**Title**
THE CONCEPT OF SETS ENCHAINED BY A STOCHASTIC PROCESS AND ITS USE IN CASCADE SHOWER THEORY



**Author**
Rankin, Bayard.

**Publication Date**
1954-07-01

UNIVERSITY OF
CALIFORNIA

# Radiation Laboratory



BERKELEY, CALIFORNIA

## DISCLAIMER

UCRL-2652
Unclassified
Physics

UNIVERSITY OF CALIFORNIA
Radiation Laboratory
Contract No. W-7405-eng-48

# THE CONCEPT OF SETS ENCHAINED BY A STOCHASTIC PROCESS AND ITS USE IN CASCADE SHOWER THEORY

Bayard Rankin
(Thesis)

July, 1954

Berkeley, California

Table of Contents

# THE CONCEPT OF SETS ENCHAINED BY A STOCHASTIC PROCESS AND ITS USE IN CASCADE SHOWER THEORY

Bayard Rankin
Statistical Laboratory
and
Radiation Laboratory, Department of Physics
University of California, Berkeley, California

July, 1954

## ABSTRACT

Almost all solutions for the number distribution of electrons and photons in cascade result from one general methodology. These solutions (including those of Nordsieck, Lamb, Uhlenbeck, Scott, Bhabha, Ramakrishnan, Janossy, Messel, Potts, and one previously given by the author) are obtained by solving for the probability, or probability density, that the electron-photon system belongs to one or another of a set of states after it has penetrated x-units of absorbing material. This general method of approach is in contradistinction to the conceivable method of determining the probability weighting for each state individually. It has the advantage of entailing fewer parameters -- just enough to specify the sets rather than the detailed configurations of states within these sets. It may have the disadvantage of leading to equations that are non-linear or otherwise difficult to solve, for the Markovian property of a cascade may be lost when the detailed description of a state is given up.

In the present paper, the above methodology is investigated for stochastic processes in general and it is shown how certain choices of sets can be made which preserve the linearity properties, though not necessarily the Markovian properties. It is then shown how in the particular case of a cascade shower these sets can be used to eliminate the need for low energy spatial parameters (those that describe the three-dimensional spreading of the shower at low energies) and also individual energy parameters for all particles. Only three parameters are retained, the number of electrons, the number of photons, and their total energy. These are enough to carry one directly to the number

distribution by means of a linear function of exponentials. No matrices or integral transforms are needed, provided the assumptions of the previous paper are retained and the bremsstrahlung spectrum is idealized in the region of photon energies less than $\epsilon$, $\epsilon > 0$. The elimination of the spatial parameters is done at some expense: The results refer only to the population of particles whose energies are above $\epsilon$. The elimination of the individual energy parameters is dependent on an assumption: For the population of particles whose energies are above , the detailed energy configuration of the particles is known and is independent of $x$ when the number of particles present at $x$ and their total energy are known. $\epsilon$ can be chosen as the critical energy.

# THE CONCEPT OF SETS ENCHAINED BY A STOCHASTIC PROCESS AND ITS USE IN CASCADE SHOWER THEORY

Bayard Rankin
Statistical Laboratory
and
Radiation Laboratory, Department of Physics
University of California, Berkeley, California

July, 1954

## INTRODUCTION

### 1.1 History

A long standing problem in theoretical physics concerns the growth and death of a population of fast electrons and photons as they pass through matter. It was first realized by a few physicists (Nordheim, Carmichael, Bhabha, and Heitler[1]) in about 1934 that successive branching processes could explain some experimental observations of Rossi and Regener, and that these branching processes were just the radiation loss by electrons and materialization by photons that the new quantum theory predicted. As Bhabha and Heitler[1] say, however, "Owing to the ill-founded suspicion in which the (quantum) theory was then held, it did not seem worth while carrying out any calculations. ..... It was believed that the direct measurements of Anderson and Neddermeyer on the energy loss by fast electrons showed that though this energy loss by radiation existed, it was much smaller for energies greater than about $10^8$ ev. than that theoretically predicted, and it was therefore assumed that the present quantum mechanics began to fail for energies greater than about this value." We find by 1937 that more experiments had been made and Bhabha and Heitler[1] and Carlson and Oppenheimer[1.1] had published their first calculations on "The Passage of Fast Electrons and The Theory of Cosmic Showers" and on "Multiplicative Showers." Their papers, while dealing with a number of aspects, inaugurated the long standing problem which is the subject of this paper: to predict the number of shower particles, electrons and photons, that emerge from a piece of absorbing material, assuming some initial conditions.

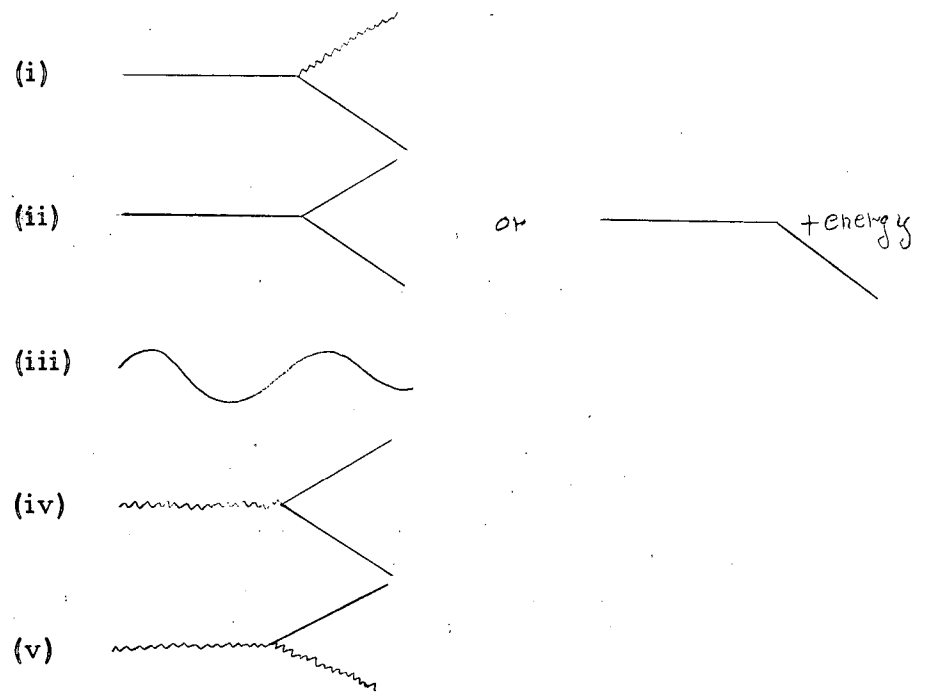
Their work and the work of the many authors* to follow them up to 1950 brought realistic results in terms of the average number of particles in the shower population as a function of the thickness. Valuable results on the second moment were also advanced in this period by Nordsieck, Lamb, Scott, and Uhlenbeck[3, 4]. More work was later done on the second moment by Bhabha and Ramakrishnan[5], Janossy and Messel[6], and Ramakrishnan and Mathews[7]. Since 1950 the efforts by Bhabha[8, 9], Ramakrishnan[10, 11], Janossy[12-14], Scott[15], and Messel and Potts[16, 17] to find the distribution of the number of particles as a function of thickness have been theoretically successful and quantum mechanically correct, but not yet have these efforts, which met extreme complexity, yielded numbers which can be compared with experiment. It is not surprising that the attempts by Bhabha, Heitler[1], Arley[18] and Furry[19] to find the distribution prior to 1950 were based on model building** and imposed severe approximations; nor is it surprising that Robert Wilson[20] has finally resorted to Monte Carlo techniques in order to bypass a general solution and force numbers for some particular cases.

* A comprehensive review of the work on the first moment is given by Bruno Rossi[2].

** "As shown in the next chapter it is possible to evaluate the average numbers of electrons and photons, with given energies, which are produced from a parent electron (or photon), with definite energy, in a certain layer of material. This calculation is, however, already so complicated that it would be quite hopeless to try to evaluate directly the probabilities of finding given numbers of electrons and photons as functions of the primary and secondary energies and of the thickness of the layer. The problem is, therefore, to construct a simplified model of the multiplication process which on the one hand retains most of its characteristic features and on the other hand permits of numerical results being obtained." Niels Arley, loc. cit. p. 88.

## 1.2 The Failure of the Obvious Markovian Description:

The reasons for the mathematical difficulties will be clear after the process for electron-photon development has been described mathematically and a general equation for it has been written down. First, it is necessary to remark that the basic processes* which contribute to the development of a shower are: (i) radiation loss by an electron (an electron branches into an electron and a photon), (ii) collision loss by an electron (an electron branches into two electrons, as in collision with a free electron, or it simply loses a fraction of its energy to excitation of an atom, as in collision with a bound electron), (iii) deflection of an electron in the Coulomb field of a nucleus without the electron losing any energy (the electron path becomes curved at random from multiple scattering), (iv) pair production by a photon (a photon branches into two electrons), and (v) Compton collision by a photon (a photon branches into an electron and a photon). Pictorially, these processes can be illustrated in the following way:




* Throughout this paper it will be assumed that energy is conserved in each of these processes. The assumption is good if the energies involved are large compared to the rest energy of the electron. Trident production by electrons and positron annihilation are ignored as are other unlikely processes.

While only two of these processes, (i) and (iv), are important at high energies where the shower begins, all processes will contribute before the energy of the shower has been completely dissipated through excitation.

Ignoring for the time being process (iii), and ignoring the fact that there is actually an angle of emission at each branch point, one can specify the state, S, of a shower at the point $x \geq 0$ of penetration by giving some particular values to a set of parameters. A suitable set of parameters might be n, m, $\overline{E}$, and $\mu$, representing, respectively, the number of electrons, the number of photons, an n + m dimensional energy vector, $\overline{E} = (E_1, E_2, \ldots, E_{n+m})$, and the ordering of electrons and photons among the energy components of $\overline{E}$. The ordering parameter, $\mu$, can take on any of $2^{n+m}$ permissible values, each value representing one of the $2^{n+m}$ ways in which n+m particles can be distributed among n+m energies. A possible state* of a shower will be written

$$S = (n, m, \overline{E}, \mu), \qquad n = 0, 1, 2, \ldots,$$
$$m = 0, 1, 2, \ldots, \; n+m > 0,$$
$$0 < E_k \leq I, \; k = 1, 2, \ldots, n+m,$$
$$\mu = 1, 2 \ldots, 2^{n+m}.$$

I is the initial energy of the shower at zero penetration.

With these conventions, one can say that a shower is completely described by the random function $\mathcal{S}(x)$, where $\mathcal{S}(x)$ may take its

* Notice that these conventions allow one state to be described by more than one value of E. One could impose the restriction $0 < E_1 \leq E_2 \leq \ldots\ldots \leq E_{n+m} \leq I$. Without this restriction the resulting functions that will be called probabilities will not be normalized but will differ from the desired values by a factor $(n!\, m!)^{-1}$. When actual solutions are obtained, the normalizing factors can be restored and in the meantime the limits of the integrations will be simpler.

values in the space $\Sigma = \{S\}$ of all possible states S. The mathematical problem is to determine the probability that $\mathcal{S}(x)$ will belong to some set of states $\Delta S$, given that $\mathcal{S}(0)$ coincided with some initial state $S_0$:

$$P(\Delta S,\ x \mid S_0,\ 0) = \Pr\left\{ \mathcal{S}(x) \in \Delta S \;\middle|\; \mathcal{S}(0) = S_0 \right\} . \quad *$$

This will be written without the initial condition appearing explicitly when the initial condition is constant and understood:

$$P(\Delta S,\ x) = P(\Delta S,\ x \mid S_0,\ 0) .$$

$\Delta S$ will be understood as the set of states, S, with common n, m, and $\mu$ but with energy components $E_k'$ taking values anywhere in the intervals $(E_k,\ E_k + e_k)$, $e_k \geq 0$, $k = 1, 2, \ldots, n+m$. That is,

$$\Delta S = \left\{ S' : \bar{E}' \in (\bar{E},\ \bar{E} + \bar{e}),\ n' = n,\ m' = m,\ \mu' = \mu \right\} ,$$

where $\bar{e} = (e_1, e_2, \ldots, e_{n+m})$. Since it can be assumed (a) that the processes (i), (ii), (iv), (v) all occur at some discrete points in space, namely the points of impact with the coulomb field of the nucleus or the electron, and (b) that before each process the particle remembers only what its energy is and whether it is an electron or a photon, the following assumptions I and II can be made.

I. <u>The process is purely discontinuous and temporally homogeneous</u>. That is to say, the conditional probability of $\mathcal{S}(x)$ being contained in $\Delta s$, given that at the fixed point y, $0 \leq y < x$, the variable $\mathcal{S}(y)$ coincided with the state S', is a continuous function* of x and y, independently, and has the limiting form:

$$P(\Delta S,\ x \mid S',\ y) = \{1 - a(S')(x-y)\}\,\delta(S',\ \Delta S) + \pi(S' \longrightarrow \Delta S)(x-y) + o(x-y) , \qquad (1.1)$$

$$\delta(S',\ \Delta S) = \begin{cases} 1 & S' \in \Delta S \\ 0 & \text{Otherwise} \end{cases} .$$

* For a precise definition of this function, see Section 2, the paragraph containing equation (2.2).

In general $o(x-y)$ will depend on $S'$ and $\Delta S$. The explicit forms of $a(S')$ and $\pi(S' \longrightarrow \Delta S)$ will be given when they are needed. At this time, it is enough to say that from the mathematical point of view, if $\Delta S$ is fixed, $\pi(S' \longrightarrow \Delta S)$ is measurable with respect to some $\sigma$-field, $\mathcal{B}$, of sets in $\Sigma$. For fixed $S'$, $\pi(S' \longrightarrow \Delta S)$ is a completely additive, non-negative function of sets in $\mathcal{B}$ with

$$\pi(S' \longrightarrow \Sigma) = a(S')$$

and

$$\pi(S' \longrightarrow S') = 0.$$

$\pi(S' \longrightarrow \Delta S)$ is independent of $x$ and $y$. From the physical point of view $P(\Delta S, x \mid S', y)$ is the probability for transition from the state $S'$ to the set of states $\Delta S$ while passing from $y$ to $x$ and $\pi(S' \longrightarrow \Delta S)$ is the cross section for transition from $S'$ to $\Delta S$.

II. The process is Markovian. That is to say, the probability distribution of $\mathfrak{S}(x)$ given $\{\mathfrak{S}(t) : t \leq y\}$ and that of $\mathfrak{S}(x)$ given $\mathfrak{S}(y)$ are the same. In other words, the future state of the system is independent of the history prior to $y$, so long as $\mathfrak{S}(y)$ is known. Symbolically:

$$\text{Dist}\{\mathfrak{S}(x) \mid \mathfrak{S}(y)\} = \text{Dist}\{\mathfrak{S}(x) \mid \mathfrak{S}(t),\ t \leq y\}\ ,\quad x \geq y \geq 0. \tag{1.2}$$

Once one is assured of properties I and II, the most obvious approach to the calculation of $P(\Delta S, x)$ is to write down and solve the diffusion equation which it satisfies. Following Feller's[21] treatment of a purely discontinuous Markoff process, one has:

$$(\partial/\partial x)P(\Delta S, x) = \int_{\Sigma} P(d\Sigma_{S'}, x)\, \pi(S' \longrightarrow \Delta S) - \int_{\Delta S} P(d\Sigma_{S'}, x)\, a(S') \tag{1.3}$$

A general solution that can be extracted from Feller's proofs concerning Equation (1.3) is:

$$P(\Delta S,\ x) = \sum_{N=0}^{\infty} P^{(N)}\ (\Delta S,\ x)\ , \qquad (1.4)$$

where $P^{(N)}\ (\Delta S,\ x)$ are given in terms of an integro-difference equation which can be solved in successive steps in N. The $P^{(N)}\ (\Delta S,\ x)$ can be interpreted as the compound probability that while penetrating x thickness of material the system will experience exactly N transitions and that $\mathcal{S}(x) \in \Delta S$, if it is known that $\mathcal{S}(0) = S_0$. Thus it appears that a formal solution for the cascade shower problem was already available in 1940. Why then have physicists continued to work on the problem for fourteen years? The answer will be clear in the sequel. The three main points of this paper are to show, first, why the Markoff property has been abandoned for much weaker properties whenever practical solutions are attempted and, second, how the Markoff property can be used to obtain a working solution in the form of linear functions of exponentials, and third what new non-Markovian techniques, suggested by present methods, can be used to expedite further the exponential solution. These three points will be made in Section 3 and Section 4, Parts 4.1 and 4.2.

The two reasons that Feller's solution* of equation (1.3) cannot be used in practice are (a) that uncommonly many recurrence steps in N are required and (b) the equation contains no corrections for multiple scattering (process iii) or angular divergence at branching. In reference to (a), the myriad steps are a consequence of the fact that collision loss (process ii) occurs very often with only a slight loss of energy at each transition. Feller's solution, as it stands, counts the number of transitions and, hence, the number of times an electron collides with an electron. In reference to (b), the neglect of the lateral spreading is not serious at high energies, but equation (1.3) follows particles down to zero energy

* The remarks made here do not imply that Feller, himself, seriously considered using his general solution for the cascade problem.

and so follows them into a region where the direction of both electrons and photons are practically random and the motion of electrons is similar to Brownian motion. The fault appears in the parameter set (n, m, E, $\mu$) which is inadequate to describe the state of the shower in the low energy region.

It now becomes apparent that either the original neglect of multiple scatterng (process iii) and angular divergence was serious and more parameters should be enumerated to describe the lateral spreading, or else equation (1.3) must be solved in some approximate manner. Outside these two alternatives, one must reach for an entirely different approach to the problem. The first alternative seems unnecessary, leading as it does to a three dimensional equation. After all, the problem at hand is to find the number of particles as a function of thickness only. Physicists have long felt that the behavior of low energy particles, the residue of a shower, should have little influence on ones mathematical investigation of high energy ones. The second alternative does not seem fruitful for reasons that will be explained immediately. Hence, the following sections will be concerned with fresh techniques that start from different bases than have already been described.

If, as a last resort, equation (1.3) is to make physical sense approximately, the solution must somehow refer only to those particles whose energies are above the region of serious lateral spreading. It could be made to do this in three possible ways. First, sum out all particles whose energies are below some appropriate cutoff energy and interpret the result as the conditional distribution of the number of particles. The condition is that their energies are not below the cutoff. Unfortunately, this still means tracing the transitions of the system through all electron-electron collisions. Alternatively, electron collisions as well as lateral spreading could be disregarded altogether for a high enough cutoff energy, but then no energy would be lost from the shower. The number of particles would increase indefinitely. In this case, the job of summing out all particles whose energies are too low would be a matter of summing out infinitely many of them. Finally, one might wish to say that collision losses occurred infinitely often along an electron's path so

that a continuous deterministic process could replace the discrete stochastic one. Little is gained by this approach. The myriad step recurrence is avoided at the expense of adding terms of the type

$$\beta \frac{\partial P(S, x)}{\partial E_k}$$

to equation (1.3). $\beta$ is the average energy lost per unit path length. The process is not longer purely discontinuous and a practical solution is not known.

The third section of this paper will describe what methods have been used up to the present in order to handle the energy loss and the awkward behavior of the shower particles at low energies. It will be seen that all of these methods are non-Markovian in structure. In the fourth section, part 4.1, a modification of the energy loss process will permit a solution in a convenient number of steps. A new random function will then be introduced. It will be Markovian and purely discontinuous but will describe the shower in such a way that the low energy particles are immediately eliminated from consideration. The serious lateral spreading does not affect the solution and a moderate amount of multiple scattering can be corrected for. The solution is different from Feller's and does not depend on an expansion in terms of the number of transitions, as in (1.4).

In the fourth section, part 4.2, a non-Markovian technique will be introduced to further simplify the solution of 4.1. The two methods of Parts 4.1 and 4.2 can be conveniently catagorized within the general logic that underlies most presently known methods. For this reason, the next section will be devoted to an abstract development of probability theory as it relates to this general logic. The terminology for comparing techniques throughout the rest of this paper will be more in keeping with the ideas introduced in the next section than with Markovian and non-Markovian ideas.

## 2 PROBABILITY THEORY

In this section some ideas will be introduced and some statements relating these ideas will be proved. As presented here, in the concise form that is convenient for reference, the ideas are purely abstract. Upon closer examination they are the probabilistic formulation of methods and techniques that either have been used in other places or will be used later in this paper to solve the cascade shower problem. It is hoped that this section will bring the various approaches to shower theory under a common heading and will lift any useful principles to an accessible place.

In order of presentation, Theorems 1 and 2 establish the basic methods that are unique to this paper. Theorem 3 establishes the essential features of the method that has been most extensively used up to the present. This latter method has not been presented in a similar light before (so far as the author is aware) and, infact, has appeared in such diverse forms that its omnipresence has not been properly appreciated. It is expressed in the form of Theorem 3 to show clearly its relationship to the methods of this paper. Theorem 4 establishes the basis for Janossy's approach. One can see that almost all methods used in the shower theory break up the observation space into subsets and then solve directly for the probability (or probability density) of the corresponding states of the electron-photon system. The difference in techniques is found in the differnce in the choice of subsets and one technique may be more powerful than another because of the judiciousness of this choice: ideally, the unnecessary parameters of the problem are eliminated (or relegated to a convenient place) by partitioning the observation space and the resulting equations become easy to solve.

In subdividing the space for the cascade shower problem, one is motivated by a desire to achieve two things: to be freed from the low energy particles' effect on the high energy shower and the high energy particles' interdependence in energy. The advantage in the first accomplishment was made clear in the introduction and that of the second comes in dismissing all of the rapidly multiplying energy parameters that encumber the equations. In effect, there is a demand to limit one's

"observations" to the set of particles whose energies are not too low and also a demand to "observe" the set of particles whose individual energies are not explicitly specified. Definitions 1 and 2 specify the choices of sets which, when introduced into the shower problem, meet the first and second demands, respectively. The corresponding theorems establish what equations relate the probability weightings of these sets. More will be said about the interpretation of the theorems in Sections 3 and 4.

In the second part of the present section a general solution will be obtained for a large and frequently met class of diffusion equations. It is a solution which can be used in the cascade shower problem, as a later discussion will show.

There will be no existence proofs given, because examples of all the newly defined entities will be encountered in the later sections.

## 2.1 The Concept of Sets Enchained by a Stochastic Process

Let $(\Omega, \mathcal{B}_\Omega, \mathbb{P})$ be a probability space, $(\Gamma, \mathcal{B}_\Gamma)$ a measurable space, and $R^+$ the space of positive real numbers. Let $G_\omega(x)$ be a measurable function defined on $\Omega \times R^+$ to $\Gamma$. The symbols $\Omega$ and $\Gamma$ represent some general spaces made up of points $\omega$ and G, respectively, while $\mathcal{B}_\Omega$ and $\mathcal{B}_\Gamma$ are some $\sigma$-fields of sets in $\Omega$ and $\Gamma$, respectively. $\mathbb{P}$ is a probability measure defined for all sets in $\mathcal{B}_\Omega$. The function $G_\omega(x)$ defined above will be called a random function or a stochastic process and will be written $G(x)$ when no confusion arises. $\Gamma$ will be called the range space of $G(x)$.

Let $\gamma = \{g\}$ be a class of sets in $\Gamma$. That is, each $g = \{G\}$ will be a set of points, G. It should be noticed that if $\gamma$ is measurable in the sense that for all $g \in \gamma$, $g \in \mathcal{B}_\Gamma$, then it is possible to construct the following $\sigma$-field, $\mathcal{B}_\gamma$, that contains the sets $\gamma = \{g\}$: $\mathcal{B}_\gamma$ contains all sets $\Delta g$ of points in $\gamma$ such that the set of all points $G \in \Gamma$ which belong to any g in $\Delta g$ belongs to $\mathcal{B}_\Gamma$. The condition of measurability will be imposed on $\gamma$ in Definition 1 and the symbol "$\Delta g$" will be used in Theorem 1 to represent the set of points G that belong to any g in $\Delta g$ as well as a set that belongs to $\mathcal{B}_\gamma$. That is, in Theorem 1 and elsewhere the statement $G(x) \in \Delta g$, $\Delta g \in \mathcal{B}_\Gamma$ will appear.

It should be further noticed that if $\gamma$ is chosen to be a measurable partition of $\Gamma$, and if the random function $G_\omega(x)$ is given, then it is always possible to treat $(\gamma, \mathcal{B}_\gamma)$ as a new measurable space and to define a new random function $\mathcal{F}_\omega(x)$ as follows: $\mathcal{F}_\omega(x)$ is the measurable function on $\Omega \times R^+$ to $\gamma$ such that $\mathcal{F}_\omega(x) = g$ whenever $G_\omega(x) \in g$. In other words, if $\varphi(G)$ denotes the function on $\Gamma$ to $\gamma$ defined by:

$$\varphi(G) = g \text{ for all } G \in g,$$

then $\mathcal{F}_\omega(x) = \varphi[G_\omega(x)]$. This new random function will appear in the first definition.

In what follows, the concept of conditional probability as introduced by Kolmogorov will be used. It can be summarized within the context of the present ideas as follows: Let $\mathcal{O}(x)$ be any $\omega$-set which is the inverse image of a set $\Delta g \in \mathcal{B}_\gamma$ as mapped by $\mathcal{F}(x)$ or the inverse image

of a set $\Delta g \in \mathcal{B}_{\mathcal{T}}$ as mapped by $\mathcal{G}_\omega(x)$. (Either function $\mathcal{F}_\omega(x)$ or $\mathcal{G}_\omega(x)$ would yield the same inverse map of $\Delta g$.) That is, $\mathcal{O}(x)$ is made up of those elements $\omega$ for which $\mathcal{F}_\omega(x) \in \Delta g$. $\mathcal{O}(x)$ may be different for different values of x. Let $\mathcal{B}$ be any $\sigma$-field of $\omega$-sets. The conditional probability of $\mathcal{O}(x)$ relative to $\mathcal{B}$ is defined at any $\omega$-function, $\Pr\{\mathcal{O}(x) \mid \mathcal{B}\}$, which is either measurable with respect to $\mathcal{B}$ or equal almost everywhere to an $\omega$-function which is, which is integrable, and which satisfies the following equation for all sets $B \in \mathcal{B}$:

$$\int_B \Pr\{\mathcal{O}(x) \mid \mathcal{B}\}\, d\mathbb{P}(\omega) = \mathbb{P}(\mathcal{O}(x) \cap B) . \qquad (2.1)$$

With one further convention, this definition will include the conditional probability of $\mathcal{O}(x)$ relative to $\{\mathcal{F}(t),\ t \in T\}$, $\{\mathcal{G}(x),\ t \in T\}$, or $\{\mathcal{F}(t_1),\ G(t_2),\ t_1 \in T_1,\ t_2 \in T_2\}$, where $T_1$, $T_2$, and T can be any index sets chosen from the positive real axis. It will be agreed that whenever a symbol appears such as "$\Pr\{\mathcal{O}(x) \mid \mathcal{F}(t),\ t \in T\}$" that it is the conditional probability of $\mathcal{O}(x)$ relative to the smallest $\sigma$-field induced by the set of functions $\{\mathcal{F}(t),\ t \in T\}$. That is to say, $\Pr\{\mathcal{O}(x) \mid \mathcal{F}(t),\ t \in T\}$ is a probability conditioned on the smallest $\sigma$-field with respect to which all functions $\mathcal{F}(t)$, $t \in T$, are measurable. If T consists only of one point, y, the conditioning $\sigma$-field of the function $\Pr\{\mathcal{O}(x) \mid \mathcal{F}(y)\}$ will, in general, be different for every different value of y. In the text the symbol "$\mathcal{O}(x)$" will generally be replaced by "$\mathcal{F}(x) \in \Delta g$" or "$\mathcal{G}(x) \in \Delta g$" in order to make explicit which $\omega$-set is being referred to.

In one instance it will be necessary to use the conditional probability of an $\omega$-set which is the inverse image of a set $\Delta G \in \mathcal{B}_{\mathcal{T}}$ as mapped by $\mathcal{G}_\omega(x)$. This conditional probability will be relative to some $\sigma$-field, $\mathcal{B}$, of $\omega$-sets and can be defined just as the above conditional probabilities of $\mathcal{O}(x)$ relative to $\mathcal{B}$.

It is necessary to remark in this place that the following has been proved for a large class of spaces and that the spaces are general enough to serve the practical needs of this paper: Corresponding to the random function $\mathcal{F}(x)$ there is, for all $x > y \geqslant 0$, a transition function $P_1(\Delta g, x \mid g', y)$, such that for fixed x, y, g' it is a non-negative completely additive function of $\Delta g$ with respect to the smallest $\sigma$-field induced by $\mathcal{F}(y)$ and such that for all $\Delta g \in \mathcal{B}_\gamma$ and $x > y \geqslant 0$:

$$P_1(\Delta g, x \mid \mathcal{F}(y), y) = P_r\{\mathcal{F}(x) \in \Delta g \mid \mathcal{F}(y)\}, \qquad (2.2)$$

with probability 1. More generally, there will be transition functions $P(\Delta g, x \mid G, y)$, $P_2(\Delta g, x \mid g, y; g_0, 0)$, $P_3(\Delta g, x \mid G, y; g_0, 0)$, and $P_4(\Delta G, y \mid g, y; g_0, 0)$, with definitions similar to $P_1(\Delta g, x \mid g, y)$, for each of the conditional probabilities $\Pr\{\mathcal{F}(x) \in \Delta g \mid \mathcal{G}(y)\}$, $\Pr\{\mathcal{F}(x) \in \Delta g \mid \mathcal{F}(y), \mathcal{F}(0)\}$, $\Pr\{\mathcal{F}(x) \in \Delta g \mid \mathcal{G}(y), \mathcal{F}(0)\}$, and $\Pr\{\mathcal{G}(x) \in \Delta G \mid \mathcal{F}(y), \mathcal{F}(0)\}$, respectively. All of these transition functions will appear in the following definitions and theorems. In instances when a symbol such as "$P(d\gamma_g, y \mid g_0, 0)$" appears under an integral sign, as in Kolmogorov's notation, it will be implicitly assumed that the integral has the usual Lebesque-Stieljes definition.

The purpose of the first definition is to place conditions on $\gamma$ and $\mathcal{G}(x)$ which permit the random function $\mathcal{F}(x)$ to be Markovian. This is done essentially through condition (iii) which is equivalent to asking that, for each fixed x and y, the conditional probability $\Pr\{\mathcal{F}(x) \in \Delta g \mid \mathcal{G}(y)\}$ is measurable not only with respect to the smallest $\sigma$-field induced by $\mathcal{G}(y)$, but also with respect to that induced by $\mathcal{F}(y)$. Intuitively speaking, condition (iii) extends the Markov property from $\mathcal{G}(x)$ to $\mathcal{F}(x)$ by demanding that any transition of the process $\mathcal{G}(x)$ from points in $g \in \mathcal{B}_\gamma$, at $x_1$, to a set $\Delta g \in \mathcal{B}_\gamma$, at $x_2$, $x_2 > x_1 \geqslant 0$, have the same probability. Though one ultimately wishes to write down an equation for the transitions of the process $\mathcal{G}(x)$ through the sets of $\gamma$, the new function $\mathcal{F}(x)$ becomes a convenient tool to help one do so. One sees in Theorem 1 that the Markovian property of $\mathcal{F}(x)$ (and also its purely discontinuous property) leads to the diffusion equation (2.3) similar to (1.3). Equation (2.3), however familiar in form, is a linear equation for the transitions of the process $\mathcal{G}(x)$ through the sets of $\gamma$. As such it provides a convenient alternative to the usual diffusion equation governing $\mathcal{G}(x)$, because any specific realization of equation (2.3) will involve only enough parameters to specify the sets of $\gamma$ rather than the detailed structure of the states, G.

The similarity between condition (iii) of Definition 1 and the basic condition for a Markov process, (1.2), should not escape the reader. In both cases a conditional probability ($\Pr\{\mathcal{F}(x) \in \Delta g \mid \mathcal{G}(y)\}$ in the first case, $\Pr\{\mathcal{G}(x) \in \Delta g \mid \mathcal{G}(t), t \leqslant y\}$ in the second) is assumed to be measurable with respect to a smaller $\sigma$-field than appears explicitly in the

notation. Intuitively speaking, a certain proportion of the events on which the probabilities are conditioned is superfluous -- in the first case, the event that $\mathcal{G}(t)$ takes a specific value within a set $g \in \mathcal{B}_\gamma$, at $t = y$, and in the second case, the event that $\mathcal{G}(t)$ takes any specific value for $t < y$. These likenesses suggest that a new term such as "enchained" should associate the new condition with the older one, the Markov "chain" condition. More generally speaking, the word "enchained," sometimes qualified, will be used in connection with any condition on the class of sets, $\gamma$, and the process $\mathcal{G}(x)$ which permit one to solve directly for the transitions of the process $\mathcal{G}(x)$ through the sets of $\gamma$.

Definition 1: The class of sets $\gamma = \{g\}$ in the space $\Gamma$ is enchained by the random function $\mathcal{G}(x)$, if the following conditions hold:

(i) $\gamma$ is measurable in the sense that $g \in \mathcal{B}_\Gamma$ for each $g \in \gamma$.

(ii) $\gamma$ forms a partition of $\Gamma$. That is to say, the sets of $\gamma$ are disjoint and their union covers the whole space $\Gamma$.

(iii) If $\mathcal{F}_\omega(x) = \varphi[\mathcal{G}_\omega(x)]$ is the random function defined earlier, then for all x, y, $0 \leqslant y \leqslant x$, and for all sets $\Delta g \in \mathcal{B}_\gamma$:

$$\Pr\left\{\mathcal{F}(x) \in \Delta g \,\middle|\, \mathcal{F}(y)\right\} = \Pr\left\{\mathcal{F}(x) \in \Delta g \,\middle|\, \mathcal{G}(y)\right\},$$

with probability 1.

Theorem 1: (i) If (a) $\mathcal{G}(x)$ is a Markovian random function with range space $\Gamma$ and if (b) $\gamma = \{g\}$ is a class of sets in $\Gamma$ which in enchained by $\mathcal{G}(x)$, then $\mathcal{F}(x)$ (the random function defined in (iii) of Definition 1) is Markovian.

(ii) Suppose (b) in the above statement is true. If (c) $\mathcal{G}(x)$ is purely discontuous and temporally homogeneous, then $\mathcal{F}(x)$ is also.

(iii) If (a), (b), and (c) in the above statements are true, then for each $g_0 \in \gamma$, for all $\Delta g \in \mathcal{B}_\gamma$, and for all $x > 0$, the function $P_1(\Delta g, x \,|\, g_0, 0)$ satisfies the equation:

$$(\partial/\partial x)P_1(\Delta g, x \,|\, g_0, 0) = \int_\gamma P_1(d\gamma_{g'}, x \,|\, g_0, 0)\, \pi'(g' \to \Delta g) - \int_{\Delta g} P_1(d\gamma_{g'}, x \,|\, g_0, 0)\, a'(g') \qquad (2.3)$$

and

$$\lim_{x \to 0} P_1 (\Delta g, x \mid g_0, 0) = \delta (g_0, \Delta g). \tag{2.4}$$

The functions $\pi' (g' \to \Delta g)$ and $a' (g')$, as functions of $g'$, are unique up to sets of measure zero and are defined by:

$$\pi' (g' \to \Delta g) = \pi (G \to \Delta g) - \pi (G \to g') \delta (g', \Delta g), \tag{2.5}$$

$$a' (g') = a (G) - \pi (G \to g'),$$

for any $G \in g'$, where $\pi (G \to \Delta g)$ and $a (G)$ are the functions appearing in the definition of $\mathcal{G}(x)$ as a purely discontinuous and temporally homogeneous process.

Proof: To prove that $\mathcal{F}(x)$ is Markovian, it is enough to show that for all $\Delta g \in \mathcal{B}_\gamma$:

$$\Pr \left\{ \mathcal{G}(x) \in \Delta g \mid \mathcal{F}(y) \right\} = \Pr \left\{ \mathcal{G}(x) \in \Delta g \mid \mathcal{F}(t),\ t \leqslant y \right\},\ t \leqslant y < x. \tag{2.6}$$

From property (iii) of Definition 1, it follows that the left hand side of (2.6) is equal to $\Pr \left\{ \mathcal{G}(x) \in \Delta g \mid \mathcal{G}(y) \right\}$ for all $\Delta g \in \mathcal{B}_\gamma$. From this and the Markov property of $\mathcal{G}(y)$, it follows that:

$$\Pr \left\{ \mathcal{G}(x) \in \Delta g \mid \mathcal{F}(y) \right\} = \Pr \left\{ \mathcal{G}(x) \in \Delta g \mid \mathcal{G}(t),\ t \leqslant y \right\}. \tag{2.7}$$

Equation (2.7) implies (2.6) because the smallest $\sigma$-fields induced by $\mathcal{F}(y)$, $\left\{ \mathcal{F}(t),\ t \leqslant y \right\}$, and $\left\{ \mathcal{G}(t),\ t \leqslant y \right\}$ have the relation:

$$\mathcal{B}_{\mathcal{F}(y)} \subset \mathcal{B}_{\left\{ \mathcal{F}(t),\ t \leqslant y \right\}} \subset \mathcal{B}_{\left\{ \mathcal{G}(t),\ t \leqslant y \right\}}.$$

If $\mathcal{G}(y)$ is purely discontinuous and temporally homogeneous, then the transition function corresponding to $\mathcal{G}(y)$ is continuous in x and y, independently, and has the limiting form:*

$$P(\Delta g,\ x \mid G,\ y) = \left\{1 - \alpha(G)\,(x - y)\right\}\ \delta(G,\ \Delta g) \qquad (2.8)$$

$$+\ \pi\,(G \to \Delta g)\,(x - y) + o(x - y),$$

where $\alpha(G)$ and $\pi\,(G \to \Delta g)$ are defined in Section 1.2. From the definition of $\gamma = \{g\}$ as a set enchained by $\mathcal{G}(x)$, it follows that for all sets $\Delta g \in \mathfrak{B}_\gamma$:

$$P_1(\Delta g,\ x \mid \varphi[\mathcal{G}(y)],\ y) = P(\Delta g,\ x \mid \mathcal{G}(y),\ y)$$

with probability 1. Hence, it is seen that for all $\Delta g \in \mathfrak{B}_\gamma$ and for almost all $G \in \Gamma$:

$$P_1(\Delta g,\ x \mid g',\ y) = P(\Delta g,\ x \mid G,\ y),\ G \in g',\ g' \in \gamma.$$

By "almost all" is meant up to a set $N(y) \in \mathfrak{B}_\Gamma$ such that $\mathbb{P}(G(y) \in N(y)) = 0$. Thus,

$$P_1(\Delta g,\ x \mid g',\ y) = \left\{1 - \alpha(G)\,(x - y)\right\}\ \delta(G,\ \Delta g) \qquad (2.9)$$

$$+\ \pi\,(G \to \Delta g)\,(x - y) + o(x - y),\ G \in g',\ g' \in \gamma.$$

This equality can also be written:

$$P_1(\Delta g,\ x \mid g',\ y) = \left\{1 - \left[\alpha(G) - \pi(G \to g')\right](x - y)\right\}\ \delta(g',\ \Delta g) \qquad (2.10)$$

$$+\ \left[\pi(G \to \Delta g) - \pi(G \to g')\ \delta(g',\ \Delta g)\right](x - y) + o(x - y),\ G \in g',\ g' \in \gamma.$$

* It is implicitly assumed through Feller's[21] paper and the present one that the term $o(x - y)$, appearing in the limiting form (2.8), is such that $o(x - y)/(x - y) \to 0$ as $(x - y) \to 0$, uniformly in G for all $G \in \Gamma$.

where the convection,

$$\delta(g', \Delta g) = \begin{cases} 1 & g' \in \Delta g \\ 0 & \text{Otherwise} \end{cases},$$

is meaningful, since the sets of $\gamma$ are disjoint. The right hand side of (2.10) is obviously constant for all $G \in g'$, $g' \in \gamma$, but it can also be shown, by using the properties of the functions $\alpha(G)$ and $\pi(G \to \Delta g)$, that this also true for the terms in square brackets. In other words, the terms in square brackets are really functions of $g'$ and $\Delta g$ only. Since it is further true that they satisfy the properties of $\alpha(G)$ and $\pi(G \to \Delta g)$ respectively, $P_1(\Delta g, x \mid g', y)$ has the limiting form of a transition function for a purely discontinuous and temporally homogeneous process. One can write:

$$P_1(\Delta g, x \mid g', y) = \{1 - \alpha'(g')(x - y)\}\ \delta(g', \Delta g) \qquad (2.11)$$

$$+ \pi'(g' \to \Delta g)(x - y) + o(x - y),$$

where the functions $\alpha'(g')$ and $\pi'(g' \to \Delta g)$ have the definitions (2.5). This establishes statement (ii) of the theorem and gives an explicit form for the transition function of $\mathcal{F}(x)$ in terms of the functions $\alpha(G)$ and $\pi(G \to \Delta g)$ that correspond to $\mathcal{G}(x)$.

To complete the proof of the theorem one can write down the Chapman-Kolmogorov equation that follows from the Markov property of $\mathcal{F}(x)$. For each $g_0 \in \gamma$, for all $x$, $y$, $0 \leqslant y < x$, and for all $\Delta g \in \mathcal{B}_\gamma$:

$$P_1(\Delta g, x \mid g_0, 0) = \int_\gamma P_1(d\gamma_{g'}, y \mid g_0, 0)\, P_1(\Delta g, x \mid g', y). \qquad (2.12)$$

This equation, together with the form (2.11) yields:

$$\{P_1(\Delta g, x \mid g_0, 0) - P_1(\Delta g, y \mid g_0, 0)\}/(x - y) \qquad (2.13)$$

$$= \int_\gamma P_1(d\gamma_{g'}, y \mid g_0, 0)\, \pi'(g' \to \Delta g) - \int_{\Delta g} P_1(d\gamma_{g'}, y \mid g_0, 0)\, \alpha'(g') + \frac{o(x - y)}{(x - y)}$$

The limiting form of (2.13) as $y \to x$ is (2.3).

The continuity condition of a purely discontinuous process implies the boundary condition (2.4).

Definition 2 is inspired by the desire to achieve under different conditions (preferably weaker) the same ultimate goal that was achieved in Theorem 1. The goal was a linear equation for determining the transitions of the Markov process $\mathcal{G}(x)$ through a class of sets $\gamma$ (that is, for determining the function $P_1(\Delta g,\ x \mid g_0,\ 0)$). The reason for Theorem 1 was to simplify the parameter problem. Its shortcomings are these: It might not always be possible to eliminate unnecessary parameters by choosing a class of sets $\gamma$ according to the conditions given in the hypothesis of the theorem. The conditions are in fact so strong that one is really brought back to the Markov property. Before stating abstractly the new conditions as they appear in Definition 2, it will now be shown how the transitions of a Markov process, $\mathcal{G}(x)$, through a class of sets $\gamma$ can be determined from a linear equation even though it is not possible to describe the transitions by means of a Markov process (such as $\mathcal{F}(x)$ of Definition 1).

Assume that the class of sets $\gamma$ is a measurable partition of $\Gamma$. According to the usual rules governing conditional probabilities, one can write down immediately the following equations for the transition functions defined earlier: (In these equations $x > y \geqslant 0$.)

$$P_1(\Delta g,\ x \mid g_0,\ 0) = \int_{\gamma} P_1(d\gamma_{g'},\ y \mid g_0,\ 0)\, P_2(\Delta g,\ x \mid g',\ y;\ g_0,\ 0), \tag{2.14}$$

$$P_2(\Delta g,\ x \mid g',\ y;\ g_0,\ 0) = \int_{\mathcal{G}'} P_3(\Delta g,\ x \mid G,\ y;\ g_0,\ 0)\, P_4(d\,\Gamma_G,\ y \mid g',\ y;\ g_0,\ 0) \tag{2.15}$$

These equations hold in general for any process, $\mathcal{G}(x)$, so long as the transition functions are defined. If it is assumed that the process $\mathcal{G}(x)$ is Markovian, the function $P_3(\Delta g,\ x \mid G,\ y;\ g_0,\ 0)$ can be written $P(\Delta g,\ x \mid G,\ y)$. The two equations can be combined and if $\mathcal{G}(x)$ is Markovian then:

$$P_1(\Delta g,\ x \mid g_0,\ 0) = \int_{\gamma} P_1(d\gamma_{g'},\ y \mid g_0,\ 0) \left[ \int_{\mathcal{G}'} P(\Delta g,\ x \mid G,\ y)\, P_4(d\,\Gamma_G,\ y \mid g',\ y;\ g_0,\ 0) \right] \tag{2.16}$$

This equation is immediately suggestive of the Chapman-Kolmogorov equation, (2.12), the difference being that a weighted transition function, the bracketed term, here replaces the usual term, $P_1(\Delta g,\ x \mid g',\ y)$. In order to obtain a diffusion equation from it, it is possible to proceed along the same steps that would be used if the more familiar equation were at hand. The basic process, $\mathcal{G}(x)$, must, of course, be purely discontinuous and temporally homogeneous. As in the proof of Theorem 1, one would invoke the asymptotic form (2.8) for the transition function $P(\Delta g,\ x \mid G,\ y)$, subtract off the proper terms as in (2.13), and pass to the limit as $y \to x$. The result would be:

$$(\partial/\partial x)\, P_1(\Delta g,\ x \mid g_0, 0) = \int_{\gamma} P_1(d\gamma_{g'},\ X \mid g_0,\ 0)\, \bar{\pi}(g' \to \Delta g,\ x) \tag{2.17}$$

$$- \int_{\Delta g} P_1(d\gamma_{g'},\ x \mid g_0, 0)\, \bar{a}(g',\ x),$$

where

$$\bar{\pi}(g' \to \Delta g,\ X) = \int_{\mathcal{G}'} \pi(G \to \Delta g)\, P_4(d\,\Gamma_G,\ x \mid g',\ x;\ g_0,\ 0),$$

$$\bar{a}(g',\ X) = \int_{\mathcal{G}'} a(G)\, P_4(d\,\Gamma_G,\ x \mid g',\ x;\ g_0,\ 0). \tag{2.18}$$

This is a linear equation for the transitions of the function $\mathcal{G}(x)$ through a class of sets $\gamma$. No assumptions have been placed on the class $\gamma$ relative to $\mathcal{G}(x)$. The only basic assumptions are: (i) $\gamma$ is a measurable partition and (ii) $\mathcal{G}(x)$ is a purely discontinuous and temporally homogeneous Markov process.

From the practical point of view, equation (2.17) would be simpler if the weighting function, $P_4(\Delta G,\ x \mid g',\ x;\ g_0,\ 0)$, were independent of $X$. Moreover, the equation can actually be solved only if the weighting function is known a peiori. Therefore, there will soon be imposed, in Definition 2, the condition that the wrighting function is independent of x. It must be implicitly assumed that the weighting function is known a priori, though such a statement is devoid of mathematical significance.

Compare new the strength of the conditions in Definitions 1 and 2. Both are conditions on equation (2.16). Both are designed to make that equation a practical one -- an integral equation from which a solvable diffusion equation can be derived. In the first case, $P(\Delta g, x \mid G, y)$ is assumed to be a measurable function of G with respect to the $\sigma$-field $\mathcal{B}_\gamma$. In the second case, $P_4(\Delta G, x \mid g', x; g_0, 0)$ is assumed to be independent of x.

Definition 2: A class of sets $\gamma = \{g\}$ in $\Gamma$ is conditionally enchained by the random function $G(x)$, if the following three properties hold:

(i) $\gamma$ is measurable in the sense of (i), Definition 1.

(ii) The sets of $\gamma$ represent a partition of $\Gamma$.

(iii) The function $P_4(\Delta G, x \mid g, x; g_0, 0)$ is independent of x for all $x \geqslant 0$:

$$P_4(\Delta G, x \mid g, x; g_0, 0) = P_4(\Delta G \mid g, g_0), \text{ say.} \tag{2.19}$$

Theorem 2: Let $G(x)$ be a purely discontinuous, temporally homogeneous Markov process with range space $\Gamma$. Suppose that there exists a class of sets $\gamma = \{g\}$ in $\Gamma$ which is conditionally enchained by $G(x)$. It follows that for all $\Delta g \in \mathcal{B}_\gamma$, $g_0 \in \gamma$, and $x > 0$, the function $P_1(\Delta g, x \mid g_0, 0)$ satisfies the equation:

$$(\partial/\partial x) P_1(\Delta g, x \mid g_0, 0) = \int_\gamma P_1(d\gamma_{g'}, x \mid g_0, 0) \overset{o}{\pi}(g' \to \Delta g) - \int_{\Delta g} P_1(d\gamma_{g'}, x \mid g_0, 0) \overset{o}{a}(g') \tag{2.20}$$

where

$$\lim_{x \to 0} P_1(\Delta g, x \mid g_0, 0) = \delta(g_0, \Delta g). \tag{2.21}$$

The functions $\overset{o}{\pi}(g' \to \Delta g)$, $\overset{o}{a}(g')$ are determined by the integrals:

$$\begin{aligned} \overset{o}{\pi}(g' \to \Delta g) &= \int_{g'} \pi(G \to \Delta g) P_4(d\Gamma_G \mid g', g_0) \\ \overset{o}{a}(g') &= \int_{g'} a(G) P_4(d\Gamma_G \mid g', g_0). \end{aligned} \tag{2.22}$$

$\pi(G \rightarrow \Delta g)$ and $a(G)$ are the functions which appear in the definition of a purely discontinuous and temporally homogeneous process as applied to $G(x)$ and $P_4(\Delta G \mid g', g_0)$ is the function defined in (iii), Definition2.

The proof of the theorem should be obvious from the discussion preceeding it.

Definition 2 was an attempt to generalize the idea of enchainment. In fact, it was shown in the discussion preceeding it, that the generalization was nearly complete: No restrictions really had to be imposed on the class of sets $\gamma$ beyond the properties of a measurable partition. The further restriction that was imposed, namely, that the weighting function be independent of x, was done to facilitate the later solution. Now a generalization in a different direction will be tried. It will no longer be required in Definition 3 that $\gamma$ should form a partition of the space $\Gamma$. Instead, a new set of conditions will be imposed. These new conditions will define the concept of local enchainment.

The idea of local enchainment comes from observing how one passes from the general Chapman-Kolmogorov equation to a diffusion equation. One does it by passing to a limit in the independent variable that corresponds to time. The resulting diffusion equation refers only to changes that occur in the stochistic process in differentially small increments of time. This observation might suggest that the linear properties of a diffusion equation can be recovered under more general conditions than go into the Chapman-Kolmogorov equation. The following definition exploits this idea and makes it useful in solving for the transitions of a process $G(x)$ through a class of sets $\gamma$. One asks in Definition 3 that some additive and disjointness properties hold in the limiting sense of Properties (v) and (vi). As the proof of Theorem 4 shows, one is then allowed to write down equation (2.33) which suggests a "local" Chapman-Kolmogorov equation for sets. Property (iv), a "local" Markov property allows one to pass from equation (2.33) to (2.30) in such a way that the limiting transition functions $\pi(g' \rightarrow g)$ and $a(g)$ are independent of the initial state of the system. Thus, the greater freedom allowed by imposing only local conditions on the process $G(x)$ permits a linear equation, similar to (1.3), to hold for a new class of sets. The same advantages against the parametric obsticles might be reached now that

were reached with the classes of sets of Definitions 1 and 2. It should be noticed, however, that a partition of the space $\Gamma$ is not required for local enchainment and consequently the equation (2.30) cannot be interpreted in terms of a process such as $\mathcal{F}(x)$.

The following concept of sets locally enchained need not be restricted to finite class of sets $\gamma$. In fact, in examples from cascade shower theory the locally enchained sets are not finite. The reason for avoiding the generalization in this paper is precisely to keep the definitions and theorems as close to intuition as possible so that the basic principles involved will be seen. A second reason is that the added complication would not be justified, since the methods of this paper are not based on local enchainment.

Definition 3: A finite class of sets $\gamma = \{g\}$ in $\Gamma$ is locally enchained by the random function $\mathcal{G}(x)$, if the following six properties hold:

(i) $\gamma$ is measurable in the sense of (i), Definition 1.

(ii) For all $G \in \Gamma$ there is at least one set $g \in \gamma$ such that $G \in g$.

(iii) For each $g_k \in \gamma$ there is some $g_\ell \in \gamma$ such that (a) $\Pr\{\mathcal{G}(y) \in g_\ell\} \neq 0$; (b) $\Pr\{\mathcal{G}(x) \in g_k \mid \mathcal{G}(y) \in g_\ell\}$, $x > y \geqslant 0$, is a continuous function of x and y, independently, and (c) for $x > y \geqslant 0$:

$$\Pr\{\mathcal{G}(x) \in g_k \mid \mathcal{G}(y) \in g_\ell\} = \pi(g_\ell \longrightarrow g_k)(x - y) + o(x - y), \qquad (2.26)$$

where $\pi(g_\ell \longrightarrow g_k)$ is a positive function of $g_\ell$ and $g_k$ and is independent of x and y.

For each $g_\ell \in \gamma$ such that (a) holds, there is some $g_k \in \gamma$ such that (b) and (c) hold.

(For each $g_k \in \gamma$ let $C(g_k)$ be the set of all $g = g_\ell$, for which (a), (b), and (c) hold and let $\overline{C}(g_k)$ be the same set with $g_k$ added.)

(iv) For each $g_k \in \gamma$ and for all $g \in C(g_k)$ and for any set, $\{\mathcal{G}(t) \in g'\}$, of events of the type $\mathcal{G}(t) \in g'$, $0 \leqslant t < y$, $g' \in \gamma$, such that $\Pr\{\mathcal{G}(y) \in g, \{\mathcal{G}(t) \in g'\}\} \neq 0$,

$$\frac{\Pr\{\mathcal{G}(x) \in g_k \mid \mathcal{G}(y) \in g\}}{\Pr\{\mathcal{G}(x) \in g_k \mid \mathcal{G}(y) \in g, \{\mathcal{G}(t) \in g'\}\}} = 1 + o(x - y), \quad 0 < y < x. \qquad (2.27)$$

(v) For each $g_k \in \gamma$:

$$\Pr\left\{ \mathcal{G}(y) \in \overline{C}(g_k) \,\middle|\, \mathcal{G}(x) \in g_k \right\} = 1 + o(x - y), \quad 0 < y < x. \tag{2.28}$$

(vi) For each $g_k \in \gamma$ and for all $g, g' \in \overline{C}(g_k)$, $g \neq g'$:

$$\Pr\left\{ \mathcal{G}(x) \in g \bigcap g' \right\} = 0, \quad x > 0 \tag{2.29}$$

Theorem 3: Let $\mathcal{G}(x)$, $x \geqslant 0$, be a random function with range space $\Gamma = \{G\}$. Suppose that there exists a finite class of sets $\gamma = \{g\}$ in $\Gamma$ which is locally enchained by $\mathcal{G}(x)$. It follows that for each $g_0 \in \gamma$ for which $\Pr\left\{\mathcal{G}(0) \in g_0\right\} \neq 0$, for all $g \in \gamma$, and for all $x > 0$, the function $P_5(g, x \,|\, g_0, 0) = \Pr\left\{ \mathcal{G}(x) \in g \,\middle|\, \mathcal{G}(0) \in g_0 \right\}$ satisfies the following equation:

$$(\partial/\partial x + \alpha(g))\, P_5(g, x \,|\, g_0, 0) = \sum_{g' \in C(g)} P_5(g', x \,|\, g_0, 0)\, \pi(g' \rightarrow g) \tag{2.30}$$

where, if $0 < y < x$,

$$\begin{aligned} \pi(g' \rightarrow g) &= \lim_{(x-y)\to 0} \Pr\left\{ \mathcal{G}(x) \in g' \,\middle|\, \mathcal{G}(y) \in g,\ \mathcal{G}(0) \in g_0 \right\} / (x - y) \\ \alpha(g) &= \lim_{(x-y)\to 0} \Pr\left\{ \mathcal{G}(x) \in \overline{g} \,\middle|\, \mathcal{G}(y) \in g,\ \mathcal{G}(0) \in g_0 \right\} / (x - y) \end{aligned} \tag{2.31}$$

and both $\pi(g' \rightarrow g)$ and $\alpha(g)$ are independent of $x$, $y$, and $g_0$. $\overline{g} = \Gamma - g$. In the limit as $x \rightarrow 0$,

$$P(g, x \,|\, g_0, 0) = \begin{cases} 1 & g = g_o \\ 0 & \text{Otherwise} \end{cases}$$

Proof: Assume that there is at least one $g_0 \in \gamma$ for which $P_r\left\{G(0) \in g_0\right\} \neq 0$. If not, the theorem is trivially true. From property (v), Definition 4, one obtains for each $g \in \gamma$, and for all $x$, $y$, $0 < y < x$:

$$P_r\left\{\mathcal{G}(x) \in g \,\middle|\, \mathcal{G}(0) \in g_0\right\} = \Pr\left\{\mathcal{G}(x) \in g,\ \mathcal{G}(y) \in \overline{C}(g) \,\middle|\, \mathcal{G}(0) \in g_0\right\} + o(x - y) \tag{2.32}$$

Using property (vi), one can write this equation as:

$$\Pr\left\{G(x) \in g \,\middle|\, G(0) \in g_0\right\} = \sum_{g' \in \bar{C}(g)} \Pr\left\{G(x) \in g,\ G(y) \in g' \,\middle|\, G(0) \in g_0\right\} + o(x - y) \tag{2.33}$$

$$= \sum_{g' \in \bar{C}(g)} \Pr\left\{G(y) \in g' \,\middle|\, G(0) \in g_0\right\} \Pr\left\{G(x) \in g \,\middle|\, G(y) \in g',\ G(0) \in g_0\right\} + o(x - y).$$

The first part of property (iii) and property (iv) allows that for each $g \in \gamma$ and for all $g' \in C(g)$:

$$\Pr\left\{G(x) \in g \,\middle|\, G(y) \in g',\ G(0) \in g_0\right\} = \pi(g' \longrightarrow g)(x - y) + o(x - y). \tag{2.34}$$

It is always true that for each $g \in \gamma$:

$$\Pr\left\{G(x) \in g \,\middle|\, G(y) \in g,\ G(0) \in g_0\right\} = 1 - \Pr\left\{G(x) \in \bar{g} \,\middle|\, G(y) \in g,\ G(0) \in g_0\right\}. \tag{2.35}$$

The second part of property (iii) and property (iv) allows that for each $g \in \gamma$ there is some $g_k$ such that:

$$\Pr\left\{G(x) \in g_k \,\middle|\, G(y) \in g,\ G(0) \in g_0\right\} = \pi(g \longrightarrow g_k)(x - y) + o(x - y).$$

This $g_k$ cannot be g, because of the continuity condition (iii, b). Hence, there is some function $a(g)$ which is non-negative, finite, and independent of x, y, and $g_0$ such that

$$\Pr\left\{G(x) \in \bar{g} \,\middle|\, G(y) \in g,\ G(0) \in g_0\right\} = a(g)(x - y) + o(x - y), \tag{2.36}$$

where $\bar{g} = \Gamma - g$.

By combining (2.35) and (2.36) and putting the result into (2.33) with (2.34), one obtains an equation which passes into (2.30) as $y \to x$. The same steps were followed in the proofs of Theorems 1 and 2. The final boundary condition of the theorem follows from the continuity property (iii, b).

The essential assumption of Definition 4 might appear somewhat artificial at first sight and can be justified only in terms of its applications. The objectives associated with enchained sets should be clear now from

the past ideas, and those of multiply enchained sets are nothing new. Condition (iii) of Definition 4 that is responsible for equation (2.38) is somewhat different from previous ones, however, and a diagram might help. This condition says that the probability for a transition of the type represented in Figure A by a double arrow is equal to a sum of probabilities, each term in the sum being a product corresponding to transitions of the type represented by a single arrow.

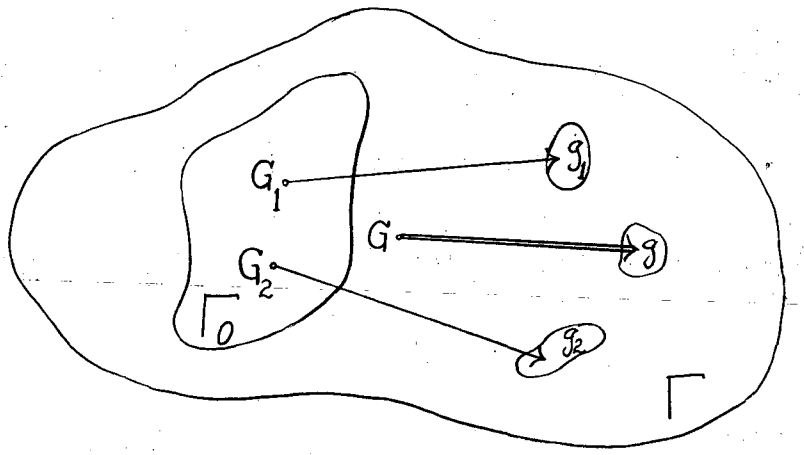



Fig. A

The advantage of Property (iii) is that it leads to equation (2.38) that involves only transitions into sets $g \in \gamma$ from points, G, of a relatively limited set $\Gamma_0$. One should refer to Section 3, Part 3.3, for a practical example.

Definition 4: A class $\gamma = \{g\}$ of sets in $\Gamma$ is multiply enchained by the random function $G_0(x)$, if the following three properties hold:

(i) $\gamma$ is measurable in the sense of (i), Definition 1.

(ii) $\gamma = \{g\}$ represents a denumerable partition of $\Gamma$ in the sense that the set $\gamma = \{g\}$ is denumerable and property (ii) of Definition 1 holds.

(iii) There exists a set of points $\Gamma_0 = \{G\}$ properly contained in $\Gamma$, $\Gamma_0 \subset \Gamma$, $\Gamma_0 \neq \Gamma$, such that the following holds: For each $g \in \gamma$, for all x, y, $0 < y < x$, and for all $G \in \Gamma - \Gamma_0$, there is a finite set of points in $\Gamma_0$, $G_1(G)$, $G_2(G)$, ..., $G_k(G)$, dependent on G (where k also depends on G) such that:

$$P(g, x \mid G, y) = \sum P(g_1, x \mid G_1(G), y)\, P(g_2, x \mid G_2(G), y) \ldots P(g_k, x \mid G_k(G), y) \tag{2.37}$$

where the sum is extended over sets of the form $(g_1, g_2, \ldots, g_k)$, each set containing k sets, $g_j \in \gamma$, $j=1, 2, \ldots, k$, which belong to $\gamma$. Each set $(g_1, g_2, \ldots, g_k)$ is dependent on g.

Theorem 4: Let $G(x)$, $x \geqslant 0$, be a purely discontinuous, temporally homogeneous Markov process whose range space is $\Gamma$. Suppose that there exists a set $\gamma = \{g\}$ of sets in $\Gamma$ which is multiply enchained by $G(x)$. If so, it follows that for each $G_0 \in \Gamma_0$, for all $x > 0$, and for all $g \in \gamma$, the transition function $P(g, x \mid G, 0)$, where $G \in \Gamma_0$, satisfies the following equation:

$$\begin{aligned} &(\partial/\partial x + \alpha(G_0))\, P(g, x \mid G_0, 0) \\ &= \int_{\Gamma} P(g_1, x \mid G_1(G), 0) \ldots P(g_k, x \mid G_k(G), 0)\, \pi(G_0 \longrightarrow d\Gamma_G) \end{aligned} \tag{2.38}$$

where $\pi(G_0 \longrightarrow \Delta G)$ and $\alpha(G_0)$ are the functions appearing in the definition of a purely discontinuous and templorally homogeneous process as applied to $G(x)$, and the sum in the integrand is defined in (iii), Definition 5. The boundary conditions are:

$$\lim_{x \to 0} P(g, x \mid G_0, 0) = \delta(G_0, g).$$

Proof: From the Markovian property of $G(x)$ one has the Chapman-Kolmogorov equation:

$$P(g, x \mid G_0, 0) = \int_{\Gamma} P(d\Gamma_{G'}, y \mid G_0, 0)\, P(g, x \mid G', y).$$

By introducing the form (2.8) (with $\Delta g$ replaced by $\Delta G$) for the first term in the integrand, one has an equation which, after the operations familiar in previous proofs yields (2.38). The only step that is new is that (2.38) is finally obtained by letting $y \longrightarrow 0$ rather than $y \longrightarrow x$.

## 2.2 A General Solution of a Class of Diffusion Equations

Let $\sum$ be a space of points S, $\sum = \{S\}$. Suppose that (i) $a(S)$ is a non-negative, finite, real function defined on $\sum$; (ii) for fixed $S \in \sum$, $\pi(S' \longrightarrow S)$ is a non-negative, bounded, real function of S' defined on $\sum$; (iii) for fixed $S' \in \sum$, the same properties hold for $\pi(S' \longrightarrow S)$ as a function of S; (iv) and finally, suppose that the same is also true of $f(S, x)$ when x is fixed, $x \geqslant 0$. Assume that for each fixed $S \in \sum$, $f(S, x)$ is a continuous function of x for $x \geqslant 0$. It will be said that S can be reached in one step from $S^{(1)}$ if $(S^{(1)} \longrightarrow S) > 0$; S can be reached (in one or more steps) from $S^{(1)}$ if there exists a finite sequence $S^{(1)}, S^{(2)}, \ldots, S^{(n)} = S$ such that $S^{(k+1)}$ can be reached in one step from $S^{(k)}$, $k = 1, 2, \ldots, n - 1$.

Suppose that for every $S \in \sum$ there is a finite sequence of points $\{S_k(S), k = 1, 2, \ldots, N(S)\}$, $S_k(S) \in \sum$ (with the same first element $S_1(S) = S_1$, for all $S \in \sum$, and with the last element $S_{N(S)}(S) = S$) such that the following is true:

(i) Every point of $\sum$ from which $S_j(S)$, $j = 1, 2, \ldots, N(S)$, can be reached is one of the points $S_i(S)$, $i < j$. In particular, $S_1$ cannot be reached from any point of $\sum$.

(ii) For every $j = 2, 3, \ldots, n$, there exists $i < j$ such that $S_j(S)$ can be reached from $S_i(S)$; and for every $j = 1, 2, \ldots, N(S) - 1$, there exists $k > j$ such that $S_k(S)$ can be reached from $S_j(S)$.

Imagine that a system is moving through a space, $\sum$, of states, S, according to some law, and that if $\pi(S' \longrightarrow S) > 0$ then the step from S' to S is allowed. A sequence satisfying properties (i) and (ii) might be called a "route" to S from $S_1$ determined by the function $\pi(S' \longrightarrow S)$. Without saying just which steps are takan by the system, it has been required that every point $S \in \sum$ can be reached by some "route" from $S_1$. Since any point $S_j(S)$ on the route to S might be reached by a step from any number of other points $S' \in \sum$, it is clear that every state on the "route" to S need not be utilized in getting to S. It is required, however, that every state, $S_j(S)$, on the route to S can be reached only from below, so to speak, or only from some part of the route already traversed. It is also required that the last point, S, can be reached from any point $S_j(S)$ on the "route".

In the following theorem the symbol " $\sum_{r}^{s}$ paths $\prod_{(i\,j)}$ " will be used. It is defined as follows. Consider the sequence of integers from r to s. A "path" will be defined as a subsequence of ascending integers chosen from r to s and including r and s. For instance, $\{n_{\ell,q},\ q = 1, 2, \ldots, N(\ell, r, s)\}$, where $n_{\ell,q} < n_{\ell,q+1}$ might be a subsequence chosen from $\{r, r+1, \ldots, s\}$. In this $\ell^{th}$ subsequence there are $N(\ell, r, s)$ integers, where $N(\ell, r, s)$ may change with $\ell$, r and s. Let the number of such sequences be $N(r \to s)$. Now for any function $Q(u, v)$ define:

$$\sum_{r}^{s} \text{paths} \prod_{(i\,j)} Q(i, j) = \sum_{\ell=1}^{N(r\to s)} \prod_{i=1}^{N(\ell,r,s)-1} Q(n_{\ell,i}, n_{\ell,i+1}). \tag{2.39}$$

Thus, " $\sum_{r}^{s}$ paths $\prod_{(i\,j)}$ " means sum over all "paths" from r to s and, corresponding to each path, take the produce of all terms, $Q(n_{\ell,i}, n_{\ell,i+1})$, involving two successive integers of the path. As an example:

$$\sum_{1}^{3} \text{paths} \prod_{(i\,j)} Q(i, j) = Q(1, 3) + Q(1, 2)\, Q(2, 3),$$

$$\sum_{1}^{4} \text{paths} \prod_{(i\,j)} Q(i, j) = Q(1, 4) + Q(1, 2)\, Q(2, 4)$$

$$+ Q(1, 3)\, Q(3, 4) + Q(1, 2)\, Q(2, 3)\, Q(3, 4).$$

Theorem 5: Let the space $\sum$ and the functions $\pi(S' \to S)$, $a(S)$, $f(S, x)$ satisfy the conditions outlined above. Suppose that for all $S \in \sum$ and all $x > 0$, the function $f(S, x)$ satisfies the diffusion equation:

$$(\partial/\partial x + a(S))\, f(S, x) = \sum_{S' \in \sum} f(S', x)\, \pi(S' \to S) \tag{2.40}$$

with boundary conditions:

$$\lim_{x \to 0} f(S, x) = \begin{cases} 1 & S = S_1 \\ 0 & \text{Otherwise} \end{cases}$$

For any fixed $S = S^* \in \sum$ let a "route" from $S_1$ to $S^*$,

This condition can be weakened by allowing $a(S_k) = a(S_\ell)$ in cases where the state $S_k$ cannot be reached from the state $S_\ell$. Such a modification in the statement of the theorem would not affect the basic logic of the proof, but would be responsible for notational complications.

$\{S_k(S^*);\ k = 1, 2, \ldots, N(S^*)\}$, be written $\{S_k,\ k = 1, 2, \ldots, N\}$. Suppose* that $a(S_k) \neq a(S_\ell)$ for $k \neq \ell$. It follows that

$$f(S_k, x) = \sum_{t=1}^{k} \beta_t(S_k)\, e^{-x a(S_t)},\quad k = 1, 2, \ldots, N, \tag{2.41}$$

where the coefficients $\beta_t(S_k)$ are determined by:

$$\beta_1(S_1) = 1$$

$$\beta_t(S_k) = \beta_t(S_t) \sum_{t}^{k} \text{paths} \prod_{(i\,j)} \frac{\pi(S_i \to S_j)}{a(S_j) - a(S_t)},\quad \begin{matrix} k = 2, 3, \ldots, N, \\ t = 1, 2, \ldots, k-1 \end{matrix} \tag{2.42}$$

$$\beta_k(S_k) = -\sum_{t=1}^{k-1} \beta_t(S_k),\quad k = 2, 3, \ldots, N.$$

Proof: In the definition of a "route" in terms of the two properties given above, there was freedom enough to allow more than one route from $S_1$ to any $S \in \sum$. For example, if $N(S) = 4$ and $S_3$ cannot be reached from $S_2$, then the indeces of $S_2$ and $S_3$ can be interchanged. On the other hand, any two routes from $S_1$ to $S \in \sum$ will contain the same points and can differ only in the order of sequence of the points. This follows from the fact that the route consists of all points from which S can be reached and only those points.

Next consider any route $\{S_k(S),\ k = 1, 2, \ldots, N(S)\}$ from $S_1$ to S. The sequence $\{S_k(S),\ k = 1, 2, \ldots, N(S) - 1\}$ will contain any route (up to ordering) from $S_1$ to $S_{N(S)-1}(S)$ as can easily be verified from the definition. In general, $\{S_k(S),\ k = 1, 2, \ldots, \gamma\}$, $\gamma \leqslant N(S)$, will contain any route (up to ordering) from $S_1$ to $S_{N(S)-1}(S)$. Therefore, for any $S = S^* \in \sum$ (and letting the route from $S_1$ to $S^*$ be written as in the statement of the theorem) the points $S' \in \sum$ which will contribute to the sum on the right hand side of (2.40) will belong to the sequence $\{S_k,\ k = 1, 2, \ldots, N\}$. For $S = S_\gamma(S^*)$, $\gamma \leqslant N(S^*)$, the points $S' \in \sum$ which will contribute to the sum on the right hand side of (2.40) will belong to the sequence $\{S_k,\ k = 1, 2, \ldots, \gamma\}$. It is also clear that for

* See footnote, page 33.

$S = S_1$, the right side of (2.40) is zero. Hence, if one keeps S* fixed, equation (2.40) can be written for all $S_k$, k = 1, 2, ..., N:

$$(\partial/\partial x + a(S_k))\, f(S_k, x) = \sum_{\ell=1}^{k-1} f(S_\ell, x)\, \pi(S_\ell \longrightarrow S_k). \qquad (2.43)$$

When $S_k = S_1$, this equation can be integrated to yield $f(S_1, x) = e^{-xa(S_1)}$. In general, it is easily shown by induction that $f(S_k, x)$, k = 1, 2, ..., N, is a linear function of exponentials and the form (2.41) holds for some, as yet undetermined, coefficients $\beta_t(S_k)$. The coefficients will depend on $S_k$, in general. After putting the form (2.41) into the right hand side of (2.43) and integrating one obtains:

$$f(S_k, x) = \sum_{\ell=1}^{k-1} \pi(S_\ell \longrightarrow S_k) \sum_{t=1}^{\ell} \frac{\beta_t(S_\ell)\, e^{-xa(S_t)}}{a(S_k) - a(S_t)} \qquad (2.44)$$

$$- e^{-xa(S_k)} \sum_{\ell=1}^{k-1} \pi(S \longrightarrow S_k) \sum_{t=1}^{\ell} \frac{\beta_t(S_\ell)}{a(S_k) - a(S_t)}, \quad k = 2, 3, \ldots, N.$$

An interchange of summation leads to:

$$f(S_k, x) = \sum_{t=1}^{k-1} \frac{e^{-xa(S_t)}}{a(S_k) - a(S_t)} \sum_{\ell=1}^{k-1} \pi(S_\ell \longrightarrow S_k)\, \beta_t(S_\ell) \qquad (2.45)$$

$$- e^{-xa(S_k)} \sum_{t=1}^{k-1} \left[a(S_k) - a(S_t)\right]^{-1} \sum_{\ell=t}^{k-1} \pi(S \longrightarrow S_k)\, \beta_t(S_\ell), \quad k = 2, 3, \ldots, N.$$

The two steps of putting the form (2.41) into the left hand side of (2.45) and then equating coefficients results in the following set of recurrence relations for the $\beta$'s:

$$\beta_1(S_1) = 1,$$

$$\beta_1(S_k) = \left[\alpha(S_k) - \alpha(S_1)\right]^{-1} \sum_{\ell=1}^{k-1} \pi(S \longrightarrow S_k)\, \beta_1(S_\ell), \quad k = 2, 3, \ldots, N,$$

$$\beta_2(S_k) = \left[\alpha(S_k) - \alpha(S_2)\right]^{-1} \sum_{\ell=2}^{k-1} \pi(S \longrightarrow S_k)\, \beta_2(S_\ell), \quad k = 3, 4, \ldots, N,$$

$$\vdots$$

$$\beta_{k-1}(S_k) = \left[\alpha(S_k) - \alpha(S_{k-1})\right]^{-1} \sum_{\ell=k-1}^{k-1} \pi(S \longrightarrow S_k)\, \beta_{k-1}(S_\ell), \quad k = 2, 3, \ldots, N,$$

$$\beta_k(S_k) = -\sum_{t=1}^{k-1} \beta_t(S_k), \quad k = 2, 3, \ldots, N.$$

The solution of this set of recurrence relations yield (2.42).

Notice that the solution given in Theorem 5 involves all possible "transitions" along the route from $S_t$ to $S_k$, $k = 2, 3, \ldots, N$; $t = 1, 2, \ldots, k-1$. The function $\pi(S_i \longrightarrow S_j)$ for many of these transitions might vanish and any product of terms in the expression for $\beta_t(S_k)$ that involves such a function will, of course, vanish. Thus a solution for any particular problem which satisifes the conditions of Theorem 5 will be completely determined by a knowledge of the transitions that are forbidden (transitions with $\pi(S_i \longrightarrow S_j) = 0$) and a knowledge of the non-zero values of the functions $\pi(S' \longrightarrow S)$ and $\alpha(S)$. In other words, a formal solution for any particular problem can be obtained from Theorem 5 if one knows for every point $S \in \sum$ the set of points $S'$ from which S can be reached in one step. This set of points, $\{S'\}$, will, of course, form a subsequence of the points on the route from $S_1$ to S.

—

# 3. THE STRUCTURE OF KNOWN SOLUTIONS

## 3.1 The Importance of Catagorizing Previous Solutions

A detailed theory with a history overlayed with extensive researches may long since have had its basic structure obscured. This seems to be the case with cascade showers. Various methods of attack evolved. Near the end of the chain Janossy's method appeared and was later elucidated in terms of a non-Markovian and regenerative process by Bartlett and Kendall.[22] Now we can look further back and say that nearly all known methods for a quantum solution of the distribution, nearly all the methods of attack on the average number of shower particles and the second moment are non-Markovian in structure. The possible value of making a general observation of this kind lies in picking out the strength of the stochastic properties on which present solutions are ultimately based. As will be seen, the properties presently used are much weaker than those actually satisfied by the shower phenomenon, itself. One might hope, after realizing this, that by making fuller use of the special properties of the shower, namely the Markovian property, that a simpler solution would result. To use stronger properties (or less general ones) is to be closer to the problem at hand and in possession of techniques that were not available before.

The possible danger in making a general observation of the kind above lies in overemphasizing a half-truth. A solution might be non-Markovian and still have advantages for other reasons. The last section was devoted to outlining the structure of cascade shower solutions from a view that cut arbitrarily through the lines of Markovian and non-Markovian distinctions. From the point of view of Section 2, it is most important to decide which of the set of sets enchained by the basic stochastic process will allow one the greatest freedom from the unwanted parameters of the problem - the direction and displacement parameters of the low energy particles and the individual energy parameters of all the particles in general. It will be seen in Section 4, Part 4.1, that Theorem 1 of Section 2 can be used to return to the Markov property even while solving for sets of particles whose specified energies are above some lower cut-off. This would appear to be a great advantage from the point of view of the last paragraph, but the energy parameters of each of the high energy particles are still present and make for bulky numerical results.

Owing to the energy dependence of the particle branchings it might appear that all parameters must inevitably be present throughout the calculations. Theorem 2 says otherwise. Banking heavily on the chance to find an enchained set of sets which will simplify the parameter problem, one will soon realize the following: If an a-priori distribution of individual particle energies is known, given information about the total energy of the system and the number of electrons and photons present, then the cross sections for the transitions of the system can be modified to allow direct calculations for the probabilities of $n$ electrons and $m$ photons, having regard only for one energy parameter, the total energy. This realization is a consequence of viewing the structure of previous solutions and continuing the essential features to a logical conclusion. By combining the advantages of Theorem 1 and Theorem 2, that is, by striving to keep the Markovian property whenever possible and yet admitting further reduction of the observation space into subsets to simplify the parameter problem, one is freed of the low energy spacial parameters and the individual energy parameters alike. Details will be given in Section 4, Part 4.2.

In this section, the consequences of analyzing previous solutions, as described above, will not be studied. Only the previous solutions themselves will be briefly reviewed. Having the results of Section 2 at free disposal, the approach will be to demonstrate how the assumptions of Theorems 3 and 4 are satisfied in the case of two broad methods of attack and how the consequences of the theorems lead to the same results that have been obtained by these methods.

### 3.2 Methods of Nordsieck, Lamb, Uhlenbeck, Scott, Bhabha, Ramakrishman, Messel, Potts, and Kendall

By far the most familiar entity in cascade theory is a function of the energy parameters variously called "product density", "moment density", or "correlation function". This function can be solved for directly and can be related to the moments of the number distribution through summation and energy integration. It appears throughout most of the work on the average number of shower particles under a special guise,

"the average number of electrons at the thickness $x$ with energy between $E$ and $E+dE$". Under a similar interpretation (where age replaces energy, time replaces thickness and the number of people replaces the number of particles) it appears in Kendall's[23] population studies. It appears in the papers of Nordsieck, Lamb, Uhlenbeck[3], Scott[4], Bhabha[8,9] and Ramakrishnan[10,11] with the interpretation, "the 'average product' of the numbers of particles at $x$ in $n$ differential energy ranges". It appears in the papers of Scott[15], Messel, and Potts[16,17] with the interpretation, "the probability of finding $n$ particles at $x$, each in preassigned differential energy ranges, without regard to the possible existence of further particles of other energies". The various techniques for solving for this function and relating it to the moments offer the only features that distinguish one of these studies from another. The fundamental entity is always the same. It would seem wise, then, to investigate the probabilistic meaning that appears everywhere in common: In the terminology of the last section, this function is the probability weighting of a set of sets locally enchained by the basic stochastic process of a cascade shower, given a knowledge of some initial conditions.

After accepting the last mentioned interpretation of the product density function, one can immediately recognize equation (2.30) of Theorem 3 (or, more exactly, its continuous counterpart) as the basic equation appearing in all of the above mentioned papers. Without further justification, though, the equation, which in reality has some very special relations to the moments of the electron distribution, here emerges quite incognito. The importance of attaching the same "locally enchained" to an otherwise familiar entity is seen in emphasizing the basic logic: Some fundamental sets are chosen whose probability weightings are solved for directly without influence from the low energy particles of the system. These probability weightings aid the direct solution for the moments of the number distribution of electrons, while one ignores the population of particles with energies below some arbitrary cut-off.

To see that the "product density" is really the probability weighting of locally enchained sets, the special notation introduced in Section 1.2 to describe a shower will be needed. Let $\mathcal{S}(x)$, $S$, $\Delta S$ and $\Sigma$ be the same as in Part 1.2. The corresponding symbols in the general

notation of Section 2.1 are $\mathcal{G}(x)$, $G$, $\Delta G$, $\Gamma$. Ignore processes (ii), (iii), and (v) listed in Section 1.2. The justification for ignoring the low energy processes is just that the method to follow is aimed at side stepping the low energy particles and finding the distribution of high energy ones directly. Corrections for process (ii) might be introduced later or included in a second modified approach modeled after the present one.

For any parameter set $(n, m, \bar{E}, \mu) \in \Sigma$, there will be a set of the form:

$$s \equiv \{ S' = (n', m', \bar{E}', \mu') : n' \geq n, m' \geq m, (\bar{E}', \mu') \supset' (\bar{E}, \mu) \},$$

where the symbol $\supset'$ means that the electron and photon energies specified by the couple $(\bar{E}', \mu')$ contain the electron and photon energies, respectively, specified by $(\bar{E}, \mu)$. In other words, for each parameter set $(n, m, \bar{E}, \mu)$, or for each state $S = (n, m, \bar{E}, \mu)$, the corresponding set, $s$, will be the set of all states S' for which there are at least $n$ electrons and at least $m$ photons and $n$ of these electrons and $m$ of these photons have energies which are specified by the parameters $\bar{E}, \mu$. The set of all such sets, s, will be denoted by $\sigma = \{s\}$.

If the assumption is made that for fixed values of x and $S_0$ and for fixed values of $n, m$ and $\mu$ in the expression for S the function $P(\Delta S, x \mid S_0, 0)$ (defined in Section 1.2) is absolutely continuous with respect to Lebesgue measure in the n + m dimensional energy space, then a probability density function $p(S, x \mid S_0, 0)$ exists. This density plays the role of the analogous function $P_r\{\mathcal{S}(x) = S \mid \mathcal{S}(0) = S_0\}$ that would have meaning if the space $\Sigma$ were discrete. That is to say, for all sets $\Delta S$ defined in Section 1.2:

$$\int_{\Delta S} p(S, x \mid S_0, 0) d\bar{E} = P_r\{\mathcal{S}(x) \in \Delta S \mid \mathcal{S}(0) = S_0\}. \qquad (3.1)$$

Similarly, a density transition function $p_0(s, x \mid s', y)$ exists which plays the role of $P_r\{\mathcal{S}(x) \in s \mid \mathcal{S}(y) \in s'\}$ and a special case of this is the "product density": $p_0(s, x \mid s_0, 0)$. More exactly, let the transition function $P(\Delta s, x \mid s', y)$ be defined like the function in equation (2.2). If $\Delta s$ represents all sets s corresponding to states in $\Delta S$, then for all $\Delta s$,

$$\int_{\Delta s} p_0(s, x \mid s', y) d\bar{E} = P(\Delta s, x \mid s', y). \qquad (3.2)$$

In terms of this density transition function, all of the properties (i) - (vi), of Definition 3, can be varified for the set of sets $\sigma = \{s\}$, assuming, of course, that Definition is modified to accomodate a density transition rather than a conditional probability. As a hint of the modifications necessary, property (iii) will have the condition (a) dropped and for each s, the set $C(s)$ will be all sets s' such that the differential cross section for transition from s' to s is non-zero. Property (vi) will require not that the probability for $\mathcal{S}(x)$ belonging to overlapping sets in $\sigma$ is zero, but that the differential probability for such events is a differential of higher order. Unfortunately a completely rigorous modification of Definition 3 would be too lengthy to include here. Intuitively speaking, the essential logic for all cases is contained in Definition 3 as it stands.

The two relations that have been used between the "product densities" and the moments have been that of Uhlenbeck[3,4], et.al., Bhabha[8,9], and Ramakrishnan[10,11]:

$$\overline{N(\epsilon, x \mid s_0)^n} = \sum_{n'=0}^{n} \sum_{m'=0}^{\infty} C(n', n, m') \int_{\bar{\epsilon}_{\mu'}}^{I\bar{1}} d\bar{E}\, p_0(n', m', \bar{E}, \mu', x \mid s_0, 0) \qquad (3.3)$$

and that of Messel and Potts[16,17]:

$$T(n, m, \epsilon, x \mid s_0) = \int_{\epsilon \bar{1}}^{I\bar{1}} d\bar{E}\, p_0(n, m, \bar{E}, \mu, x \mid s_0, 0) . \qquad (3.4)$$

The symbols in 3.3 are defined as follows: $\overline{N(\epsilon, x|s_0)^n}$ is interpreted as the $n^{th}$ moment of the number distribution of electrons whose energies lie above $\epsilon$, assuming that the initial state coincided with $s_0$. ($s_0$ might correspond to a single electron or a single photon of energy I.) $C(n', n, m')$ are known constants dependent on n', n, and m'. $\bar{\epsilon}_{\mu'}$ is an n'+m' - dimensional energy vector in which $\epsilon$ appears for each electron component and 0 appears for each photon component, as determined by $\mu'$. $\bar{1}$ is the n'+m' dimensional unit vector and I is the upper limit on the energy components as defined in 1.2. The symbol on the left of (3.4) is defined as the n, $m^{th}$ factorial moment of the number distribution of electrons and photons whose energies lie above $\epsilon$, assuming some initial condition, $s_0$.

The chief disadvantages in the general method of this section are (a) the equations governing the "product densities" are, in reality, large sets of simultaneous integro-differential-difference equations. (b) As a consequence of (a) the solution for the "product density" corresponding to n electrons and m photons generally involves matrices of order n+m and inverse integral transforms of multiplicity n+m+1. (c) The solution for the probabilities, themselves, of, say, n electrons must be carried out through the moments. Unfortunately, many moments are needed because of the irregularities in the magnitudes of the probabilites (the probabilities for an odd number of electrons being depressed because of pair production). (d) With the handicaps of (a), (b), and (c), only asymptotic cross sections can be used, that is, full screening approximations, valid for high energies, are used throughout. (e) Finally, Compton effect and multiple scattering (processes (v) and (iii))are ignored and collision loss (process (ii) ) can be corrected for only approximately.

3.3 The Method of Janossy.

Let $\mathcal{S}(x)$, S, $\Delta S$ and $\Sigma$ be as in part 1.2. Ignore processes (ii), (iii), and (v). (Corrections for process (ii) will not be discussed here, though Janossy does take it into account.) The sets in $\Sigma$ whose probability weightings Janossy[12-14] chooses to solve for will immediately be defined: Let $\epsilon$ be a number in the interval $0 \leq \epsilon \leq I$. For every parameter set $(n, m, \epsilon)$, $n = 0, 1, 2, \ldots$, $m = 0, 1, 2, \ldots$, there will be a set $s = \{S'\}$ which contains all states with (a) exactly n electrons and m photons whose energies are greater than $\epsilon$ and (b) zero or more electrons and photons whose energies are less than or equal to $\epsilon$. The set $\{s\}$ of all such sets s will be denoted by $\sigma = \{s\}$. It can be verified by the reader that all three properties in Definition 4 are satisfied by $\sigma$, where the subset $\Sigma_0$ of $\Sigma$ (corresponding to $\Gamma_0$) can be taken as all states of the type $S' = (1, 0, E, \mu)$ or $S'' = (0, 1, E, \mu)$, $0 \leq E \leq I$, that is, all states made up of exactly one electron or one photon of any specified energy. The meaning of property (iii), Defintion 4, in the case of cascade showers, is that the future states of the system, given any initial state, are determined by the events initiated by all of the electrons and photons taken individually. Every

particle initiates its own shower when it branches and the compound effect of these individual showers yields the whole. The significance of property (iii) is that it permits equation (2.38) to be written, an equation involving transition functions conditioned only on single particle states.

In Janossy's treatment, where the only branching processes involved are processes (i) and (iv), only two terms enter in each product of the summands of Equation (2.38). These correspond to events initiated by the two branch products from the initial state $G_0$. The awkward summation of Equation (2.38) can be eliminated by transforming the equation to one satisfied by moment generating functions. The non-linearity remains throughout and, once having resorted to generating functions, one must ultimately recover the probabilities through the moments. The apparent simplification, the disappearance of the individual energy parameters of each particle, does not seem to have brought one closer to a knowledge of the basic probability weightings of the sets $\sigma = \{s\}$ . Most of the objections made at the end of Part 3.2 hold now for Janossy's technique.

It is interesting to see from another point of view, as in the papers of Bartlett and Kendall[22], and Ramakrishnan and Mathews[7], that Janossy's method is based on a regenerative process. A random function $\mathcal{Y}(x)$, $x \geq 0$, is <u>regenerative</u>, if there is some $y > 0$ such that for all $x > y$:

$$\text{Dist}\{\mathcal{Y}(x) \mid \mathcal{Y}(y)\} = \text{Dist}\{\mathcal{Y}(x) \mid \mathcal{Y}(t),\ t \leq y\}.$$

Any point, y, satisfying this definition is called a point of regeneration. Such a point in the development of a cascade shower described by the random function $\mathcal{Y}(x)$, is the point of the first collision and this is true even if the shower is incompletely described (or described by a new random function) as belonging to sets of the type s. From this viewpoint, Janossy's method is clearly based on properties that are weaker than Markovian, for in the basic Markovian process, $\mathcal{Y}(x)$, every point y is a point of regeneration (see Assumption II, Part 1.2).

Though Janossy developed his theory independently, the idea of a regenerative process seems to have been developed in other contexts by Palm[24], Bellman and Harris[25], and others.

## 3.4 Other Known Methods.

It will be remarked, without elaboration, that Bartlett and Kendall[22] have investigated the characteristic functional in relation to a cascade shower and Ramakrishnan[26] has investigated Janossy's equations in the special case when $\epsilon = 0$. The first approach is more general than Janossy's and, hence, should meet at least as many difficulties in a practical solution; the second approach is much less general and avoids the difficulties at the expense of accuracy. The assumptions of the latter approach will admit realistic answers only for small penetration thicknesses, as the author points out.

# 4. METHODS BASED ON ENCHAINED SETS AND CONDITIONALLY ENCHAINED SETS

## 4.1 A Markovian Solution Based on Theorems 1 and 5.

It was pointed out at the beginning of Section 3 that nearly all known solutions of the cascade shower problem are non-Markovian in structure. Section 3 was then devoted to showing that the majority of solutions have used locally enchained sets or multiply enchained sets. One should now glance at Theorems 3 and 4, which involve the latter concepts: It is not necessary, under the hypotheses of the theorems, for the functions satisfying Equations (2.30) and (2.38) to admit interpretation as transition functions that correspond to a Markovian process. Thus, the original statement about non-Markovian structure.

Granted these things, it may now be useful to apply Theorem 1 to the cascade problem and arrive at a function that has immediate physical meaning, that can be solved for directly through equation (2.3), and which must be interpreted as the transition function for a Markovian random function. An approach of this kind was made in an earlier paper by the present author[27]. The basic model in that paper included the exact cross section for process (iv), radiation, and permitted approximations under which the divergence of the radiation cross section cancelled out. The diffusion equations were a genuine simultaneous set and had to be solved by matrix methods. In the present paper a model will be introduced that imposes a slightly unrealistic view of the radiation spectrum but that permits a solution to be carried out in terms of Theorem 5. The advantage is that the probability weighting for n electrons & m photons of specified energies can be directly expressed as a linear function of exponentials. Compton collision, a process ignored in the previous paper, will be included in the model of the section. Collision loss and multiple scattering will be present as before.

Let $\mathfrak{S}(x)$, S, $\Delta S$, $\Sigma$ and I be as in Section 1, Part 1.2. Let $\epsilon$ be any number in the interval $0 < \epsilon < I$. For any parameter set $(n, m, E, \mu) \in \Sigma$, $0 < \epsilon < E_k$, $k = 1, 2, \ldots, n+m$, $n > 0$, $m > 0$, there will be a set s of states S', $s \equiv \{S'\}$, defined as follows: $s \equiv \{S'\}$

is the set of all states S' for which (i) there are at least n electrons and at least m photons and (ii) n of these electrons and m of these photons have energies which are specified by the parameters $(\bar{E}, \mu)$, and (iii) the remaining electrons and photons whose energies are not specified by $(\bar{E}, \mu)$ have energies less than or equal to $\epsilon$. Let $\sigma'$ be the set of all sets s so defined and let $s^o = \Sigma - \sigma'$. It follows that $\sigma$, that is defined as $\sigma'$ with $s^o$ added, will represent a partition of $\Sigma$ in the sense of Defintion 1. The event that the system belongs to the set of states $s^o$ corresponds to the event that no electrons or photons are present with energies above $\epsilon$.

The reader can easily verify that the set of sets $\sigma$ satisfy all the conditions of Definition 1. $\sigma$ is then enchained by the basic stochastic process, $\mathcal{S}(x)$, of a cascade shower. Theorem 1 can be applied and, moreover, the cross sections for transition in Equation (2.3) will not be influenced by the low energy behavior of the particles. If $\Delta s$ represents the sets of all sets s corresponding to points S in $\Delta s$, then the solution of Equation (2.3) will yield $P(\Delta s, x \mid s_0, 0)$ : the probability that the system of electrons and photons at depth x will be made up of n electrons and m photons with energies in some specified energy intervals above $\epsilon$ ; this is regardless of how many electrons or photons are present with energies less than or equal to $\epsilon$.

It will be desirable to apply the solution of Theorem 5 to Equation (2.3). In order to do this, it will be necessary to introduce some assumptions about the basic processes (i) - (v), Section 1. Physically, speaking the assumptions are these:

I'. For convenience of notation, shower penetration will be measured in radiation lengths and the initial energy I will be taken as unit energy.

I. All five processes described in Section 1, Part 1.2, are operative.

II. The effect of angular emission at branching is negligible for processes (i), (ii), (iv), (v), when the primary energies of the processes are above $\epsilon$. The effect of multiple scattering (process (iii) by an electron of energy u, $u > \epsilon$, is to increase the cross sections for radiation and collision loss by the factor $\lambda(u)$, where $b\,\lambda^{-1}(u)$ is the average distance that an electron of energy u would penetrate into the scattering medium after entering

normally to the surface and traveling $b$ radiation lengths (the primary electron is identified with its secondary if radiation occurs); $b$ is some constant which for simplicity can be chosen as 1. Hence, the shower of electrons and photons above $\epsilon$ energy is one dimensional in the sense that no angular or lateral parameters are necessary, and the parameter set $S = (n, m, \bar{E}, \mu)$ is adequate for describing the shower.

III. The differential cross sections (to be defined below, (4.5), (4.6), in terms of transition functions for radiation loss and pair production (processes (i) and (iv) ) are given by the following, in which $\epsilon < u \leq 1$ :

$$w^{(1)}(u-v, v) = \begin{cases} g^{(1)}(u, v/u)/v & ; \quad \epsilon < v \leq u \\ \int_0^{\epsilon} dv\, g^{(1)}(u, v/u)/\epsilon & ; \quad v = \epsilon \\ 0 & ; \quad \text{Otherwise}, \end{cases} \tag{4.1}$$

$$w^{(2)}(u-v, v) = g^{(2)}(u, v/u)/u \quad ; \qquad 0 \leq v \leq u, \tag{4.2}$$

where $u$ represents the primary energy and $v$, in the first case, represents the photon energy and, in the second case, the electron energy. The g's, as functions of $v/u$, are slowly varying over most of the range and $g^{(2)}(u, v/u)$ is symmetric about the point $v/u = 1/2$, as it is, for instance, when computed under the Born approximation.

IV. As for the process (ii), an electron of energy $u$ can lose energy (in addition to radiation losses) only in amounts $\epsilon$ and the cross section per radiation length for this loss is $e(u)/\epsilon$ , where $e(u)$ is the true collision loss per radiation length for an electron of energy $u$.

V. In every Compton collision (process (v) ) the full energy, $u$, of the photon is transmitted to the electron and the cross section for an event of this type is the total cross section for Compton collision, say, $ph(u)$.

VI. The integrands of certain integrals, to be discussed later (see (4.20), (4.22), (4.23) ) are approximately linear over regions of length $\epsilon$. A special approximation will be made for integrals of type (4.21). (This assumption is introduced here, because its justification essentially depends on the energy variation of the basic processes.)

VII. $\epsilon \gg \mu c^2/I$, where I is the initial energy of the shower at zero penetration and $\mu c^2$ is the rest energy of the electron. AssumptionsIII, IV, and VI are better, if it is also true that $\epsilon << 1$.

The justification of these assumptions will be made during the discussion of the solution that is a consequence of them. It should be noticed that the material difference between these assumptions and those made previously[27] is the accomodation for Compton collision, Assumption V, in this paper and modification of the radiation cross section in Assumption III.

In order to solve equation (2.3) under Assumptions I-VII it would be convenient to solve for a probability density (see equation 3.1) rather than for the transition function $P(\Delta s, x|s_0, 0)$. Unfortunately, the transition function can no longer be considered absolutely continuous as it was in Section 3 (see the paragraph containing Equation 3.1), for Assumptions III and IV allow discrete amounts of energy (specifically $\epsilon$ - amounts) to be lost from the system of electrons and photons whose energies are above $\epsilon$. It can be seen that for fixed initial conditions and for fixed penetration thicknesses, the probability weighting for finding N particles of specified energies above $\epsilon$ (regardless of how many particles are below $\epsilon$ ) is not a continuous function of energy over N-dimensional energy space. Instead, the probability weighting is distributed continuously over the N-dimensional space except along the hyperplanes $H(1 - \nu\epsilon) = \{\bar{E}' : \bar{E}' \cdot \bar{1} = 1 - \nu\epsilon\}$, where the set of N-dimensional energy vectors, $\bar{E}'$, satisfy the relation $\bar{E}' \cdot \bar{1} = \sum_{k=1}^{N} E_k' = 1 - \nu\epsilon$, $\nu = 0, 1, 2, \ldots$ On these hyperplanes the probability weighting is distributed in N-1 dimensions. For example, suppose that a shower is started by an electron of unit energy. If one ignores how many shower particles have energies below $\epsilon$, then at any specified thickness there will be positive probability of finding a single electron with energy $1$, $1 - \epsilon$, $1 - 2\epsilon$, etc. These

probabilities correspond to the event that no process has occurred, one collision process or one radiation process with energy loss $\epsilon$ has occurred, and so forth. At the same time there will be a probability density distributed between $\epsilon$ and 1 which is due to the fact that continuous amounts of energy can be lost from the system through the creation of particles with energies less than $\epsilon$. As a consequence of the fact that the transition function is not absolutely continuous, it will be necessary to solve for a continuous and a discontinuous part. That is, the probability weightings along the hyperplanes will have to be solved separately from the probability weightings elsewhere. The solution will presently be carried out after defining what is meant by the continuous and discontinuous parts.

One may begin to doubt the advantages of Assumptions like III and IV, if they are responsible for discontinuities. Then again, it would be best, one might think, to ignore the discontinuities, if they are present and $\epsilon$ is small. Simply treat the transition function as though it were absolutely continuous. This was done, in fact, in the earlier paper[27] mentioned above. As will be found, however, there are even practical advantages to the present more accurate approach of accepting the full consequences of Assumptions III and IV. For instance, the high energy part of the shower will be determined almost completely by the discontinuous part, the easiest part to solve for. Continuous modes of energy loss from the system become important only at low energies, or, in other words, toward the end of the shower.

The continuous and discontinuous parts of the transition function will now be defined, as well as the differential cross sections appearing in Assumption III. Let H represent the sum of all the n+m-1 -dimensional hyperplanes defined above: $H \equiv \sum_{\nu=0}^{1/\epsilon} H\,(1 - \nu\epsilon)$. In this notation, the dimensions of H are implicitly understood. The symbol H will always appear in context with other symbols such as s, $\Delta$s, S, etc., and it will be understood that the dimensions of these symbols are consistent. That is to say, when s is a set in n+m dimensions, H has n+m-1 dimensions. For any parameter set $(n, m, \overline{E}, \mu) \in \Sigma$, $0 < \epsilon < E_k$, $k = 1, 2, \ldots, n+m$, $n > 0$, $m > 0$, let the corresponding set s defined above be written $s(n, m, \overline{E}, \mu)$ and as special cases let $s_1 = s(1, 0, 1, 1)$, $s_2 = s(0, 1, 1, 1)$. Let $\Delta_H s = \{ s(n', m', E', \mu') : \overline{E}' \in H,\ n' = n,\ m' = m,\ \mu' = \mu \}$.

It will now be possible to define the continuous and discontinuous parts of the transition function, $p_2(s(n, m, \bar{E}, \mu), s \mid s_j, 0)$ and $p_1(s(n, m, \bar{E}, \mu), x \mid s_j, 0)$, $j = 1, 2$, respectively. For all fixed $s \in \sigma$, $n+m > 1$, let the functions be continuous in x, for all $x \geq 0$. For all fixed values of $x > 0$, n, m, $n+m > 1$, and $\mu$, let $p_2(s, x \mid s_j, 0)$ be a measurable function of $\bar{E}$ and let $p_1(s, x \mid s_j, 0)$ be measurable with respect to the Borel field defined on H. Finally, for all $x > 0$, for all sets of the type $\Delta s$, $n+m > 1$, and for $j = 1, 2$, let:

$$\int_{\Delta s} p_2(s, x \mid s_j, 0)\, d\bar{E} = P(\Delta s - \Delta_H s \mid s_j, 0), \tag{4.3}$$

$$\int_{\Delta s \cap \Delta_H s} p_1(s, x \mid s_j, 0)\, d\bar{E}\Big|_H = P(\Delta s \cap \Delta_H s \mid s_j, 0), \tag{4.4A}$$

$$p_1(s, x \mid s_j, 0) = 0, \quad s \notin \Delta_H s. \tag{4.4B}$$

The integral in (4.4A) is an $n+m-1$ dimensional integral extended over the intersection of the hyperplanes $\Delta_H s$ and $\Delta s$. The symbol "$d\bar{E}|_H$" reminds one that the differential really involves $n+m-1$ independent variables. When $n+m = 1$ the continuous part is still defined as above. For all fixed values of E, $E > \epsilon$, and for $j = 1, 2$, let $p_1(s^0, x \mid s_j, 0)$, $p_1(s(1, 0, E, \mu), x \mid s_j, 0)$ and $p_1(s(0, 1, E, \mu), x \mid s_j, 0)$ be continuous functions of x, $x \geq 0$. For all $x > 0$ and all $E > \epsilon$ let:

$$p_1(s^0, x \mid s_j, 0) = P(s^0, x \mid s_j, 0), \tag{4.4C}$$

$$p_1(s(1, 0, E, \mu), x \mid s_j, 0) = P(s(1, 0, E, \mu), x \mid s_j, 0), \tag{4.4D}$$

$$p_1(s(0, 1, E, \mu), x \mid s_j, 0) = P(s(0, 1, E, \mu), x \mid s_j, 0). \tag{4.4E}$$

$\mu$ can be set equal to 1 in the above. It should be noticed that the right

hand side of (4.4D) and (4.4E) is zero except perhaps for $E = 1 - \nu\epsilon$, $\nu = 0, 1, 2, \ldots, 1/\epsilon$ .

From the definition of the discontinuous part, it is possible to define the differential cross sections for pair production and radiation used in Assumption III:

$$w^{(1)}(1-v, v) = \lim_{x \to 0} p_1(s(1, 1, \bar{E}, \mu), x|s_1, 0) / x , \tag{4.5}$$

$$w^{(2)}(1-v, v) = \lim_{x \to 0} p_1(s(2, 0, \bar{E}, \mu), x|s_2, 0) / x, \tag{4.6}$$

where $0 \leq v \leq 1$ and $\bar{E} = (1-v, v)$ and the value of $\mu$ is understood to associate the photon with $v$ in the first place and the electron with $v$ in the second place. Since the unit chosen is entirely arbitrary, this definition extends to differential cross sections for initiating particles of arbitrary energy. Similar definitions would hold for the collision loss and Compton collision cross sections but will not be given here.

The following notational definitions will also be used:

$$p_1^{(j)}(0, 0, x) = p_1(s^0, x|s_j, 0), \tag{4.7}$$

for $j = 1, 2$; $x \geq 0$;

$$p_1^{(j)}(n, m, \bar{E}, \mu, x) = p_1(s(n, m, \bar{E}, \mu), x|s_j, 0), \tag{4.8}$$

$$p_2^{(j)}(n, m, \bar{E}, \mu, x) = p_2(s(n, m, \bar{E}, \mu), x|s_j, 0), \tag{4.9}$$

for $j = 1, 2$; $n = 0, 1, 2, \ldots$, $m = 0, 1, 2, \ldots$, $n+m > 0$; $E_k > \epsilon$ , $k = 1, 2, \ldots, n+m$; $\mu = 1, 2, \ldots, 2^{n+m}$; $x \geq 0$.

$$\psi^{(1)}(u,\ v/u) = \mathcal{K}(u) g^{(1)}(u, v/u)/v, \tag{4.10}$$

for $\epsilon < u \leq 1, \quad \epsilon < v \leq u;$

$$\psi^{(2)}(u,\ v/u) = 2g^{(2)}(u, v/u)/u, \tag{4.11}$$

for $\epsilon < u \leq 1, \quad 0 \leq v \leq u$. In all of the following definitions the range of u is $\epsilon < u \leq 1$ :

$$\varphi_1^{(1)}(u) = \int_0^{\epsilon} dv\, \mathcal{K}(u)\, g^{(1)}(u, v/u)/\epsilon\ , \tag{4.12}$$

$$\varphi_2^{(1)}(u) = \mathcal{K}(u)\, e(u)/\epsilon\ , \tag{4.13}$$

$$\varphi_1^{(2)}(u) = \begin{cases} \int_{u-\epsilon}^{\epsilon} dv\, g^{(2)}(u, v/u)/u, & \epsilon < u \leq 2\epsilon\ , \\ 0 & \text{otherwise}\ , \end{cases} \tag{4.14}$$

$$\varphi_2^{(2)}(u) = p\ell(u)\ , \tag{4.15}$$

$$\alpha^{(1)}(u) = \int_{\epsilon}^{u} dv\, \psi^{(1)}(u, v/u) + \sum_{s=1}^{2} \varphi_s^{(1)}(u)\ , \tag{4.16}$$

$$\alpha^{(2)}(u) = \int_0^{u/2} dv\, \psi^{(2)}(u, v/u) + \varphi_2^{(2)}(u)\ . \tag{4.17}$$

The function $p_1^{(j)}(n, m, \bar{E}, \mu, x)$ will be zero unless the energy vector, $\bar{E}$, belongs to some hyperplane $H(1 - \nu\epsilon)$, $\nu = 0, 1, \ldots$. Hence, it will be convenient to have a symbolism to show which hyperplane $\bar{E}$ belongs to. This will be accomplished by writing $\bar{E}_\nu$ when $\bar{E}_\nu \cdot \bar{1} = 1 - \nu\epsilon$. Notice that $\nu$ does not refer to the dimensions of $\bar{E}_\nu$. The function $p_2^{(j)}(n, m, \bar{E}, \mu, x)$ may be non-zero for all values of $\bar{E}$, but from the practical point of view it will be enough to know the limiting values of this function as $\bar{E}$ approaches $\bar{E}_\nu$, $\nu = 1, 2, 3, \ldots$, that is, the limiting values along the surfaces of the hyperplanes. It is only necessary to establish which way the limit is to be taken. If one writes $\bar{E}_r$ when $\bar{E}_r \cdot \bar{1} = 1 - r$, $r \geq 0$, then $p_2^{(j)}(n, m, \bar{E}_\nu, \mu, x)$ will be understood to mean:

$$p_2^{(j)}(n, m, \bar{E}_\nu, \mu, x) = \lim_{r \nearrow \nu} p_2^{(j)}(n, m, \bar{E}_r, \mu, x), \tag{4.18}$$

where the rising arrow indicates that the limit is taken through increasing values of r. The convention (4.18) means that the surfaces of the hyperplanes are approached through decreasing values of total energy, $\bar{E} \cdot \bar{1}$. $\nu$ is an index of the total energy of the system and will play an important part in the solution of the ensuing equations.

The boundary conditions of the problem now take the form:

$$p_i^{(j)}(n, m, \bar{E}_\nu, \mu, 0) = \delta(n + j - 2)\,\delta(m + 1 - j)\,\delta(i - 1),$$

$$p_i^{(j)}(n, m, \bar{E}_\nu, \mu, x) = 0 \quad \text{For all } x \geq 0, \text{ if } \nu < 0, \text{ or } E_k < \epsilon, \tag{4.19}$$

$$k = 1, 2, \cdots, n + m, \text{ or } n < 0, \text{ or } m < 0.$$

The limiting form of equations (2.3) becomes equations in the continuous and discontinuous parts, (4.7) - (4.9), when the sets $\Delta s$ (or in the general notation, $\Delta g$) represent the set of all sets s corresponding to points S in $\Delta S$, and when the energy intervals, $(\bar{E}, \bar{E} + \bar{e})$, appearing in the

definition of the sets $\Delta S$ (Part 1.2) are taken to be arbitrarily small. This limiting form can be written down immediately and the $\pi'(g' \to \Delta g)$ and $\alpha'(g')$ are now expressed in terms of the differential cross sections, (4.10) - (4.17). The functions $f_1^{(j)}(n, m, \bar{E}_\nu, \mu, x)$ can be solved for without any knowledge of the functions $f_2^{(j)}(n, m, \bar{E}_\nu, \mu, x)$ and, hence, one set of equations will involve only the discontinuous part and will involve no integrals. On the other hand, the equations for $f_2^{(j)}(n, m, \bar{E}_\nu, \mu, x)$ will include some terms involving the discontinuous parts, because it is possible for transitions to take place which carry the system from the hyperplanes into the continuous region. The latter set of equations will also involve integrals of the the types:

$$\int_{E_k}^{E_k+\epsilon} du\, f_2^{(j)}(n-1, m+1, \bar{E}', \mu', x)\, \psi^{(2)}(u, E_k/u), \tag{4.20}$$

$$\int_{E_k}^{E_k+\epsilon} du\, f_2^{(j)}(n+1, m-1, \bar{E}', \mu', x)\, \psi^{(1)}(u, E_k/u), \tag{4.21}$$

$$\int_{\epsilon}^{2\epsilon} du\, f_2^{(j)}(n+1, m, \bar{E}', \mu', x) \sum_{s=1}^{2} \varphi_s^{(1)}(u), \tag{4.22}$$

$$\int_{\epsilon}^{2\epsilon} du\, f_2^{(j)}(n, m+1, \bar{E}', \mu', x)\, \varphi_1^{(2)}(u). \tag{4.23}$$

in which the variable of integration, u, is considered to be a component of the vector variable $\bar{E}'$, $\mu'$ orders the electrons and photons among the components of $\bar{E}'$, and $E_k$ is some energy component of the energy vector $\bar{E}_\nu$ in $s(n, m, \bar{E}_\nu, \mu)$. The first integral above represents the fact that

a transition can take place whose differential cross section is $\psi^{(2)}(u, E_k, /u)$ and which carries the system from the set of states $s(n-1, m+1, \bar{E}', \mu')$ to the set of states $s(n, m, \bar{E}_\nu, \mu)$. The values of the energy vector, $\bar{E}'$, and ordering parameter, $\mu'$, need not be explicitly specified because there is only one set of states $s(n-1, m+1, \bar{E}', \mu')$ which can be carried by $\psi^{(2)}(1, E_k/u)$ into $s(n, m, \bar{E}_\nu, \mu)$. Physically speaking, the first integral means that a photon can make an electron pair one member of which has less that $\epsilon$-amount of energy and the other member of which has $E_k$-amount. The electron with energy $E_k$ appears as a defining member in the new set of states $s(n, m, \bar{E}_\nu, \mu)$. The other integrals have similar interpretations. For instance, the integral (4.23) corresponds to a transition in which the total number of particles with energies above $\epsilon$ decreases by 1. A photon may be responsible for such a transition when it makes an electron pair and both members of the pair have less than $\epsilon$ amount of energy.

Assumption VI, above, will now be understood to mean that all integrals of the type (4.20), (4.22), and (4.23) that appear in the equations for $p_2^{(j)}(n, m, \bar{E}_\nu, \mu, x)$ can be replaced by the product of the length of the interval of integration and the mean value of the integrand within the limits of integration. Assumption VI will also mean that $p_2^{(j)}(n+1, m-1, \bar{E}', \mu', x)$, appearing in (4.21) can be replaced by the constant value, $p_2^{(j)}(n+1, m-1, \bar{E}'_{\nu-1}, \mu', x)$, everywhere in the interval $(E_k, E_k + \epsilon)$. If this last special assumption is not imposed, the resulting equations would be a simultaneous set in spite of the special form for the radiation cross section chosen in Assumption III.

If full advantage is taken of all the Assumptions I - VII and the special notational definitions (4.7) - (4.17) are used, then the equations for the continuous and discontinuous parts of the transition function can be written as follows: (The equations hold for any values of the parameters and variables for which the functions are defined.)

$$(\partial/\partial x)p_1^{(j)}(0,0,x) = \sum_{r=1}^{2} p_r^{(j)}(1,0,1/\epsilon-2,1,x)\sum_{s=1}^{2}\varphi_s^{(1)}(2\epsilon)\left[\delta(r-1)+\frac{\epsilon}{2}\delta(r-2)\right]$$
$$+ p_2^{(j)}(1,0,1/\epsilon-1,1,x)\sum_{s=1}^{2}\varphi_s^{(1)}(\epsilon)$$
$$+ p_2^{(j)}(0,1,1/\epsilon-1,1,x)\,\frac{\epsilon}{2}\,\varphi_1^{(2)}(\epsilon) \qquad (4.24)$$

$$\left(\partial/\partial x + \sum_{k=1}^{n}\alpha^{(1)}(E_k) + \sum_{k=n+1}^{n+m}\alpha^{(2)}(E_k)\right) p_i^{(j)}(n,m,\bar{E}_\nu,\mu,x) \qquad (4.25)$$

$$= \sum_{k=1}^{n}\sum_{\ell=n+1}^{n+m} p_i^{(j)}(n,m-1,\bar{E}'_\nu,\mu',x)\,\psi^{(1)}(E_k+E_\ell, E_\ell/E_k+E_\ell)$$
$$+ \sum_{k=2}^{n}\sum_{\ell=1}^{k-1} p_i^{(j)}(n-2,m+1,\bar{E}'_\nu,\mu',x)\,\psi^{(2)}(E_k+E_\ell, E_\ell/E_k+E_\ell)$$
$$+ \sum_{k=1}^{n} p_i^{(j)}(n-1,m+1,\bar{E}'_\nu,\mu',x)\left[\varphi_2^{(2)}(E_k)+\frac{\epsilon}{2}\psi^{(2)}(E_k,1)\delta(i-2)\right]$$
$$+ \sum_{k=1}^{n} p_i^{(j)}(n,m,\bar{E}'_{\nu-1},\mu',x)\sum_{s=1}^{2}\varphi_s^{(1)}(E_k+\epsilon)$$
$$+ \left\{\sum_{r=1}^{2}\sum_{k=1}^{n} p_r^{(j)}(n-1,m+1,\bar{E}'_{\nu-1},\mu',x)\,\psi^{(2)}(E_k+\epsilon,E_k/E_k+\epsilon)\left[\delta(r-1)+\frac{\epsilon}{2}\delta(r-2)\right]\right.$$
$$+ \sum_{r=1}^{2}\sum_{k=n+1}^{n+m} p_r^{(j)}(n+1,m-1,\bar{E}'_{\nu-1},\mu',x)\left[\psi^{(1)}(E_k+\epsilon,E_k/E_k+\epsilon)\,\delta(r-1)\right.$$
$$\left.\left. + \int_{E_k}^{E_k+\epsilon}\psi^{(1)}(u,E_k/u)\,du\,\delta(r-2)\right]\right.$$
$$+ \sum_{r=1}^{2} p_r^{(j)}(n+1,m,\bar{E}'_{\nu-2},\mu',x)\sum_{s=1}^{2}\varphi_s^{(1)}(2\epsilon)\left[\delta(r-1)+\frac{\epsilon}{2}\delta(r-2)\right]$$
$$+ p_2^{(j)}(n+1,m,\bar{E}'_{\nu-1},\mu',x)\,\frac{\epsilon}{2}\sum_{s=1}^{2}\varphi_s^{(1)}(\epsilon)\left[1-\delta(\nu-1)\right]$$
$$\left. + p_2^{(j)}(n,m+1,\bar{E}'_{\nu-1},\mu',x)\,\frac{\epsilon}{2}\,\varphi_1^{(2)}(\epsilon)\left[1-\delta(\nu-1)\right]\right\}\delta(i-2)$$

In order to see that equation (4.25) is a special case of equation (2.40) already solved in general, consider the following. Let $R \equiv (i, j, n, m, \overline{E}_\nu, \mu)$ be a set of parameters whose values may fall anywhere in the domain of definition of the function $p_i^{(j)}(n, m, \overline{E}_\nu, \mu, x)$, for any $x>0$. For a fixed value of j, the space $\{R\}^{(j)}$ of all possible values of the parameter set, R, will correspond to $\Sigma$ of Part (2.2). For any fixed values of the parameters, $R^* \in \{R\}^{(j)}$, there is a finite sequence $\{R_k(R^*)\}$, $R_k(R^*) \in \{R\}^{(j)}$, of parameter sets which depend on $R^*$ and which satisfies all the conditions of a "route", as defined in Part 2.2. (When j = 1 or 2 the point $S_1$ of Part (2.2) is $R_1 \equiv (1, 1, 1, 0, 1, 1)$ or $R_2 \equiv (1, 2, 0, 1, 1, 1)$, respectively.) The function $\pi(S' \to S)$ has now the various forms $\pi[(n, m-1, \overline{E}'_\nu, \mu') \to (n, m, E, \mu)] =$ $\psi^{(1)}(E_k + E_\ell, E_\ell / E_k + E)$, $\pi[(n, m, \overline{E}'_{\nu-1}, \mu') \to (n, m, \overline{E}_\nu, \mu)] =$ $\sum_{S=1}^{2} \varphi_S^{(1)}(E_k + \epsilon)$, etc. As an example, a "route" from $R_2$ to $(1, 2, 2, 0, E, \mu)$ is specified by the sequence of Table 1. Table 1 can immediately be extended to represent the "route" from $R_1$ or $R_2$ to any point R. In doing so, it is noticed that the order of sequence of the points along any route is completely determined by the parameters $(i, n, m, \nu)$ only.

Table 1.

| | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| i | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| n | 0 | 1 | 0 | 1 | 2 | 0 | 1 | 0 | 1 | 2 | 3 | 0 | 1 | 2 | 0 | 1 | 2 | |
| m | 1 | 0 | 2 | 1 | 0 | 1 | 0 | 3 | 2 | 1 | 0 | 2 | 1 | 0 | | | | |
| $\nu$ | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | | | | |

As remarked at the end of Part 2.2, every point on the route to R can be reached by one step transitions from a subsequence of points lying somewhere on the route before that point and those subsequences alone give the special character to the Solution (2.41), (2.42) for a particular case. The subsequence of points corresponding to the point $(1, 2, 2, 0, E, \mu)$ has been especially marked in Table 1. This as well as any other can be determined immediately by reading the right hand side of equation (4.25).

The Assumptions II - VII that make equations (4.24), (4.25) possible all involve the quantity $\epsilon$. The physical justification for the assumptions will be in order when this quantity is properly chosen. For instance Assumption II requires that multiple scattering is not so great for energies above $\epsilon$ that it cannot be corrected for by the factor $\mathcal{A}(u)$, roughly interpreted as the ratio of integrated to projected path length for electrons. Assumption III requires that the effect of a delta function placed at $v = \epsilon$ in the bremsstrahlung spectrum will not significantly alter the shower, provided this delta function, which replaces the region of the spectrum below $\epsilon$, has the weight of the average energy lost to this region. Assumption IV, an assumption about the way that energy is lost through electron-electron collision, allows that, as far as the shower is concerned, $\epsilon$-amount of energy could be lost in each such collision and all would be the same. This assumption, of course, includes the proviso that such artificial collisions occur few and far enough between to permit the average energy loss per radiation length by electrons to be correct. Assumption V implicitly demands that the effect of Compton collision on the shower is slight enough to permit an idealization of the energy transfer. The effect will be smaller, as is known, the larger $\epsilon$. Assumption VI requires that the energy variation of certain functions is not too great over intervals of length $\epsilon$. Assumption VII asks that the rest mass of the electron can be neglected in comparison to other energy parameters of the problem - all of which are greater in magnitude than $\epsilon$. For instance, when a photon materializes into an electron-positron pair, the fraction of its energy that goes into the rest mass of the pair can be neglected. It is immediately seen that Assumptions II, V, and VII are better justified the larger $\epsilon$, while Assumptions III, IV, and VI are more consistent with a small $\epsilon$. These opposing requirements need not be incompatible from a practical point of view:
If $\epsilon$ is chosen in the region of the critical energy (the region where collision losses are nearly equal to radiation losses) and if the initial energy of the shower (which establishes the unit of energy) is chosen at least 10 times the critical energy, then all requirements are approximately satisfied, simultaneously. This means that the methods of this section are valid for showers whose initiating energies are about 70 Mev or greater in lead or 1000 Mev or greater in air, provided $\epsilon$ is chosen near the critical energy.

The assumptions which require special comment are II - VI. In reference to II, it is necessary to point out why a correction for multiple scattering takes precedence over a correction for angular emission at branching: It does, because, according to Rossi and Greisen[28], the average angle of emission per radiation length for primary electrons and primary photons is roughly one-fortieth the root mean square angle of scattering of electrons in one radiation length. In reference to III, the modification found there eliminates an otherwise divergent term but at the same time keeps the radiation process intact in the sense of preserving the correct average energy loss to radiation. Assumption IV was introduced to eliminate the myriad - step transitions encountered in the trial solution of Part 1.2 and yet to permit the correct average amount of energy to be lost from the shower per unit path length. Assumption V is justified for two reasons: One sees, first, that Compton scattering that decreases rapidly above $\epsilon$ (the critical energy) has a cross section there one-fifth as great in air as the cross section for pair production: The effect of Compton scattering on the shower cannot be very great. Second, one calculates from the Klein-Nishina formula that at 20 Mev more than 70% of the photon energy is transferred to the electron in roughly 70% of all Compton collisions. It is sufficient to assume that all the photon energy is transferred to the electron all of the time, especially since the most important contribution of this process to the shower is to introduce charge.

The justification of Assumption VI must ultimately rest on the behavior of the functions $g^{(1)}(u, v/u)$ and $g^{(2)}(u, v/u)$ as functions of the fractional energy, $v/u$, and especially on their behavior near the end point $v/u = 1$. Unfortunately, these functions have been computed[29] for high primary energies under the Born approximation, which is not valid near the end point cited above. More recent calculations[30], while not limited by the Born approximation, are still not valid when, for instance, the electron radiates the most of its energy. Experiments[31, 32] which have been performed on the radiation and pair production spectrums shed no light on this point. The resolution in determining the secondary energies is too broad to resolve a rapid change in curvature. The experiments are not inconsistent with relatively uniform spectrums and non-zero cut-offs. In view of our

knowledge of non-relativistic radiation spectrums, it would seem wisest to modify the Bethe-Heitler calculations so that $g^{(1)}(u, v/u)$ and $g^{(2)}(u, v/u)$ are weak functions of $v/u$ when the Born approximation is no longer valid. The Bethe-Heitler values should be further normalized in accordance with attenuation experiments[33-38] and more accurate theory.[39]

Having justified equations (4.24), (4.25), one may now commence to wonder what the advantages are in obtaining them. Even though a formal solution, (2.41), (2.42), is immediately applicable, one can justly surmise that a numerical problem, perhaps insurmountable, lies ahead. Even though all of the objections to known solutions laid down at the end of Part 3.2 are now largely eliminated, it is still necessary to carry along all of the individual energy parameters throughout the calculations. The presence of the energy vector $\overline{E}_\nu$ is the one impractical element in equation (4.25). In the next part of this section it will be shown how the application of Theorem 2 will make the solution developed so far a practical reality. With the use of Theorem 2 and an assumption about our knowledge of a certain distribution, the equation (4.25) will be replaced with one identical in form except that the vector $\overline{E}_\nu$ and parameter $\mu$ become one parameter, the total energy, the summation signs disappear, and the cross sections become modified.

4.2 A Solution Based on Theorems 1, 2, and 5.

Instead of applying Theorem 2 in its abstract form to the solution developed in the preceding part of this section, it might be well to give it as much intuitive meaning as possible. At the same time the method of Part 4.1, generalized in Theorem 1, will be given similar broad meaning. In the terminology of Part 1.2, let $\Sigma$ represent the space of all possible states, S, in which a cascade shower might fall. For the present argument, consider the states, S, as points in the plane, though they are actually rather complicated vectors. Recall that these vector-points should actually contain angular and spatial components to describe the three-dimensional spreading of the low energy particles through the absorbing medium, though the corresponding parameters were not displayed in the definition of a

state, S, given in Part 1.2. (The incompleteness in the description of a shower as supplied by the parameters (n, m, $\bar{E}$, $\mu$) is just the basis for proceeding with the problem as follows.) As in Fig. 1 the development of a shower would normally progress from state to state and Fig. 1 would be an adequate illustration of a particular transition if one imagined that the state vectors, S, really contained all the components necessary for a complete description.

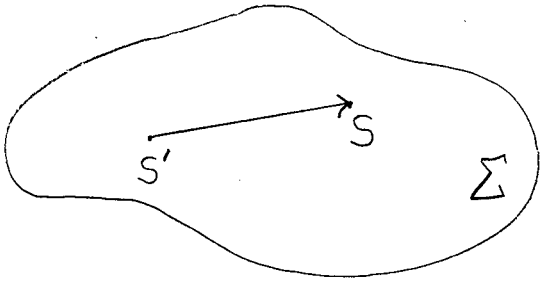



Fig. 1

The diffusion equation for a shower would involve transition probabilities, $\pi$ (S' → S), corresponding to the transition illustrated.

Due to the desire to avoid including all the low energy spatial parameters, it was found convenient in Part 4.1 to consider the transitions of a shower from set to set, as illustrated in Fig. 2.

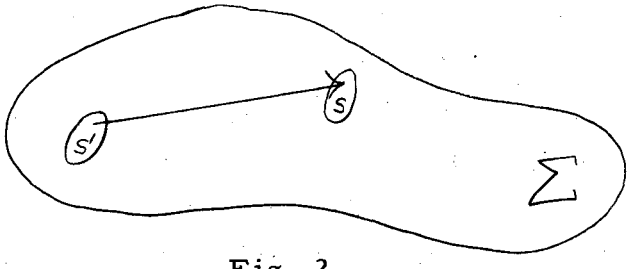



Fig. 2

The sets were chosen to include a certain number of high energy particles of specified energies and an arbitrary number of low energy particles. By choosing them in this way, the transition probabilities, $\pi$ (s' → s), for motion from set to set were not affected the changes taking place between low energy particles within the sets. In fact, the transition probabilities could be determined by the transition to s from any state S with s'. That is, the particular configuration of low energy particles within s' at the time of transition was immaterial. The sets chosen from $\Sigma$ in Part 4.1 were said to be <u>enchained</u> by the basic process of the cascade shower because the transitions of the system from set to set could be treated identically to the transitions of a system through a Markov chain of events (to use the terminology of a discrete space).

Suppose that a similar technique of solving for transitions from sets to sets be used now to eliminate, if possible, the individual energy parameters. If the method of sets worked for the low energy spatial parameters, it might work for others as well. One could keep the sets, s, that have already eliminated the need for spatial parameters, and group these into larger sets, r, according to the number of high energy particles present and the total energy represented by the high energy particles. See Fig. 3.

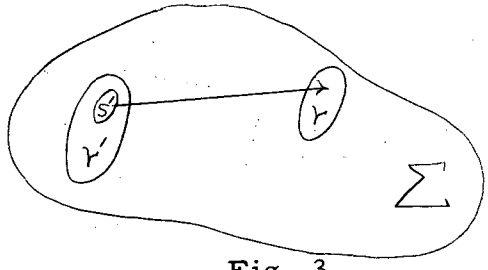



Fig. 3

Though it is too much to ask that, in analogy to the behavior of the transition between the sets, s, the probabilities for transitions between these larger sets will also remain unaffected by the configuration of particles within them, still the following will be true: If there is an a priori knowledge of the conditional probabilities corresponding to various high energy particle configurations within each set r (in other words, a knowledge of the energy distribution among high energy particles conditioned on the event that they belong to r), then the transition from a set, r', to a set, r, can be determined by averaging. One will take as $\pi(r' \to r)$ the average value of $\pi(s' \to r)$ for all s' in r', weighting each $\pi(s' \to r)$ according to the a priori conditional distribution. These ideas are just the substance of Theorem 2 intuitively expressed. The sets r, so chosen, are called conditionally enchained, because a probabilistic knowledge of the configurations that are conditioned to exist within each set permits the transition probabilities to be treated as before -- as though referring to a Markov chain.

In return to the actual problem, let all the notations be the same as in Part 4.1 and let all the assumptions of 4.1 hold. Let $r(n, m, T_\nu)$, $\nu = 0, 1, 2, \ldots$, $n = 1, 2, \ldots$ $m = 1, 2, \ldots$, $n+m > 0$, represent the set of all sets, $s(n, m, \bar{E}_\nu, \mu)$, for which $\bar{E}_\nu \cdot \bar{1} = \sum_{k=1}^{n+m} E_k = 1 - \nu\epsilon = T_\nu$.

Let $r^o = s^o$. Let $\Delta\bar{E}_\nu$ be an n+m-dimensional energy interval of the type appearing in the definition of $\Delta S$, Part 1.2. Assume that there is known the distribution of energies among the particles, given a knowledge of the number of electrons and photons and their total energy. Assume that this distribution is independent of x. In other words, for each $\Delta\bar{E}_\nu$ and $\mu$, the probability, $P(\Delta\bar{E}_\nu, \mu | n, m, T_\nu)$, that $\mathcal{S}(x)$ belongs to $\Delta S$, given a knowlege of n, m, and $T_\nu$ is known and is independent of x. Such a function corresponds to $P_4(\Delta G | g, g_o)$ that appears in Definition 2, property (iii). If it exists and is known, then the set, $\rho = \{r\}$, of sets r is conditionally enchained by $\mathcal{S}(x)$, because all properties of Definition 2 will be satisifed. The justification of the essential assumption about $P(\Delta E_\nu, \mu | n, m, T)$ will be left till later.

Let $\Delta r$ be defined as any set of the form:

$$\Delta r \equiv \{r(n', m', T') : T' \in (T, T + e), n' = n, m' = m\}, \quad e > 0.$$

With the use of Theorem 2, it is possible to obtain a diffusion equation for the transition function, $P(\Delta r, x | r_o, o)$, which in the general notation of that theorem has the form $P(\Delta g, x | g_o, o)$. The hypothesis of the theorem stating that the basic process is Markovian will be satisfied in the present case even though the sets, r, are sets of sets, s, and not sets of points, S. This is true because $\rho = \{r\}$ is conditionally enchained by a process ($\mathcal{F}(x)$ in the general notation of Definition 1) that Theorem 1 proved to be Markovian. As in Part 4.1, when the argument referred to the functon $P(\Delta s, x | s_o, o)$, it would be convenient to solve for a probability density. Again, however, there is a continuous and discontinuous part. The reasons are the same and the points of discontinuity are the same, namely, the zero-dimensional hyperplanes $H(1 - \nu\epsilon) = 1 - \nu\epsilon$, $\nu = 0, 1, 2, \ldots$ The definitions of the continuous and discontinuous parts of the transition function $P(\Delta r, x | r_o, o)$ are completely analogous to the corresponding definitions of Part 4.1. Consequently they will not be reproduced here. Since $s_1$ and $s_2$ belong to $\rho = \{r\}$, one can define $r_1 = s_1$, $r_2 = s_2$ and write:

$$f_1^{(j)}(0, 0, x) = p_1^{(j)}(0, 0, x) \quad , \tag{4.7A}$$

in correspondence to (4.7). Let $f_1^{(j)}(n, m, T, x)$ and $f_2^{(j)}(n, m, T, x)$, $n = 1, 2, \ldots$, $m = 1, 2, \ldots$, $n+m>0$, $T > \epsilon$, $x \geq 0$, be the continuous and discontinuous parts in the present case, corresponding to (4.8) and (4.9). $f_1^{(j)}(n, m, T_\nu, x)$ will be zero except perhaps for $\nu = 0, 1, 2, \ldots$ $f_2^{(j)}(n, m, T_\nu, x)$ can be defined in direct analogy to (4.18). Integrals similar to (4.20) - (4.23) will again appear in the equations for the continuous part and will have to be approximated through an assumption almost identical to Assumption VI. (See the discussion following the integrals (4.20)-(4.23).) The equations themselves that are satisfied by $f_i^{(j)}(n, m, T_\nu, x)$ are so much like (4.24) and (4.25) that they will be written down immediately along with the boundary conditions. Their solution follows the same pattern as that of (4.24), (4.25), even to the use of Table 1, Part 4.1. Suffice it only to say that the weighted cross sections, such as $\psi^{o(1)}$, that appear in place of the cross sections of (4.25), have an obvious construction out of the latter. Their general definition is given by (2.22). Argument variables seem unnecessary for it will be clear that each weighted cross section represents a transition from the set of states parametrized by the arguments appearing in the accompanying function. For example:

$$\psi^{o(1)} = \int_{r'(n,m-1,T_\nu)} \pi[\, s(n,m-1,\bar{E}_\nu,\mu) \to r(n,m,T_\nu)\,]\, P_4(d\sigma_s | r', r_o), \qquad (4.26)$$

where,

$$\pi[\, s(n,m-1,\bar{E}_\nu,\mu) \to r(n,m,T_\nu)\,] = \sum_{k=1}^{n} \int_{\epsilon}^{E_k - \epsilon} d\upsilon\, \psi^{(1)}(E_k, \upsilon/E_k).$$

The equation follows:

$$f_i^{(j)}(n,m,T_\nu,0) = \delta(n+j-2)\,\delta(m+1-j)\,\delta(\nu)\,\delta(i-1), \qquad (4.27)$$

$$f_i^{(j)}(n,m,T_\nu,x) = 0, \quad \text{For all } x \geq 0, \text{ if } \nu < 0 \text{ or } n < 0 \text{ or } m < 0.$$

$$(\partial/\partial x) f_1^{(j)}(0,0,x) = \sum_{r=1}^{2} f_r^{(j)}(1,0,1/\epsilon-2,x) \sum_{s=1}^{2} \varphi_s^{(1)}(2\epsilon)[\delta(r-1)+\tfrac{\epsilon}{2}\delta(r-2)]$$
$$+ f_2^{(j)}(1,0,1/\epsilon-1,x) \sum_{s=1}^{2} \varphi_s^{(1)}(\epsilon)$$
$$+ f_2^{(j)}(0,1,1/\epsilon-1,x)\, \tfrac{\epsilon}{2}\, \varphi_1^{(2)}(\epsilon) . \qquad (4.28)$$

$$(\partial/\partial x + \alpha(n,m,T_\nu))\, f_i^{(j)}(n,m,T_\nu,x) \qquad (4.29)$$
$$= f_i^{(j)}(n,m-1,T_\nu,x)\, \psi^{(1)}$$
$$+ f_i^{(j)}(n-2,m+1,T_\nu,x)\, \psi^{(2)}$$
$$+ f_i^{(j)}(n-1,m+1,T_\nu,x)\,[\varphi_2^{(2)} + \tfrac{\epsilon}{2}\psi^{(2)}\delta(i-2)]$$
$$+ f_i^{(j)}(n,m,T_{\nu-1},x) \sum_{s=1}^{2} \varphi_s^{(1)}$$
$$+ \Big\{ \sum_{r=1}^{2} f_r^{(j)}(n-1,m+1,T_{\nu-1},x)\, \psi^{(2)}\,[\delta(r-1)+\tfrac{\epsilon}{2}\delta(r-2)]$$
$$+ \sum_{r=1}^{2} f_r^{(j)}(n+1,m-1,T_{\nu-1},x)\,[\psi_1^{(1)}\delta(r-1) + \psi_2^{(1)}\delta(r-2)]$$
$$+ \sum_{r=1}^{2} f_r^{(j)}(n+1,m,T_{\nu-2},x) \sum_{s=1}^{2} \varphi_s^{(1)}\,[\delta(r-1)+\tfrac{\epsilon}{2}\delta(r-2)]$$
$$+ f_2^{(j)}(n+1,m,T_{\nu-1},x)\, \tfrac{\epsilon}{2} \sum_{s=1}^{2} \varphi_s^{(1)}\,[1-\delta(\nu-1)]$$
$$+ f_2^{(j)}(n,m+1,T_{\nu-1},x)\, \tfrac{\epsilon}{2}\, \varphi_1^{(2)}\,[1-\delta(\nu-1)] \Big\}\delta(i-2) .$$

The final link in the justification of equations (4.27)-(4.29) concerns the weighting function $P(\Delta\bar{E}_{\nu}, \mu | n, m, T)$, its existence and our cognizance of it. Its existence depends only on whether such a function exists that is independent of x, for in the proof of Theorem 2 it is seen that there would always be such a function which depended on x. That is, the conclusion of Theorem 2 could be reached without a condition on the weighting function provided the weighted cross sections (2.22) in the conclusion were x-dependent. From this point of view, the existence of the function must be postulated. One can only say that, intuitively, it appears to be independent of x. If it is not, it can be substituted for one that is, and the answer will suffer an approximation from this source.

Our knowledge of the function will come from experience, previous exact calculations (such as those of Part 4.1), or experiment and will not be deeply probed here. The main purpose of this part of Section 4 is to show that it is not necessary to carry all the energy parameters throughout all the calculations. If one carried the energy parameters through one simple case by the method of Part 4.1, sufficient information would be had about the weighting function to facilitate any further calculations. Another purpose of this Part is to demonstrate by example the potential use of any detailed knowledge conditioned on partial knowledge that relates to the energy distribution among particles. For instance, it was purely for simplicity that the number of particles and their total energy were chosen as the only parameters from which to infer the energy distribution. It would have been even more realistic to choose four parameters, the number of electrons, the number of photons, the total energy of the electrons and that of the photons. The generalities of Theorem 2 permit any such modification of the method presented here.

It is worth noticing the kinship between the method of this Part and that suggested by the present author[40] to simplify the earlier matrix solution.[27] The present method rests on stronger theoretical grounds and is more effective in eliminating unwanted parameters.

## 5. SUMMARY

In the Introduction the basic problem of electron-photon showers was introduced. It was shown how the collision-loss process was a first cause for mathematical difficulties. The behavior of the low energy particles was a second cause. In Part 4.1 the first of these difficulties was eliminated by a convenient idealization of the physical process of collision-loss and the second difficulty was overcome by partitioning the space of state vectors so as to make the presence of low energy parameters unnecessary. After reviewing the known solutions to the problem in Section 3, it was noticed that large simultaneous sets of equations had always been used and asymptotic cross sections had been necessary in order to permit integral transform methods. In Part 4.1 a slight idealization of the radiation process and an approximation to an integral made the simultaneous equations unnecessary. A general solution developed in Section 2 could be immediately applied to yield a direct answer for the probabilities in terms of a linear function of exponentials. At the same time, the presence of the idealized collision-loss process and the partitioned space mentioned above made it possible to use energy dependent cross sections rather than asymptotic ones. Multiple scattering and Compton effect could be corrected for so that the solution of Part 4.1 was valid for showers initiated by high energy primaries and medium energy primaries of about 70 Mev in lead.

As in most of the solutions examined in Section 3, the solution of Part 4.1 contained all of the individual energy parameters of each particle. Though the results of Part 4.1 were desirable from many points of view, they were somewhat impractical for this reason. The use of Theorem 2 made it possible to eliminate all but one of the energy parameters and thus bring Part 4.1 onto a realistic ground. The application of Theorem 2 was made in Part 4.2 and was reminiscent of the application of Theorem 1 in Part 4.1: A partitioned space was used once again to eliminate unwanted parameters.

Throughout the paper an awareness of other techniques was kept alive by the use of terminology and ideas that were developed in Section 2. Section 2 was responsible for holding some general logic that lay beneath all previous methods and that could be extended to the framework of Section 4.

## REFERENCES


1. H. J. Bhabha and W. Heitler, Proc. Roy. Soc. (London) A159, 432 (1937).

1.1. J. F. Carlson and H. R. Oppenheimer, Phys. Rev. 51, 220 (1937).

2. Bruno Rossi, "High Energy Particles", Prentice-Hall, New York (1952).

3. A. Nordsieck, W. E. Lamb and G. E. Uhlenbeck, Physica 7, 344 (1940).

4. W. T. Scott and G. E. Uhlenbeck, Phys. Rev. 62, 497 (1942).

5. H. J. Bhabha and A. Ramakrishnan, Proc. Indian Acad. Sci. 32, 141 (1950).

6. L. Janossy and H.Messel, Proc. Phys. Soc. (London) A63, 1101 (1950).

7. Alladi Ramakrishnan and P. M. Mathews, Prog. Theor. Phys. 11, 95 (1954).

8. H. J. Bhabha, Proc. Roy Soc. (London) A202, 301 (1950).

9. H. J. Bhabha, Proc. Indian Acad. Sci. 32, 154 (1950).

10. A. Ramakrishnan, Proc. Cambridge Phil. Soc. 46, 595 (1950).

11. A. Ramakrishnan, Proc. Cambridge Phil. Soc. 48, 451 (1952).

12. L. Janossy, Proc. Phys. Soc. (London) A63, 241 (1950).

13. L. Janossy, Proc. Phys. Soc. (London) A66, 117 (1953).

14. L. Janossy, Acta Phys. Hung. 2, 289 (1952).

15. W. T. Scott, Phys. Rev. 82, 893 (1951).

16. H. Messel and R. B. Potts, Phys. Rev. 86, 847 (1952).

17. H. Messel and R. B. Potts, Phys. Rev. 87, 759 (1952).

18. Niels Arley, "On the Theory of Stochastic Processes and Their Application to the Theory of Cosmic Radiation", G.E.C. Gads Forlag, Copenhagen (1948).

19. W. H. Furry, Phys. Rev. 52, 569 (1937).

20. R. R. Wilson, Phys. Rev. 86, 261 (1952); Phys. Rev. 86, 590 (1952).

21. W. Feller, Trans. Amer. Math. Soc. 48, 488 (1940).

22. M. S. Bartlett and D. G. Kendall, Proc. Cambridge Phil. Soc. 47, 65 (1951).
23. D. G. Kendall, J. Roy. Statist. Soc. (B) 11, 230 (1949).
24. C. Palm, Ericsson Technics, No. 44, 1, (1943).
25. R. Bellman and T. E. Harris, Proc. Nat. Acad. Sci., Washington, 34, 601 (1948).
26. Alladi Ramakrishnan, J. Roy. Statist. Soc. (B) 13, 131 (1951).
27. Bayard Rankin, Univ. of Calif. Rad. Lab. Report, No. 2322.
28. B. Rossi and K. Greisen, Rev. Mod. Phys. 13, 240 (1941).
29. e.g., W. Heitler, "The Quantum Theory of Radiation", second edition, Oxford University Press (1949).
30. H. A. Bethe and L. C. Maximon, Phys. Rev. 93, 768 (1954).
31. W. Hartsough, M. Hill and W. M. Powell, Phys. Rev. 81, 213 (1951).
32. J. W. DeWire and L. A. Beach, Phys. Rev. 83, 476 (1951).
33. G. D. Adams, Phys. Rev. 74, 1707 (1948).
34. J. L. Lawson, Phys. Rev. 75, 433 (1949).
35. R. L. Walker, Phys. Rev. 76, 527 (1949).
36. J. W. DeWire, A. Ashkin and L. A. Beech, Phys. Rev. 83, 505 (1951).
37. A. I. Berman, Phys. Rev. 85, 774 (1952).
38. E. S. Rosenblum, E. F. Shrader and R. M. Warner, Phys. Rev. 88, 612 (1952).
39. Handel Davies, H. A. Bethe, and L. C. Maximan, Phys. Rev. 93, 788 (1954).
40. Bayard Rankin, Phys. Rev. 95, 657 (A) (1954).