so that the theoretical z-spectrum (for z measured in the preorbit) consists of

(i) a concentration of ρT "new" comets in the energy state x, and

(ii) a continuous distribution of "old" comets with density

$$\rho T \frac{\partial}{\partial z} C(z|x)\, dz, \qquad 0 < z < \infty. \tag{4.2}$$

Oort [11] has suggested that "new" comets which are making their first close passage by the sun may be intrinsically more luminous than "old" comets which have already had this experience and have used up some of the relevant constituents. If this is so then, as Oort has pointed out, the concentration of "new" comets would occur with augmented frequency in the empirical z-spectrum. On the other hand Lyttleton [8] refers to the possibility that "the chance of discovering a comet may depend on its period." This might be a reference to the need for the ascertainment correction which we have just made, or it might refer to a possibility of z-dependence for the chance of discovery *given* that perihelion occurs during the period T. Whether or not such a refined ascertainment correction is necessary I do not know; I have not made one.

We shall now illustrate the use of the above formulas when the perturbation distribution has the double exponential form. Two assumptions, (O) and (L), will be made about the initial energy state x; they correspond *very roughly* to the theories of Oort and Lyttleton, respectively. In formulating them and in the subsequent numerical work we shall employ 10^{-5} (astronomical unit)$^{-1}$ as a practical unit for the energy state z. On this scale the empirical evidence [6] leads to a value of about 75 for the standard deviation $\sigma = b\sqrt{2}$ of the perturbation distribution, or say, $b = 50$.

ASSUMPTION (O): *the initial energy state x is very small when compared with b; we shall take $x = 0+$, that is, we let $x \downarrow 0$ in all the preceding formulas.*

ASSUMPTION (L): *the initial energy state x is appreciably larger than b; we shall take a token value of $x = 200$ units.*

Assumption (O) is appropriate in and might be said to be fundamental to Oort's theory; the evidence in its favor will be discussed in a moment. In Lyttleton's theory $x = v^2/600$ a.u.$^{-1}$, where v, in km/sec, is the velocity relative to the sun of the initiating dust cloud. Lyttleton ([7], pp. 82–83) calculates an upper limit of about six km/sec for v, implying an upper limit of about 6000 units for x, and remarks that "had the result come out to, say, less than a tenth of this value [of v], it might have suggested, in the absence of a fuller investigation than at present seems possible, that the process needed exceptionally small though not impossible conditions of relative velocity." Now a reduction of v by a factor of ten implies a reduction of x by a factor of one hundred, so that in his theory $x = 60$ corresponds to "exceptionally small though not impossible conditions of relative velocity." It is not clear where the margin should be drawn, and no doubt general agreement cannot be hoped for at present, but it seems reasonable to conclude that the extreme hypothesis (O) is as inadmissible on Lyttleton's theory as it is vital to Oort's. On the other hand Lyttleton ([7], p. 104) takes as typical a value of $x(= 215)$ corresponding to a period of 10^4 years. This corresponds to a choice for v of about one km/sec. It therefore seems reasonable to take $x = 200$ units to represent the conditions of Lyttleton's theory in the present rough calculations.

Finally, we require an estimate of k, the disintegration probability. We have already mentioned Oort's estimate of 0.019. As it is the square root of k which occurs in our formulas we shall use the values $k = 0.04$, 0.01, 0.0025, and zero.

On assumption (O) we put $x = 0+$, so that $x < z$; thus only the first of the formulas (3.13) is relevant. The continuous part of the expected z-spectrum is then

$$\rho T(1 - \sqrt{k}) \exp\left(-\frac{z}{b}\sqrt{k}\right)\frac{dz}{b}, \qquad z > 0, \tag{4.3}$$

while to this we must add a concentration at $z = 0$ of amount ρT (to be enhanced if we think that fresh comets are intrinsically more luminous). When there is no disintegration, that is, when $k = 0$, the expected z-spectrum (4.3) becomes one with a uniform density. This latter result (equal numbers of comets in equal intervals of the energy axis, when there is no disintegration) was stated by Russell in 1920.

On assumption (L) we shall have $x = 200$, and because only the section $0 < z < 150$ of the z-spectrum is at all reliably known from observation (the existing evidence relating to $z > 150$ is biased by selection effects), we shall have $z < x$; thus only the second of the formulas (3.13) is now relevant, and for the expected z-spectrum we obtain

(4.4)

$$\rho T(1 - \sqrt{k}) \exp\left(-\frac{x}{b}\sqrt{k}\right)\left[\frac{1}{\sqrt{k}} \sinh\left(\frac{z}{b}\sqrt{k}\right) + \cosh\left(\frac{z}{b}\sqrt{k}\right)\right]\frac{dz}{b}, \quad 0 < z < x.$$

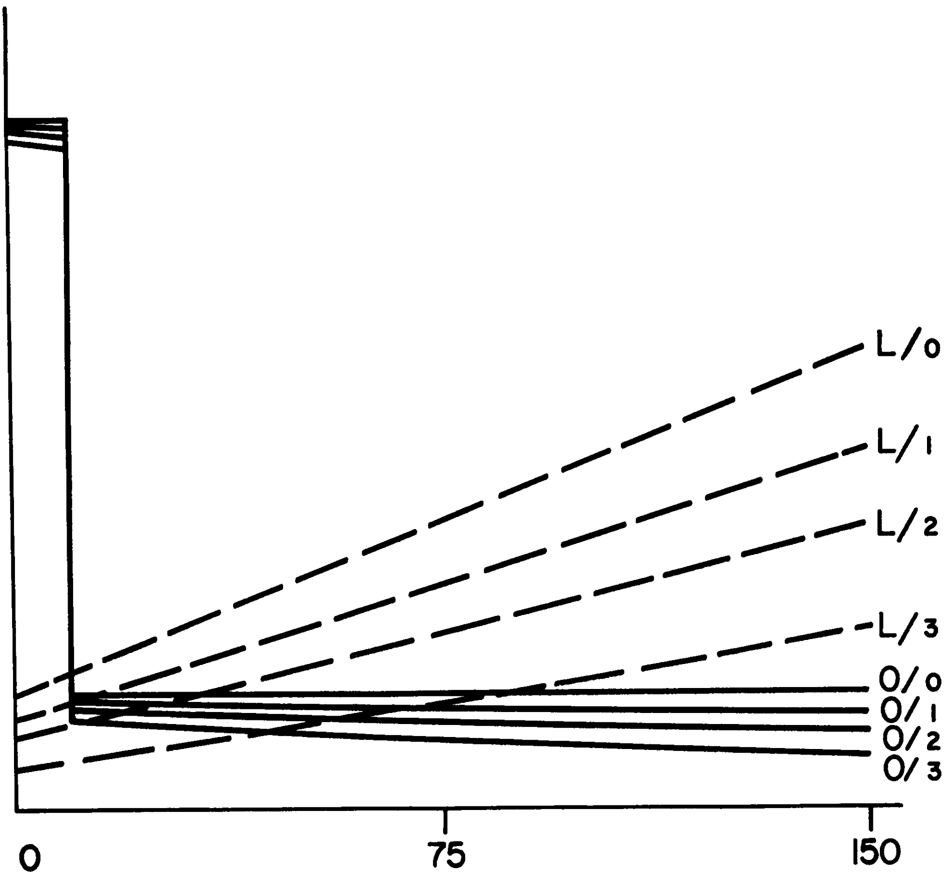

FIGURE 2

Theoretical z-spectra.

The major difference between (4.3) and (4.4) is that in (4.3) the density of the expected z-spectrum is constant or falls off slowly, like $1 - (z/b)\sqrt{k}$, when z increases, while in (4.4) it increases quite rapidly, like $1 + z/b$. This is illustrated in figure 2, where the monotonic-increasing curves relate to assumption (L), while the monotonic-decreasing (or constant) curves relate to assumption (O). For convenience of identification the curves are labeled L/n and O/n, where the initial letter indicates the hypothesis made about x, and the final figure indicates the value of k, thus: (0) $k = 0$, (1) $k = 0.0025$, (2) $k = 0.01$, (3) $k = 0.04$. The horizontal scale gives z in $10^{-5} \times$ a.u.$^{-1}$, while the vertical scale gives the density of the continuous part of the expected z-spectrum, as a multiple of ρT. For the curves associated with hypothesis (O), the concentration at $z = 0$ has been allowed for by spreading it uniformly over the interval $0 < z < 10$.

The expected z-spectrum is not shown in figure 2 beyond $z = 150$ units, because we have no reliable estimate of the observed spectrum beyond this point. It is, however, likely that one will become available in the near future, and then it will be worthwhile computing the expected density for $z > x$ under hypothesis (L); the formula for this is

(4.5)

$$\rho T(1 - \sqrt{k}) \exp\left(-\frac{z}{b}\sqrt{k}\right)\left[\frac{1}{\sqrt{k}}\sinh\left(\frac{x}{b}\sqrt{k}\right) + \cosh\left(\frac{x}{b}\sqrt{k}\right)\right]\frac{dz}{b}, \qquad z > x.$$

This is to be associated with a concentration of amount ρT at $z = x$, just as (4.3) is associated with a concentration of the same amount at $z = 0$.

With regard to the z-spectrum under hypothesis (O), let us note that the concentration at $z = 0$ is equal in amount to the "mass" in the continuous part of the z-spectrum over the range $0 < z < b$, approximately.

So much for the theoretical z-spectrum. Let us now take a look at the observations.

In [6] we made use of the rigorously calculated preorbits of 24 long-period comets; for one of these (Comet 1882 II) the original (preorbit) value of z was 1215 units, and we now exclude it from consideration. The remaining 23 comets, their energy states in the preorbit (z_-) and at perihelion (z_0), are shown in table I.

In the last column of table I the figures in parentheses show the numbers of comets with values of z_- in the stated intervals, based on rigorous calculations of the preorbits for *all* comets for which $z_0 < 200$. These figures are taken from column two of the table on page 366 of Oort's paper [14]; here "all" comets mean "all comets between 1850 and 1936 for which z_0 was known with a mean error of less than 10 units or for which, if no mean error was known . . . the observations . . . extended over at least six months."

It is known (see [6]) that $z_0 = z_- - 50 + w$, where w is a random variable with mean value zero and with a standard deviation of 50 units, approximately. Thus, except for rare exceptions, we shall have $z_0 < z_- - 50 + 100 = z_- + 50$, and so, if $z_- < 150$, then we can fairly safely assert that $z_0 < 200$. This means

TABLE I

VALUES OF z_- AND z_0 FOR 23 COMETS

Range of Values of z_-	Comet	z_-	z_0	Number in the Range
$z < 5$ units	1899 I	-2.7	-107.3	12 (13)
	1898 VII	-1.6	-60.7	
	1886 I	-0.7	-69.4	
	1922 II	$+0.4$	-38.1	
	1902 III	$+0.5$	$+8.1$	
	1914 V	$+1.2$	-14.7	
	1919 V	$+1.6$	-19.3	
	1863 VI	$+1.7$	-49.5	
	1907 I	$+2.5$	-49.9	
	1897 I	$+4.0$	-87.2	
	1889 I	$+4.2$	-69.2	
	1932 VI	$+4.4$	-59.5	
$5 < z < 10$	1925 I	$+5.4$	-56.7	4 (4)
	1886 IX	$+6.3$	-57.7	
	1890 II	$+7.2$	-21.5	
	1853 III	$+8.3$	-81.9	
$10 < z < 15$	1925 VII	$+11.5$	-27.3	1 (2)
$15 < z < 20$	1908 III	$+15.8$	-73.3	1 (2)
$20 < z < 25$	1936 I	$+20.5$	-48.7	2 (1)
	1904 I	$+21.7$	-50.4	
$25 < z < 50$	1886 II	$+31.7$	-47.7	1 (3)
$50 < z < 75$	1905 VI	$+62.1$	-14.2	2 (2)
	1910 I	$+69.2$	$+21.4$	
$75 < z < 100$	none			0 (1)
$100 < z < 150$	none			0 (3 below 200)

that the figures collected by Oort and shown in parentheses in table I give an unbiased estimate of the z-spectrum over the range $z < 150$, but not perhaps beyond this.

The outstanding features of the empirical z-spectrum are (i) the absence of any significantly negative values of z_-, and (ii) the big concentration of values below $z = 5$. It is now widely accepted that hyperbolic original orbits virtually do not exist, and so we can take all entries in the table for which $z < 5$ as belonging to the interval $0 < z < 5$; effect (ii) then becomes the more pronounced. This fact, the concentration of preorbits in the smallest energy class, is the observation on which the whole of Oort's theory is built, but its reality has been questioned by Lyttleton [8]. Lyttleton's main objections, and some replies which might be made to them, are as follows.

(a) The concentration of preorbits in the lowest energy class may be the result of selecting for analysis those perihelion orbits which are most nearly hyperbolic. (Undoubtedly this criticism was appropriate at an earlier stage, but it loses its force now that the calculations have been extended to cover all sufficiently well-observed comets for which $z_0 < 200$.)

(b) The energy interval in question has a width of only five units, while the error in the determination of z_- may be of the order of 10 units. (But the effect of random errors in z_- would be to diminish and not to enhance such a peak.)

This question will no doubt be settled in the course of time by the accumulation of evidence. At the moment it seems reasonable to accept the phenomenon, at least tentatively, as a real one.

If we do so, then we are obliged to follow Oort in concluding that an appreciable number of the 12 or 13 comets in the lowest energy interval are "new" comets which have not previously passed through a (close) perihelion. For if these comets had all passed through the planetary zone in previous revolutions then the over-all energy perturbation (with a root-mean-square value of about 75 units) would have eliminated any such concentration in the 0 to 5 class. It now becomes very desirable to examine the physical characteristics of the comets in the 0 to 5 class, to see if they, or a significant proportion of them, differ in any marked way from what we may call "old" comets; such an investigation has been carried out by Oort and M. L. Schmidt [12], but we shall not summarize it here. A study of the distribution of aphelia for this group of objects would also be of interest.

The relation between the theoretical and empirical z-spectra may be studied by comparing figure 2 with figure 3, in which we have a histogram showing the frequency with which the observed values of z_- (based on the figures in parentheses in the last column of table I) fall into various subintervals of the range $0 < z < 150$. Near $z = 0$ we have used intervals of length five units, although the error of determination may be of the order of 10 units. Thus the peak of over-all width of about 20 units which appears in figure 3 would be compatible with a concentration at $z = 0$. There does not appear to be any prospect of reconciling the observed z-spectrum with hypothesis (L), except perhaps by giving up the hypothesis of statistical equilibrium, made in section 2 when we put $(\infty) = \infty$, but qualitatively at least the agreement with hypothesis (O) is striking. We proceed to examine this more carefully.

It does not seem likely that the present very sparse data could possibly discriminate between the various values for the disintegration constant k, and in fact from figure 2 we see that k has only a small effect [when combined with hypothesis (O)] over the range $0 < z < 150$. Let us first consider hypothesis (O) combined with $k = 0$ (curve O/0 in figure 2). The continuous z-spectrum is then

$$0.02\, \rho T\, dz, \qquad 0 < z < \infty, \tag{4.6}$$

and this has to be combined with a concentration ρT at $z = 0$. Before attempting to compare this with observation we should first convolve it with a (say Gaussian)

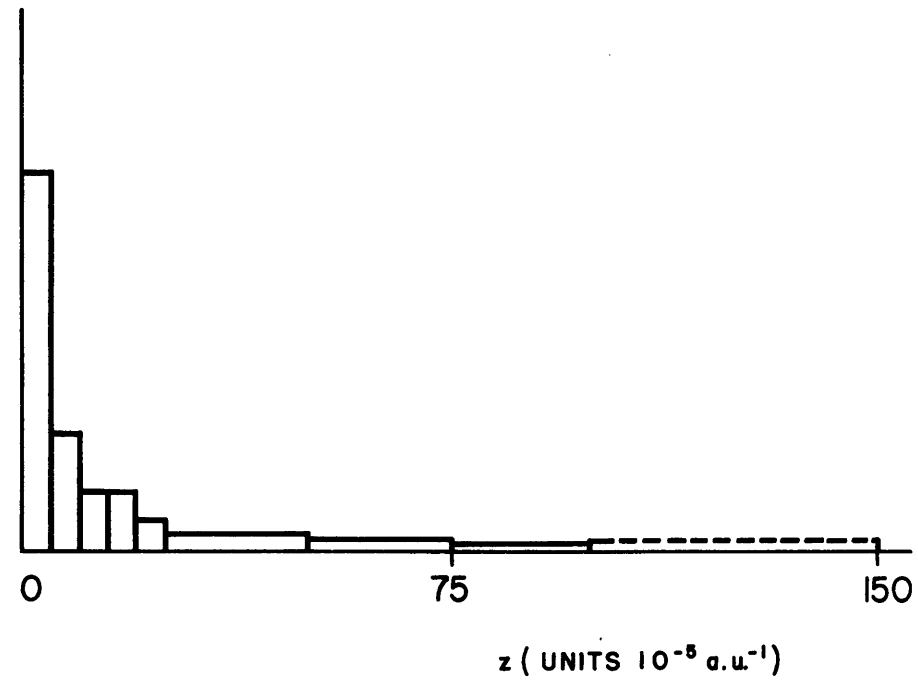

FIGURE 3

The observed z-spectrum.

error distribution having a standard deviation of the order of 10. We can best eliminate our ignorance of the form and precise magnitude of the error distribution by grouping together all the values of z in the interval $0 < z < 25$, and then we find that the observed and expected numbers in the various ranges are as shown in table II. The flatness of the z-spectrum for $z > 25$ is well supported by

TABLE II

COMPARISON OF THE OBSERVED z-SPECTRUM WITH HYPOTHESIS O/0

Range of Values of z	Observed Number	Expected Number
$0 < z < 25$	22	$1.5\ \rho T$
$25 < z < 50$	3	$0.5\ \rho T$
$50 < z < 75$	2	$0.5\ \rho T$
$75 < z < 100$	1	$0.5\ \rho T$
$100 < z < 150$	≤ 3	$1.0\ \rho T$

the empirical figures but is hardly worth a formal significance test with such small numbers. The expected numbers for the intervals $0 < z < 25$ and $25 < z < 150$ are $1.5\ \rho T$ and $2.5\ \rho T$, while the observed numbers are 22 and 6 to 9, respectively. Thus the concentration at $z = 0$ is four to six times as prominent as would be expected, unless we allow a disintegration effect. From figure 2 it can be seen that with a disintegration rate $k = 0.04$, which is somewhat higher than that envisaged by Oort (but just about compatible with the data which he employed—11 disintegrations in 576 perihelion passages), the continuous part of the z-spectrum can be depressed relative to the concentration at $z = 0$; in fact the expected numbers in the two ranges just mentioned now become each equal to $1.4\ \rho T$.

We conclude that the section $0 < z < 150$ of the observed z-spectrum is compatible with Oort's assumption $x = 0$, together with a disintegration rate of the order of 0.04, provided that we assume the "fresh" comets to be intrinsically about four times more luminous than "old" comets.

This is entirely in accordance with Oort's conclusions ([11], p. 105, column 1). Oort and his colleagues at Leiden (especially E. H. Bilo, Mrs. J. van Houten, and H. C. van de Hulst) are currently engaged in a rigorous determination of the empirical z-spectrum over a much wider range of z-values. When this is available a more thorough analysis along the present lines would be worthwhile.

To conclude this discussion it may be useful to give a brief résumé of the deductions made by Oort on the basis of the existence of a peak at $z = 0$ in the z-spectrum, identified with comets entering the planetary zone for the first time. (I refer to the parallel survey of Lyttleton's theory [5] available elsewhere in this Symposium volume.) The first inference is that in the outskirts of the solar system there must exist a swarm of what we may call "precomets," physically identical with the comets with which we are familiar, except that they will on the whole be intrinsically more luminous than "old" comets. It is supposed that these precomets are moving in elliptic orbits with perihelion distances so large that they are quite undetectable by us except when by chance one of them is thrown by stellar perturbations into a new orbit passing ultimately within an astronomical unit or so of the sun. It is assumed that the swarm extends to a radial distance of the order of 200,000 astronomical units; this value is dictated by the large values of $2a$ which have been calculated for some cometary preorbits, and by Oort's estimate that a swarm with a radius of this order could just keep in existence in the face of the disrupting forces of stellar encounters for a period of the order of the age of the solar system. The swarm is supposed to have spherical symmetry, and to possess an isotropic velocity distribution. It is argued that these properties would be acquired by the swarm because of stellar perturbations, if they were not present initially, and the assumption of spherical symmetry is roughly in accordance with the observed distributions of the poles of the orbital planes and the directions of aphelia for long-period comets. Oort first tried a Maxwellian distribution of velocities, but this led to unacceptably high densities in the swarm; he finally assumed a velocity distribution of the form

$$3\,V^2 \frac{dV}{U^3}, \qquad 0 < V < U, \tag{4.7}$$

where U is a constant multiple of the square root of $(R - r)/r$ (r being distance from the sun, and $R = 200{,}000$). This velocity distribution is stable when combined with a suitable density distribution. The arbitrary constant in the latter can be estimated from the observed number (about 100) of new observed comets passing through perihelion per century with perihelion distances of 1.5 a.u. or less. In this way Oort finds the number of comets in the swarm to be of the order of 10^{11}, but with a *total* mass which is only about one-tenth of that of the *earth.*

REFERENCES

[1] K. L. CHUNG and W. H. J. FUCHS, "On the distribution of values of sums of random variables," *Mem. Amer. Math. Soc.*, No. 6 (1951), 12 pp.

[2] W. FELLER, "On the integro-differential equations of purely discontinuous Markoff processes," *Trans. Amer. Math. Soc.*, Vol. 48 (1940), pp. 488–515.

[3] M. FRÉCHET, "Les éléments aléatoires de nature quelconque dans un espace distancié," *Ann. Inst. H. Poincaré*, Vol. 10 (1948), pp. 215–310.

[4] I. V. GALIBINA, "The determination of original and future orbits of some long-period comets," *Bjull. Inst. Teor. Astron.*, Vol. 6 (1958), pp. 630–678.

[5] J M. HAMMERSLEY, "On the statistical loss of long-period comets from the solar system. II," *Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1961, Vol. 3, pp. 17–78.

[6] D. G. KENDALL, "The distribution of energy perturbations for Halley's and some other comets," *Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1961, Vol. 3, pp. 87–98.

[7] R. A. LYTTLETON, *The Comets and Their Origin*, Cambridge, Cambridge University Press, 1953.

[8] ———, "The origin of comets," *Observatory*, Vol. 72 (1952), pp. 33–35.

[9] ———, "On the statistical loss of long-period comets from the solar system. I.," *Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1961, Vol. 3, pp. 229–244.

[10] S. G. MAKOVER, "Ob opredelenii pervonačal'nyh orbit dolgopernofičeskih komet," *Bjull. Inst. Teor. Astron.*, Vol. 6 (1955), pp. 244–248.

[11] J. H. OORT, "The structure of the cloud of comets surrounding the solar system, and a hypothesis concerning its origin," *Bull. Astr. Inst. Netherlands*, Vol. 11 (1950), pp. 91–110.

[12] J. H. OORT and M. SCHMIDT, "Differences between new and old comets," *Bull. Astr. Inst. Netherlands*, Vol. 11 (1950), pp. 259–269.

[13] J. H. OORT, "Origin and development of comets," *Observatory*, Vol. 71 (1951), pp. 129–144.

[14] ———, "Inferences on the origin oı comets derivable from the distribution of the reciprocal major axes," *Mem. Soc. Roy. Sci. Liège* Vol. 13 (1953), pp. 364–371.

[15] H. ROBBINS, "Mixture of distributions," *Ann. Math. Statist.*, Vol. 19 (1948), pp. 360–369.

[16] H. N. RUSSELL, "On the origin of periodic comets," *Astr. J.*, Vol. 33 (1920), pp. 49–61.

[17] E. STRÖMGREN, "Ueber den Ursprung der Kometen," *Publ. Kφbenhavns Obs.*, Vol. 19 (1914), pp. 1–62.

[18] A. J. J. VAN WOERKOM, "On the origin of comets," *Bull. Astr. Inst. Netherlands*, Vol. 10 (1948), pp. 445–472.