Information Content of Particle Tracks

WALTER H. BARKAS
Lawrence Radiation Laboratory, University of California, Berkeley, California
(Received May 26, 1961)

The information content of a track is analyzed with respect to the prime track-variable g and to the particle velocity on which g depends. Quantities are operationally defined that are applicable to emulsion, bubble-chamber or cloud-chamber tracks inclined with arbitrary dip angles. The theory is developed of the projected linear structure of such particle tracks. Previously derived connections between the true value of g and measurable track features are reviewed. A new and independent estimate of g based on the mean blob length is introduced. The two independent quantities, mean gap length and mean blob length, each yield measurements of g. These are combined into an estimate of maximum likelihood. It is argued that in a practical sense this exhausts the information content of the track. The statistical error of this result is evaluated. It is found that correct utilization of the information in the measured blob lengths greatly reduces the error. Suggestions are made regarding technique for the reduction of error in g and in particle masses estimated from grain-density measurements.

I. INTRODUCTION

An exact treatment of the statistical structure of particle tracks in nuclear-research emulsions recently was attempted.[1] Many of the proofs appear to be valid also for bubble tracks and tracks in Wilson chambers, and they have had considerable experimental verification.[2–8]

In this paper we recapitulate these results and find additional track quantities. Then we construct a likelihood function of the grain (bubble, droplet) information in the track. This function yields the expectation value of the grain (bubble, droplet) density, and also measures the statistical reliability of the result. By means of it we learn what are the optimum conditions for observation.

In the case of emulsion, the linear density, g_0, of silver-halide crystals that develop is the primary track variable. In a bubble chamber, g_0 represents the linear density of sites on which bubbles grow, and in a cloud chamber it is the linear density of points on which liquid condenses. (By "grain" hereafter we shall mean the quantity whose density is g_0 irrespective of which of these instruments was used.)

It is assumed that the conditions of observation do not change along the track. No variations of the sensitivity, illumination, etc., are considered, and possible distortions of the track not discussed. In this paper, too, we limit the investigation to the *linear* track structure. Further information is contained in the track width, in the delta rays, and in the scattering. Moreover, it is sometimes possible to decrease the effective track saturation by letting track elements diffuse radially before counting them. These techniques are left for other theoretical studies.

The connection between the grain density and the particle velocity usually is an empirical one. The dependence on velocity is strong. The grain density over a wide range tends to be proportional to the inverse square of the velocity.

The problem now posed is how best to determine g_0 from the measureable features of the track granularity. Besides this immediate goal, we seek criteria for the improvement of emulsion quality. We wish to know, in general terms, how to alter the emulsion in manufacture to increase the information density attainable. The parallel problem in bubble chambers is how to adjust the temperature, the age of the tracks, the amount of expansion, and the optics so as to obtain the optimum bubble-image size as well as bubble density. In a Wilson chamber, the gas pressure, the expansion ratio, and the age of the tracks are among the variables, the adjustment of which also can optimize the information density existing in track photographs.

II. STATISTICAL GEOMETRY OF PARTICLE TRACKS

In this section we review and supplement operational definitions and mathematical results of reference 1 that are needed for this analysis.

A track is seen as a somewhat indefinite locus of grain images distributed generally along the path of the particle that produced it. The grain centers are displaced by small random amounts from the most probable particle trajectory. The images generally are not all of the same size. Some grains occult others or fuse with them. Owing also to imperfect optical resolution, clusters of several unresolved grains may be present in the image of a track segment. Such clotting of the grain images always causes a loss of information

[1] Walter H. Barkas, Lawrence Radiation Laboratory Report UCRL-9181 (unpublished). Also, Third International Conference on Nuclear Photography, Moscow, 1960 (unpublished); and Lawrence Radiation Laboratory Report UCRL-8687 (unpublished).
[2] D. J. Holthuizen, Nuovo cimento **17**, 928 (1960).
[3] Akbar Ahmadzadeh and Nripendra N. Biswas, Nuovo cimento **19**, 958 (1961).
[4] B. Hahn and E. Hugentobler, International Conference on Instrumentation for High-Energy Physics, Berkeley, September, 1960 (unpublished).
[5] D. A. Glaser, D. C. Rahm, and C. Dodd, Phys. Rev. **108**, 1046 (1957).
[6] C. O'Ceallaigh, Suppl. Nuovo cimento **12**, 412 (1954).
[7] P. H. Fowler and D. H. Perkins, Phil. Mag. **46**, 587 (1955).
[8] J. Patrick and W. H. Barkas, Lawrence Radiation Laboratory Report UCRL-9692 (unpublished).

about the grain density, but by analysis we can greatly reduce the information loss in partially saturated tracks.

Let $(\pi/2)-\delta$ be the angle between the direction of motion of the particle and the line of sight. Then δ is the "dip angle" in unprocessed emulsion. (Since all particle trajectories are inclined more or less, we shall from the outset treat the general case of tracks with arbitrary dip angles.) For the analysis, let the track grain images be projected on a plane perpendicular to the line of sight. All measurements are to be made in this plane.

A "resolution distance" a is defined as the minimum distance between centers of two grain images at which they can be resolved into two objects separated by a gap. If c is the distance, projected on the most probable particle trajectory, between the centers of consecutive track grains, and if this exceeds the resolution distance, a gap of length $c-a$ is said to exist in the track. The grains are not all of the same size, but a mean value α of the resolution distance exists. The quantity α is also called the *mean grain diameter*, but it has more significance than this name implies. When referred to emulsion, for example, the effects of several diverse quantities are lumped in this parameter. Among these are the original halide crystal size, the amount of physical development sustained by the silver grains, and the optical resolution of the observer-microscope instrument. When certain equipment is used, α is even affected by the reaction time of the observer.

A "cluster of class C_l" is a segment of track bounded by gaps with lengths exceeding l, and in which no gap longer than l exists. The density $H(l)$ of such clusters is also equal to the average number of gaps with lengths exceeding l in unit track length. The cluster density when $l=0$ has the special name "blob density" or "gap density." The blob density is symbolized by $B=H(0)$.

An important measureable track quantity is the *lacunarity* L. This is the average fraction of the track segment that consists of gaps. Thus

$$L=-\int_0^\infty \cdot l\frac{dH}{dl}dl. \tag{1}$$

The quantity $1-L$ is known as the track *opacity*.

It was shown, as a fundamental result of reference 1, that the density of clusters of class C_l is

$$H(l)=ge^{-g(\alpha+l)}, \tag{2}$$

where l is the projected gap length, $g=g_0\sec\delta$, and g_0 is the value that would be found for g were the track not inclined.

The original treatment was carried out for primary emulsion grains that were traversed by the moving particle. Because of the finite noninterpenetrating volumes of the silver-halide crystals, this proof required close analysis. The extension to secondary grains, the positions of which are in effect unrestricted, was an obvious step. In bubble- and cloud-chamber tracks the g sites are almost dimensionless, and have a Poisson distribution. For them, Eq. (2) also applies.

Inclination of the track shortens most gaps and closes some entirely. Typically, a grain center is displaced from the particle trajectory. This makes it possible for a gap to appear in an inclined track that, in accord with the definitions above, would not have been counted as a gap were the track not inclined. The effect of the inclination on a track of grain density g_0 is to produce a grain structure pattern in its projection that is statistically equivalent to that in an uninclined track having a grain density g. This *exactness* of the correspondence between the structure of the projection of the inclined track and a flat one of higher grain density was not brought out in reference 1.

The same track with different optical resolution will have an observed structure corresponding to a changed value of α, but, of course, g remains unaltered. The theory of the track structure must contain this invariance.

The gap density in the track projection is

$$B=ge^{-g\alpha}, \tag{3}$$

and the lacunarity is

$$L=e^{-g\alpha}. \tag{4}$$

Quantities like L and B that depend on g are often referred to as *ionization parameters*. Several other track quantities also have simple expectation values. For example, that for the number of grains per blob is $e^{g\alpha}$. The expression $1-e^{-g\alpha}$ for the track opacity is also the probability for no gap to be left between successive developed grains. While these expectation values are exact, the complete distribution function of L or B can be derived only approximately by introducing a blob model. (See Appendix.)

The product $(\alpha+l)H(l)$ is a universal function of $(\alpha+l)g$. Any observation of $H(l)$, therefore, is a measure of g.

The expectation value, $\langle l\rangle$, of the gap length is equal to $1/g$. If the measured mean gap length is designated $\bar{l}$, we are led to an important estimate of the grain density which we shall call g_1:

$$g_1=1/\bar{l}. \tag{5}$$

A new result of this work is the utilization of the mean blob length to estimate g. The expectation value, $\langle b\rangle$, of the blob length is $(e^{\alpha g}-1)/g$. A second important estimate, g_2, therefore can be derived from the observed mean blob length $\bar{b}$:

$$(e^{g_2\alpha}-1)/g_2=\bar{b}. \tag{6}$$

The numerical relation between g and the mean blob length is given in Table I.

For what follows it is essential that g_2 be independent of g_1. Their relationship, therefore, must be elaborated. For g_1 and g_2 to be independent, it is necessary that no

TABLE I. The important track quantities are tabulated as functions of the track lacunarity. The table also relates them to each other. The quantity $\Lambda\sigma_g^2/g$ gives: (a) the error for the maximum likelihood solution using model 1 for σ_b, (b) the maximum likelihood solution using model 2, and (c) the error when g_1 alone is measured. The corresponding values of ω and p are also tabulated.

				Combined g_1 and g_2						g_1 alone
				(a) Model 1			(b) Model 2			(c)
L	αg	αB	$\langle b\rangle/\alpha$	ω	$\Lambda\sigma_g^2/g$	p	ω	$\Lambda\sigma_g^2/g$	p	$\Lambda\sigma_g^2/g$
0.00	∞	0.0000	∞	0.000	∞	0.00	0.000	∞	0.000	∞
0.05	2.9957	0.1498	6.342	0.143	2.86	0.350	0.139	2.77	0.361	20
0.10	2.3026	0.2303	3.908	0.212	2.12	0.472	0.202	2.02	0.495	10
0.15	1.8971	0.2846	2.987	0.271	1.81	0.553	0.255	1.70	0.587	6.67
0.20	1.6094	0.3219	2.485	0.326	1.63	0.614	0.304	1.52	0.659	5.00
0.25	1.3863	0.3466	2.164	0.376	1.51	0.664	0.349	1.40	0.717	4.00
0.30	1.2040	0.3612	1.938	0.425	1.42	0.706	0.392	1.31	0.765	3.33
0.35	1.0498	0.3674	1.769	0.471	1.35	0.742	0.435	1.24	0.805	2.86
0.40	0.9163	0.3665	1.637	0.517	1.29	0.774	0.476	1.19	0.840	2.50
0.45	0.7985	0.3593	1.531	0.561	1.25	0.803	0.517	1.15	0.870	2.22
0.50	0.6932	0.3466	1.440	0.604	1.21	0.828	0.558	1.12	0.896	2.00
0.55	0.5978	0.3288	1.369	0.646	1.17	0.851	0.599	1.09	0.918	1.82
0.60	0.5108	0.3065	1.305	0.688	1.15	0.873	0.640	1.07	0.937	1.67
0.65	0.4308	0.2800	1.250	0.728	1.12	0.893	0.682	1.05	0.953	1.54
0.70	0.3567	0.2497	1.201	0.769	1.10	0.911	0.724	1.03	0.966	1.43
0.75	0.2877	0.2158	1.158	0.808	1.08	0.928	0.768	1.02	0.977	1.33
0.80	0.2231	0.1785	1.120	0.848	1.06	0.944	0.812	1.01	0.986	1.25
0.85	0.1625	0.1381	1.079	0.886	1.04	0.960	0.857	1.01	0.992	1.18
0.90	0.1054	0.0949	1.054	0.925	1.03	0.973	0.903	1.00	0.997	1.11
0.95	0.0513	0.0487	1.027	0.963	1.01	0.986	0.951	1.00	0.999	1.05
1.00	0.0000	0.0000	1.000	1.000	1.00	1.000	1.000	1.00	1.000	1.00

information about the blob lengths be obtainable from the gap lengths, and that the blob lengths suffice for a measure of g when no knowledge exists of the gap lengths. These conditions are satisfied. When one measures the gap lengths, he gains no information about the blobs. A measured blob length, therefore, is entirely new information. Moreover, the value g_2 can be calculated from the mean blob length while making no reference to the gap lengths or to the length of the track segment (which contains both the gap and blob lengths).

A peculiarity of the exponential gap distribution is that, if all gaps are shortened by the same amount, the mean gap length remains unaltered. It follows from this circumstance that whatever the distribution of the amounts by which gaps are shortened, the mean gap length is unaffected. This is a very useful deduction. It means, for example, that the growth of grains and bubbles and displacements caused by their crowding do not affect the mean gap length. This quantity $\bar{l}=L/B$, therefore, is an excellent measure of g when there are many gaps in the track. Such a measurement also does not require knowledge of α. On the other hand, when the grain density is high, it is the mean blob length that contains most of the grain-density information. To use it, however, the quantity α must also be known.

An estimate g_B of the grain density is found from the blob density:

$$g_B e^{-g_B\alpha}=B.$$

Another, g_L, is obtained from the lacunarity:

$$g_L=-(\ln L)/\alpha.$$

The estimates g_B and g_L are combinations of g_1 and g_2.

Long blobs are distributed exponentially. Their distribution is approximately $qe^{-qb}db$ in the interval of blob length between b and $b+db$. The blob coefficient, q, is related to an estimate, g_q, of the grain density by

$$q=g_q/(e^{\alpha g_q}-1-\alpha g_q).$$

A measurement of g_q contains part of the information in g_2.

The gap coefficient,[6,7] which we designate g_g, is

$$g_g=\ln[H(l_1)/H(l_2)]/(l_2-l_1). \tag{7}$$

Although this fact was not known at the time that it was introduced, g_g also is an estimate of the true grain density. It contains part of the information in g_1.

As g is varied, B passes through a maximum, $B_{\max}$, when $\alpha g=1$, so that

$$\alpha=(eB_{\max})^{-1}. \tag{8}$$

A measurement of $B_{\max}$ under normal observing

conditions is a correct procedure for determining α, at least for tracks with $g \approx \alpha^{-1}$. Another measure of α is $\alpha = -(L/B)\ln L$—an important formula.

The cluster lengths or blob lengths have a somewhat more complex behavior than the gap lengths. The probability that the next j developed grains that follow a developed grain will leave no gap is $(1-e^{-\alpha g})^j$. The mean length of a blob consisting of $j+1$ grains takes the form $\alpha + j\beta$, where β is the average length added to a blob by the addition of a grain. Then the expectation value, $\langle b \rangle$, of b is[1]

$$\langle b \rangle = \alpha + (e^{\alpha g}-1)\beta, \tag{9}$$

and

$$\beta = (e^{\alpha g}-1-\alpha g)/g(e^{\alpha g}-1) \tag{10}$$

since $\langle b \rangle = (e^{\alpha g}-1)/g$.

Let the distance parallel to the particle path between the centers of the first and last grains in a blob be x. The frequency with which blobs of $j+1$ grains occur relative to those with j grains is $1-e^{-\alpha g}$. Therefore, the fraction of blobs having $b-\alpha$ greater than x varies like e^{-qx} for $x \gg \alpha$.

There are two kinds of linear structure elements, blobs and gaps, that alternate in a track segment. The elementary track cell is comprised of a blob and an adjacent gap. The track is generated by a repetition of this unit. The length of a cell is a random variable equal to the sum of two random variables: the blob length and the gap length. The distribution function of each depends on g. Their observed distributions provide all the information regarding g.

In the course of this work it was surprising to discover that the whole information content of the gaps resided in the *mean gap length*. Thus the mean gap length is an example of a *sufficient statistic* in the terminology of Sir Ronald Fisher, who in 1920 first found this most efficient estimate.[9] A sufficient statistic contains all the information in the observations from which it is derived. The mean gap length is such a statistic because the gap lengths have an exponential distribution (see below). The estimate g_1 derived from the gaps, therefore, completely exhausts their information content relative to g. The variance of g_1, based on N cells, has the irreducible minimum of g^2/N.

The situation with respect to the blob lengths, on the other hand, has complicating elements. The blob-length distribution function depends on the grain-size distribution, and even for idealized models, an analytic blob-length distribution is difficult to derive. One may deduce, however, that the blob-length distribution falls exponentially for long blobs, and when g is large, the distribution is approximately exponential. The blob lengths contain the bulk of the track information only when g is high, but when it is high, the mean blob length approaches a sufficient statistic. Now, in addition, if a large number of blobs are used to make the estimate of $\bar{b}$, by the central limit theorem,[10] $\bar{b}$ will approach a Gaussian distribution. Its distribution function then is written

$$(N^{\frac{1}{2}}/(2\pi)^{\frac{1}{2}}\sigma_b)\exp[-N(\bar{b}-\langle b \rangle)^2/2\sigma_b^2],$$

where σ_b^2 is the variance of b, and N is the number of cells in the track segment. This expression describes the distribution of $\bar{b}$, given $\langle b \rangle$, or the distribution of $\langle b \rangle$, given a measurement, $\bar{b}$. The distribution of expectation values is assumed to be such that *a priori* every $\langle b \rangle$ is equally probable.

Since the distribution function of b is intractable for any except artificial models, the calculated moments of b beyond the first must be approximations. The variance of b, however, is a readily observed quantity. Moreover, σ_b/α is a function of the lacunarity that may be measured on any tracks. We therefore can choose to consider both α and σ_b/α as calibration data describing the instrument.

The mean blob length and mean gap length provide independent and efficient[11] estimates g_2 and g_1 of g. We now wish to combine them so as to obtain the best estimate of g.

III. MAXIMUM LIKELIHOOD ESTIMATE OF GRAIN DENSITY

We have distribution functions for the gap lengths, and for the mean blob length. Then the *likelihood function*[11] of the configuration of gaps and blobs observed in N cells can be constructed as follows:

$$P = \frac{1}{\sigma_b}\left[\frac{N}{2\pi}\right]^{\frac{1}{2}} g^N \exp\left[-g\sum_1^N l_i\right]\exp\left[\frac{-N(\bar{b}-\langle b \rangle)^2}{2\sigma_b^2}\right], \tag{11}$$

where l_i is the length of the ith gap. The *sufficiency* of the mean gap length as a statistic now is easy to prove. We have merely to replace $\sum_1^N l_i$ by $N\bar{l}$. The only gap information that appears in the likelihood function then is $\bar{l}$.

A particular value of g in Eq. (11) maximizes P. When N is sufficiently large (see Appendix B), this value, g', is the one of maximum likelihood. The function P also estimates the probability that a value of g other than the mode could be the true value. Any desired confidence intervals can be quoted with a knowledge of this function.

Let $W = \ln P$. Then the condition of maximum likelihood is $\partial W/\partial g = 0$, or

$$\frac{1}{g} - \bar{l} + \frac{\bar{b}-\langle b \rangle}{\sigma_b^2}\frac{d\langle b \rangle}{dg} = 0, \tag{12}$$

since

$$\langle b \rangle = (e^{\alpha g}-1)/g, \quad d\langle b \rangle/dg = (1-e^{\alpha g}+\alpha g e^{\alpha g})/g^2.$$

[9] R. A. Fisher, Monthly Notices Royal Astron. Soc. **80**, 758 (1920).

[10] H. Cramèr, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, New Jersey, 1946).

[11] R. A. Fisher, Proc. Cambridge Phil. Soc. **22**, 700 (1925).

Also

$$\bar{l}=1/g_1, \quad \text{and} \quad \bar{b}=(e^{\alpha g_2}-1)/g_2.$$

The two estimates g_1 and g_2 are supposed to differ little from each other, so that a solution g' of Eq. (12), valid through the first order in $g_2-g_1\equiv\epsilon$, will suffice. The solution is the linear combination

$$g'=\omega g_1+(1-\omega)g_2=g_2-\omega\epsilon, \tag{13}$$

with

$$\frac{\omega}{1-\omega}=\frac{L^2(\ln L)^2}{(1-L+\ln L)^2}\left(\frac{\sigma_b}{\alpha}\right)^2.$$

The weighting of g_1 and g_2 is not critical, and any reasonably good measurement of L may be used in the expression for ω.

When an empirical value of σ_b has not been obtained, a theoretical estimate must be used. In the Appendix, limits between which σ_b lies have been estimated for any physical tracks. One estimate based on a completely random grain spacing in the blob has been made by Stapp.[12] This model we shall designate model 1. In reference 1, σ_b^2 was obtained for a completely ordered spacing. The results using this model, designated model 2, are also given. All real tracks should exhibit blob-variance behavior intermediate between these extremes. The more inclined the track, the better it should be approximated by model 1.

Equation (13) provides a best estimate of g'. It remains to calculate its reliability. The second derivative of W with respect to g provides a measure of the width of the probability peak. The likelihood function, Eq. (11), approaches a Gaussian as N becomes large. For a Gaussian, the variance of g' is given by

$$\sigma_{g'}^2=-(\partial^2W/\partial g^2)^{-1}, \tag{14}$$

at the maximum of W. We adopt this expression for the error. Then we find for a track length Λ,

$$\Lambda\sigma_{g'}^2/g' \; (\equiv 1/p)=\omega/L. \tag{15}$$

This function, corresponding to the theoretical limits (model 1 and model 2) for σ_b^2, is also included in Table I. In order to demonstrate the substantial gain effected by introducing the mean blob length, a column is also given that is the calculated uncertainty remaining in the grain density when the mean gap length alone is used. It can be seen that a very important amount of information has been salvaged by utilizing the blob information—especially in near-saturated tracks. Moreover, the required measurements are of types that are efficiently made with automatic track analysis equipment. The requirement that N be large (say 10 or more) for the Gaussian to represent well the mean blob length distribution, seldom limits the applicability of the theory.

In the error estimates, no allowance has been made for the uncertainty of α or for systematic errors of other sorts. For minimum error, the calibration measurements of α and σ_b are to be made, independently of g_1 and g_2 on other track segments. If a faulty measuring technique is employed, $-(L/B)\ln L(=\alpha)$ may vary with the dip of the track. In other circumstances α could depend slightly on g. For example, in a bubble chamber the energy required to produce the bubbles might so lower the temperature in the vicinity of a saturated track that the bubble size is reduced. For these reasons it is good technique to make the calibration measurements on tracks similar to the one in which g' is to be measured, and, of course, with the identical equipment. If no separate estimate of α is available, the likelihood function can be considered to depend on g and α. Then its maximum as a function of both parameters may be found. There is generally a loss of information when such a procedure is necessary, however.[13]

To measure $\bar{b}$, one could observe only the track and gap lengths along with the blob density. Then $\bar{b}$ would be estimated from $(1-L)/B$. A wise check would be to measure the blobs themselves, because measurement errors can tend to be systematic. This should especially be done if g_1 and g_2 fail to agree as well as expected.

The results of reference 1 giving statistical errors in ionization parameters were based on inexact assumptions and are superseded by these results.

The possibility that the instrument sensitivity may vary with track position, particularly with depth, has previously been mentioned. This effect must be eliminated by empirical correction. Each estimage of g requires multiplication by a factor $f(r,g)$, where r represents the point coordinates. As indicated, f may also be a function of g.

Some numerical data drawn from grain, bubble, and droplet measurements are now cited in order that the reader may be oriented as to orders of magnitude.

For Ilford K.5 emulsion, using optics of high numerical aperture, α is about 0.48 micron. At the minimum of ionization $g\approx 2000/\text{cm}$ (5000, perhaps, if the emulsion is hypersensitized). In K.5 emulsion g saturates at a value of 60 000–70 000 per centimeter for singly charged particles.

In a propane bubble chamber α can be about 0.03 cm, and at the minimum of ionization, g may be 20–30/cm. In a bubble chamber g does not saturate, but for $g>100/\text{cm}$ the lacunarity becomes very low.

On a photograph in which a cloud-chamber track image is reduced to $\frac{1}{10}$ actual size, $\alpha\approx 10^{-3}$ cm, and at the minimum of ionization, g is perhaps 250/cm on the film. There is no saturation of g in a cloud chamber and

[12] Henry P. Stapp, Lawrence Radiation Laboratory (private communication).

[13] A. Ahmadzadeh, Lawrence Radiation Laboratory Report UCRL-9527 (unpublished). Quite another approach to this problem was taken by Ahmadzadeh in this document which since has been withdrawn. His form of the likelihood function of g makes redundant use of gap data and consequently weights the data incorrectly. The form of the likelihood function itself is an approximation which underestimates the error in the result.

the droplets can be allowed to diffuse. When droplets are individually countable, the present methods of analysis do not provide any additional information, but the diffused image of the track then limits the accuracy of a curvature measurement.

IV. MASS ESTIMATION

A principle of mass-ratio determination that uses grain-density information alone is the following:

Segments of track having the same initial and terminal grain densities have lengths in proportion to the masses of the equally-charged particles that produced them.

A corollary is: If tracks of stopping particles of equal charge have equal mean grain densities, the track lengths are proportional to the particle masses.

As an application of the preceding theory we shall apply this corollary. Let the track of a particle that comes to rest be broken into segments of equal projected lengths. Such segments are to be short enough so that the average grain density in one of them is negligibly different from that at its center. For protons in emulsion, a length of about 100 microns might be suitable. We let the segment length be unity in what follows.

The grain densities in different parts of a track are not known equally well. They must be weighted by their reciprocal variances. Thus, for n uninclined segments,

$$\bar{g}_0(n)=\sum_{i=1}^{n}\left(\frac{g'}{\sigma_{g'}^2}\right)_i \Big/ \sum_{i=1}^{n}\left(\frac{1}{\sigma_{g'}^2}\right)_i. \tag{16}$$

The expression $(g'/\sigma_{g'}^2)_i$ for the ith cell can be designated p_i and the measured g' in this cell is g_i'. Now from Eq. (15),

$$(g'/\sigma_{g'}^2)_i=L_i/\omega_i=p(L_i), \tag{17}$$

where L_i and ω_i are the measured values in the ith cell.

Suppose that when the particle mass is known we symbolize the grain density by γ and reserve g for the grain density in the track of the particle of unknown mass. Then for a known particle, the weighted mean grain density, $\bar{\gamma}_0$, in the terminal n cells (length, $\sum_1^n \sec\delta_i$) of its track is

$$\bar{\gamma}_0\left(\sum_1^n \sec\delta_i\right)=\sum_1^n p_i\cos\delta_i \Big/ \sum_1^n\left(\frac{p_i}{\gamma_i}\right). \tag{18}$$

Here the argument $(\sum_1^n \sec\delta_i)$ of the function $\bar{\gamma}_0$ is the track length and, as for g_0, the subscript zero indicates that the quantity is the calculated value for an uninclined track. The angle δ_i is the inclination angle of the track in the ith cell. The average of $\bar{\gamma}_0$ for many particles we designate $(\bar{\gamma}_0)_{\text{av}}$. Each average is understood to be made for a fixed argument.

The track of an unknown particle is also to be segmented into portions of unit projected length. The available length is broken into n' segments like those comprising the track of the particle of known mass. In each interval, g and δ are measured. Then $p\cos\delta$ and p/g are determined. Using these numbers we calculate a single number,

$$\bar{g}_0\left(\sum_1^{n'} \sec\delta_j\right)=\sum_1^{n'} p_j\cos\delta_j \Big/ \sum_1^{n'}\left(\frac{p_j}{g_j'}\right). \tag{19}$$

To apply the corollary condition above, the value of $\bar{\gamma}_0$ or $(\bar{\gamma}_0)_{\text{av}}$ is tabulated or graphed as a function of its argument. Then we find what value of $\sum_{i=1}^n \sec\delta_i$ makes

$$\bar{\gamma}_0(\sum_{i=1}^n \sec\delta_i) \quad \text{or} \quad (\bar{\gamma}_0)_{\text{av}} \quad \text{equal to} \quad \bar{g}_0(\sum_{i=1}^{n'} \sec\delta_j).$$

When $\bar{g}_0$ and $\bar{\gamma}_0$ are equal, the ratio of the track lengths,

$$\sum_{j=1}^{n'} \sec\delta_j \Big/ \sum_{i=1}^{n} \sec\delta_i,$$

is the estimated mass ratio. The mass-ratio limits corresponding to a confidence interval of $\pm r$ standard deviations are found from the following condition:

$$\bar{g}_0(\sum_{j=1}^{n'} \sec\delta_j)=\bar{\gamma}_0(\sum_{i=1}^{n} \sec\delta_i) \pm r[(\sum_{i=1}^{n}(p_i/\gamma_i))+(\sum_{j=1}^{n'}(p_j/g_j))]^{-\frac{1}{2}}. \tag{20}$$

This error estimate neglects the energy-loss straggling.

Means for mass estimation that combine grain density with multiple scattering or curvature in a magnetic field have been developed, and are especially valuable if the particle fails to come to rest. The result is always improved if the properly weighted combination of g_1 and g_2 is used in preference to another estimate of g. Of course, it may not be necessary to segment the tracks of high-velocity particles if the velocity changes little in the observed portion of the track. On the other hand, it could be advisable to segment tracks in bubble or cloud chambers when the track-to-camera distance and the track aspect change along the track.

ACKNOWLEDGMENT

Thanks are owed to Dr. Henry P. Stapp for his efforts to construct mathematical track models, and for very useful discussions on the subject of likelihood theory.

MATHEMATICAL APPENDIX

A. Blob Length Variance

Let $y_1, y_2, \cdots, y_n$ be the projections on the particle path of the distances between successive grain centers in a blob. The expectation value of y_i is $\langle y\rangle$, and its

variance is σ_y^2. The mean square blob length is

$$\langle b^2\rangle = e^{-g\alpha}\sum_{n=0}^{\infty}\langle(\alpha+y_1+y_2+\cdots+y_n)^2\rangle(1-e^{-g\alpha})^n$$

$$=\alpha^2+(2\alpha\langle y\rangle+\sigma_y^2)e^{-g\alpha}\sum_{n=0}^{\infty}n(1-e^{-g\alpha})^n$$

$$+\langle y\rangle^2 e^{-g\alpha}\sum_{n=0}^{\infty}n^2(1-e^{-g\alpha})^n.$$

The mean blob length also can be calculated:

$$\langle b\rangle = e^{-g\alpha}\sum_{n=0}^{\infty}\langle(\alpha+y_1+y_2+\cdots+y_n)\rangle(1-e^{-g\alpha})^n$$

$$=\alpha+\langle y\rangle e^{-g\alpha}\sum_{n=0}^{\infty}n(1-e^{-g\alpha})^n.$$

Then $\sigma_b^2=\langle b^2\rangle-\langle b\rangle^2$ can be found, and in general,

$$\sigma_b^2=e^{g\alpha}(1-e^{-g\alpha})\sigma_y^2+e^{2g\alpha}(1-e^{-g\alpha})\langle y\rangle^2.$$

This formula neglects the variance of the diameter of a single grain, which could be included as an added term, but is hardly justified. Stapp[12] has made calculations based on a model in which the grain centers have a Poisson distribution, and in which each grain has the diameter α. To avoid interpenetration of the grains, they can be thought to be displaced from the particle trajectory. For this model we calculate

$$\langle y\rangle\equiv\beta=(1/g)[1-\alpha g e^{-\alpha g}/(1-e^{-\alpha g})],$$

$$\langle y^2\rangle=[2(1-e^{-\alpha g}-\alpha g e^{-\alpha g})-\alpha^2 g^2 e^{-\alpha g}]/g^2(1-e^{-\alpha g}),$$

and

$$\sigma_y^2=(1/g^2)[1-\alpha^2 g^2 e^{-\alpha g}/(1-e^{-\alpha g})^2].$$

When expressed as a function of the lacunarity, the quantity σ_b^2/α^2 then can be written $(1-L^2+2L\ln L)/L^2(\ln L)^2$.

Stapp's model (model 1) permits the maximum variance of y. The noninterpenetrability of crystals in emulsion and the finite size of bubbles in a bubble cluster tends to reduce this variance. Barkas[1] calculated σ_b for a model (model 2) in which the term containing σ_y^2 was omitted. This yields for σ_b^2/α^2 the expression $(1-L+L\ln L)^2/[L^2(\ln L)^2(1-L)]$.

The two models probably represent opposite extremes, and in actual tracks intermediate behavior should be observed. The results of numerical calculations for these models are included in Table I.

B. Gaussian Approximation Error

Let m be the number of grains in a typical blob. The expectation value, $\langle m\rangle$, of this number is $e^{\alpha g}$, and an estimate M derived from the mean length, $\bar{b}$, of N blobs is $e^{\alpha g_2}$. For model 2, the distribution function, $P(M,\langle m\rangle)$ of $\langle m\rangle$ can be written down exactly. It is

$$P(M,\langle m\rangle)=(\langle m\rangle-1)^{N(M-1)}/\langle m\rangle^{NM}.$$

In this expression M is a sufficient statistic.

We put $M-\langle m\rangle=\epsilon$. Then

$$N(\bar{b}-\langle b\rangle)^2/2\sigma_b^2=\epsilon^2/2\sigma_\epsilon^2.$$

Then also the distribution function $P(M,\langle m\rangle)$ can be developed in powers of ϵ. It takes the unnormalized form

$$Q(M,\epsilon)\approx\exp(-\epsilon^2/2\sigma_\epsilon^2)\left[1-\frac{N(2M-1)\epsilon^3}{3M^2(M-1)^2}+\cdots\right].$$

with $\sigma_\epsilon^2=M(M-1)/N$.

Whereas the modal value of $\langle m\rangle$ is M, the presence of the second term in the brackets indicates that the mean value of ϵ is not zero. This term measures the deviation of the model 2 distribution function from the assumed Gaussian. The mean value of ϵ is approximately $-(2M-1)/N$. On the other hand, its statistical uncertainty, $\sigma_\epsilon=[M(M-1)/N]^{\frac{1}{2}}$, always exceeds the above systematic effect when N is greater than $4+[1/M(M-1)]$.

When M is small, N must be large, but then the weight given the blob information is negligible. The Gaussian, therefore, is probably always a satisfactory approximation for $N>4$. Model 2 describes the real blob structure well enough so that this result can be applied with confidence.