RELEASED FOR ANNOUNCEMENT
IN NUCLEAR SCIENCE ABSTRACTS

ORNL-RSIC-14

Contract No. W-7405-eng-26

Neutron Physics Division

THE EXPONENTIAL TRANSFORM AS AN IMPORTANCE-SAMPLING DEVICE

— A REVIEW —

Francis H. Clark

Note:

This work partially supported by
U. S. ARMY BALLISTIC RESEARCH LABORATORIES
under AEC-BRL Agreement 40-40-64

JANUARY 1966

OAK RIDGE NATIONAL LABORATORY
Oak Ridge, Tennessee
operated by
UNION CARBIDE CORPORATION
for the
U.S. ATOMIC ENERGY COMMISSION

LEGAL NOTICE

This report was prepared as an account of Government sponsored work. Neither the United States, nor the Commission, nor any person acting on behalf of the Commission:
A. Makes any warranty or representation, expressed or implied, with respect to the accuracy, completeness, or usefulness of the information contained in this report, or that the use of any information, apparatus, method, or process disclosed in this report may not infringe privately owned rights; or
B. Assumes any liabilities with respect to the use of, or for damages resulting from the use of any information, apparatus, method, or process disclosed in this report.
As used in the above, "person acting on behalf of the Commission" includes any employee or contractor of the Commission, or employee of such contractor, to the extent that such employee or contractor of the Commission, or employee of such contractor prepares, disseminates, or provides access to, any information pursuant to his employment or contract with the Commission, or his employment with such contractor.

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

b
D

12

TABLE OF CONTENTS

	PAGE
Abstract	1
I. What Is the Transform	1
II. Development and Background	3
III. Solutions Incorporating the Exponential Transform	6
IV. Forms of the Transform	9
V. Tracking	13
VI. Some Difficulties Associated with the Transform	15
VII. Variations from the Exponential Transform	21
VIII. Some Results	23
Acknowledgement	37
References	38
APPENDIX A. Space Integral of the Boltzmann Equation	40
APPENDIX B. Spatial Discontinuities in g	42
APPENDIX C. Equivalence of the Importance-Sampled Boltzmann Solution and the Transformed Solution	43
APPENDIX D. Choosing from the Track Length Distributions	46

1
2
3

4
5

6
7

THE EXPONENTIAL TRANSFORM AS AN IMPORTANCE-SAMPLING DEVICE

— A REVIEW —

Francis H. Clark

ABSTRACT

The exponential transform is reviewed, with emphasis on its use as a guide to effective importance sampling in the solution of the Boltzmann equation by Monte Carlo methods. Contributions of various workers are assembled, along with numerical results. Special consideration is given to approximate forms and to effective practical methods. Problems related to negative effective cross sections, tracking across discontinuities, directional biasing in inhomogeneous media, and high variance in back-scattered components are specifically treated.

I. WHAT IS THE TRANSFORM

The exponential transform is one of numerous devices employed to reduce variance in Monte Carlo solutions of the Boltzmann equation. It tends to promote transport of particles in preferred directions and discourage transport in directions opposed to the preferred. [The preferred direction might be a function of the spatial coordinates.]

The Boltzmann equation is

$$\vec{\omega} \cdot \nabla N(\vec{r}, \vec{\omega}, E) + \sigma(\vec{r}, E) N(\vec{r}, \vec{\omega}, E) = \sum_{i=1}^I \int \int \sigma(\vec{r}, E') \alpha_i(E')$$

$$\left[\frac{\sigma_{si}(\vec{r}, E')}{\sigma(\vec{r}, E')} \right] K_i(E', \vec{\omega}' \rightarrow E, \vec{\omega}) N(\vec{r}, \vec{\omega}', E') dE' d\vec{\omega}' + S(\vec{r}, \vec{\omega}, E) \quad , \quad (1)$$

where

$\vec{r}, \vec{\omega}$, and E are the independent variables, phase space coordinates denoting location, direction, and energy (in particular, $\vec{\omega}$ is a unit vector),

N = particle flux density in phase space ,

$\sigma(\vec{r}, E)$ = total macroscopic cross section ,

$\sigma_{si}(\vec{r}, E)$ = macroscopic reaction cross section for i^{th} material,*

$\alpha_i(E')$ = number of particles emerging from i^{th} process per particle entrant with energy E' ,

$K_i(E', \vec{\omega}' \rightarrow E, \vec{\omega})$ = probability density for change of coordinates from $E', \vec{\omega}'$ to $E, \vec{\omega}$ in i^{th} process ,

$S(\vec{r}, \vec{\omega}, E)$ = source density.

Let us now define

$$\varphi(\vec{r}, \vec{\omega}, E) = N(\vec{r}, \vec{\omega}, E) e^{g(\vec{r}, \vec{\omega}, E)} \quad . \quad (2)$$

We then have

$$\vec{\omega} \cdot \nabla \varphi(\vec{r}, \vec{\omega}, E) + \left[\sigma(\vec{r}, E) - \vec{\omega} \cdot \nabla g(\vec{r}, \vec{\omega}, E) \right] \varphi(\vec{r}, \vec{\omega}, E) =$$

$$\sum_{i=1}^I \int \int \left[\sigma(r, E') - \vec{\omega} \cdot \nabla g(\vec{r}, \vec{\omega}', E') \right] \alpha_i(E')$$

$$\left[\frac{\sigma_{si}(\vec{r}, E')}{\sigma(\vec{r}, E') - \vec{\omega} \cdot \nabla g(\vec{r}, \vec{\omega}', E')} \right] e^{g(\vec{r}, \vec{\omega}, E) - g(\vec{r}, \vec{\omega}', E')}$$

$$K_i(E', \vec{\omega}' \rightarrow E, \vec{\omega}) \varphi(\vec{r}, \vec{\omega}', E') dE' d\vec{\omega}' + S(\vec{r}, \vec{\omega}, E) e^{g(\vec{r}, \vec{\omega}, E)} \quad . \quad (3)$$

We shall often refer to the coefficient of φ in the second term on the left-hand side of (3) as σ^* or an "effective" total cross section.

*Each distinct reaction is considered to define a separate material.

Equation (2), along with the transformed Eq. (3), comprises the exponential transform. The transforming function $g(\vec{r}, \vec{\omega}, E)$ is allowed to have an arbitrary dependence on $\vec{\omega}$ and E . It is required to be differentiable in the spatial coordinates everywhere except across separated closed surfaces,* where it can have finite discontinuities. Such discontinuities will contribute to the complexity of the weight functions associated with the track distributions derived in Appendix A. The treatment of φ at such surfaces is given in Appendix B.

II. DEVELOPMENT AND BACKGROUND

The exponential transform was first proposed by Herman Kahn in a general exposition of the Monte Carlo method in application to transport problems.¹ Kahn used the specialized form

$$g(\vec{r}, \vec{\omega}, E) = \vec{\epsilon} \cdot \vec{r} C, \quad (4)$$

$$\vec{r} = \vec{i}x + \vec{j}y + \vec{k}z, \quad (5)$$

where

$\vec{\epsilon}$ = some specified constant unit vector,

C = constant governing the magnitude of the transforming factor.

Form (4) brings out the principal features of the transform but does not demonstrate its full flexibility. The vector $\vec{\epsilon}$ is in the direction in which one is interested in studying the particle transport. If there is no such direction, (4) is inappropriate. With definition (4) Eq. (3) is specialized in the following two ways:

$$e^{g(\vec{r}, \vec{\omega}, E)} = e^{\vec{\epsilon} \cdot \vec{r} C}, \quad (6)$$

$$\vec{\omega} \cdot \nabla g(\vec{r}, \vec{\omega}, E) = \vec{\omega} \cdot \vec{\epsilon} C, \quad (7)$$

$$-1 \leq \vec{\omega} \cdot \vec{\epsilon} \leq 1. \quad (8)$$

*Or open surfaces which extend to infinity.

It is apparent from (3) that a negative "effective" total cross section can result whenever

$$|C| > \sigma(\vec{r}, E) \quad . \quad (9)$$

This problem has been considered by a number of workers.

Goertzel and Kalos² in a review article on Monte Carlo reformulated the exponential transform with some particular attention to the negative "effective" total cross section.

Steinberg and Liedtke in describing a gamma-ray Monte Carlo transmission program³ examined various aspects of importance sampling and made the important observation that many kinds of track length sampling schemes closely approximated the exponential transform in effect. That is, not only might the exponential transform be effected by reasonably simple importance-sampling schemes, but further simplifications can be made by using rougher and simpler importance-sampling schemes. It is well known that in many instances an elaborate distribution can be replaced by a simpler one retaining the essential features of the elaborate scheme without substantial loss of effectiveness.

Perkins and Burrell⁴ have produced a gamma-ray Monte Carlo program incorporating the exponential transform and have published results giving some indication of its effectiveness. They suggest an approach to prove the equivalence of the exponential transform to certain importance-sampling schemes and they incorporate the transform by importance sampling. Scheme (4) is used and called the additive method. It is compared with another scheme, which they called multiplicative. In the latter scheme, when there are no spatial discontinuities in the total cross section (e.g., a single medium), g is represented by

$$g(\vec{r}, \vec{\omega}, E) = \sigma(\vec{r}, E) \vec{e} \cdot \vec{r} \quad C \quad . \quad (10)$$

For systems where there are cross-section discontinuities, the multiplicative scheme as used does not strictly correspond to any g and is therefore not truly, in such cases, an exponential transform. But it is an excellent approximation to the transform.

Leimdorfer has written a series of notes containing contributions to the useful development of the transform. He has given an explicit proof, applicable to many forms of the function $g(\vec{r}, \vec{\omega}, E)$, of the equivalence of the transformed equation to an importance-sampled solution of the Boltzmann equation.⁵ He has offered two ways of dealing with negative "effective" total cross sections,^{5,6} of which one opens the way to much stronger biasing ($C > 1$) than had been used previously. Leimdorfer has used the transform to improve the computation of reflection⁷ as well as transmission. In addition to the commonly used form (4) he has experimented with an approximately "spherically outward" form of the transform.⁸ That is, he has used a form given by

$$g(\vec{r}, \vec{\omega}, E) = \sqrt{x^2 + y^2 + z^2} C, \quad (11)$$

although his track computation procedure was approximate. (Kalos, too, has experimented with such a form.⁹) All of these features are incorporated in machine programs, for which results are liberally quoted in the above cited references.

Eriksson¹⁰ has recent Leimdorfer's work in a greater formality, making some additions. He has somewhat generalized the proof of the equivalence of the transform to an importance-sampling scheme, and has offered the scheme

$$g(\vec{r}, \vec{\omega}, E) = - \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2} C \quad (12)$$

for directing effort toward the point (x_0, y_0, z_0) .

Beach, Theus, Plawchan, and Faust produced one of the earliest programs which incorporated the exponential transform¹¹ using scheme (4). Later work by Beach and Pieper¹² incorporated scheme (10).

Loechler and MacDonald describe an approximate exponential transform capability in the FMC program¹³ and Wells describes sampling schemes similar to those of Liedtke and Steinberg, i.e., approximations to the transform,

in the C-18 program.¹⁴ We do not have available results which isolate the effect of the transform in these programs.

III. SOLUTIONS INCORPORATING THE EXPONENTIAL TRANSFORM

Equation (1) can be solved by standard Monte Carlo procedures as follows. Initial coordinates of a history are chosen randomly from the source distribution

$$\frac{S(\vec{r}, \vec{\omega}, E)}{\int S(\vec{r}', \vec{\omega}', E') d\vec{r}' dE' d\vec{\omega}'} ; \quad (13)$$

track lengths are chosen randomly from the distribution

$$\sigma(s, E) e^{-\int_0^s \sigma(\xi, E) d\xi} , \quad (14)$$

$$\vec{r}_i = \vec{r}_{i-1} + \vec{\omega} s , \quad (15)$$

where $\sigma(s, E)$ is shorthand notation for $\sigma(\vec{r}_{i-1} + \vec{\omega}s, E)$ and $\vec{r}_j$ is the location of j^{th} collision point of the history. Directions and energies after a collision are determined by random selection from the distribution

$$K_\ell(\vec{E}', \vec{\omega}' \rightarrow E, \vec{\omega}) \quad (16)$$

the index ℓ itself determined by random selection from the distribution function

$$L(\ell) = \sum_{\lambda=1}^I \frac{\sigma_{s\lambda}(\vec{r}, E)}{\sigma(\vec{r}, E)} \delta(\lambda - \ell) \quad . \quad (17)$$

At each collision point the weight is adjusted by the factor

$$\alpha_i(E') \quad . \quad (18)$$

Equation (3) can be solved in an entirely analogous fashion by taking for the source distribution

$$\frac{S(\vec{r}, \vec{\omega}, E) e^{g(\vec{r}, \vec{\omega}, E)}}{\int S(\vec{r}', \vec{\omega}', E') e^{g(\vec{r}', \vec{\omega}', E')} d\vec{r}' dE' d\vec{\omega}'} \quad , \quad (19)$$

and for the track length distribution (see Appendix A)

$$\left[\sigma(\vec{r} + \vec{\omega}s, E) - \frac{dg(\vec{r} + \vec{\omega}s, \vec{\omega}, E)}{ds} \right] e^{-\int_0^s \left[\sigma(\vec{r} + \vec{\omega}\xi, E) - \frac{dg(\vec{r} + \vec{\omega}\xi, \vec{\omega}, E)}{d\xi} \right] d\xi} \quad . \quad (20)$$

The kernel (16) is used for choosing direction and energy, and (17) for interacting material. The weight factor becomes

$$\alpha_i(E) \frac{\sigma(\vec{r}, E) e^{[g(\vec{r}, \vec{\omega}^*, E) - g(\vec{r}, \vec{\omega}, E)]}}{\sigma(\vec{r}, E) - \vec{\omega} \cdot \nabla g(\vec{r}, \vec{\omega}, E)} \quad (21)$$

(quantities marked * are after-collision coordinates). It is sometimes necessary to multiply (20) by a normalization factor $1/M$. In such cases an additional factor M appears in (21). (See Appendix A.)

We are, of course, at liberty to employ importance sampling. That is, where the problem in some particular formulation calls for a random

selection of $\vec{\eta}$ from the probability distribution function $p_1(\vec{\eta})$, we elect instead to choose $\vec{\eta}$ from the probability distribution $p_2(\vec{\eta})$ and modify the weight of the sample item by the factor

$$w_2(\vec{\eta}) = \frac{p_1(\vec{\eta})}{p_2(\vec{\eta})} . \quad (22)$$

The construction of $p_2(\vec{\eta})$ is restricted by the requirement that $p_2(\vec{\eta})$ be nonzero wherever $p_1(\vec{\eta})$ is nonzero.

Let us now consider a solution of (1) with the following scheme of importance sampling. Initial history coordinates are chosen from (19) with weight adjustment given by the ratio of (13) to (19); track lengths from (20) with weight adjustment in the ratio of (14) to (20); energies and directions after scattering from (16); reaction index ℓ from (17); and multiplicity weight adjustment from (18).

The solution to (1) so produced is identical to the solution to (3) already described. For if these procedures are followed, and (1) is solved with source normalization

$$\int S(\vec{r}, \vec{\omega}, E) d\vec{r} dE d\vec{\omega} ,$$

and (3) is solved with source normalization

$$\int S(\vec{r}, \vec{\omega}, E) e^{g(\vec{r}, \vec{\omega}, E)} d\vec{r} d\vec{\omega} dE$$

and with the same set of random numbers, and, further, if the collision densities of the φ variable (i.e., weights) computed in the solution of (3) at every point in phase space where an event occurred are transformed to collision densities for the N variable by means of Eq. (2), this set of densities will have weights and coordinates identical with those of the set computed in the importance-sampled solution of (1). This identity is proved in Appendix C.

IV. FORMS OF THE TRANSFORM

It is shown in Appendices B and C that for quite general functions $g(\vec{r}, \vec{\omega}, E)$ we obtain the properties of the transform developed up to this point, i.e., that it be meaningful over the whole domain, and that the importance-sampled Boltzmann solution be equivalent to the transformed solution. We shall now consider the advantages of specialized forms.

Let us assume that $\sigma(\vec{r}, E)$ may change discontinuously at certain boundaries but remains constant within the contained volumes. Most actual systems conform to this assumption. For such systems consider the form

$$g(\vec{r}, \vec{\omega}, E) = \sigma(\vec{r}, E) \nu(\vec{r}, E) f(\vec{r}) \quad . \quad (23)$$

We shall require that $\nu(\vec{r}, E)$ also be a spatial constant within boundaries but may be discontinuous at boundaries. The effective total cross section is, in general,

$$\sigma^*(\vec{r}, \vec{\omega}, E) = \sigma(\vec{r}, E) - \vec{\omega} \cdot \nabla g(\vec{r}, \vec{\omega}, E) \quad (24)$$

and, for the restricted form (23), it is

$$\sigma^*(\vec{r}, \vec{\omega}, E) = \sigma(\vec{r}, E) [1 - \nu(\vec{r}, E) \vec{\omega} \cdot \nabla f(\vec{r})] \quad (25)$$

except at boundaries of discontinuity.

One advantage of form (25) is that it is easy to choose ν and f such that

$$\nu \omega \cdot \nabla f < 1 \quad , \quad (26)$$

assuring that

$$\sigma^* > 0 \quad . \quad (27)$$

Forms other than (25) often lead to negative values of σ^* . Dealing with, or avoiding, such negative values has been one of the larger problems faced by users of the transform. That problem will be discussed in a later section.

Another advantage of (25) is that it is simply a multiplicative correction on σ . It is sometimes much simpler to alter an existing program to accept such a multiplicative change than it might be for an additive change.

Perkins and Burrell⁴ have made some efficiency comparisons, defining g on the one hand by (4), leading to

$$\sigma^* = \sigma - c \vec{\omega} \cdot \vec{\epsilon} , \quad (28)$$

and on the other hand by

$$g(\vec{r}, \vec{\omega}, E) = \sigma(\vec{r}, E) M \vec{\epsilon} \cdot \vec{r} , \quad (29)$$

a special form of (23) which leads to

$$\sigma^* = \sigma(1 - M \vec{\omega} \cdot \vec{\epsilon}) . \quad (30)$$

On the basis of the comparison made, these workers were unable to determine that either method led to clearly better results. In dealing with specialized forms we shall most often use (25). Where g has discontinuities serious tracking problems arise, and, of course, in a multi-medium problem form (25) will introduce discontinuities into g . This problem will be discussed in the next section.

We shall consider two analytic forms for $f(\vec{r})$. They are

$$f(\vec{r}) = \vec{\epsilon} \cdot \vec{r} , \quad (31)$$

where $\vec{\epsilon}$ = constant unit vector, and

$$f(\vec{r}) = \lambda \sqrt{(x - x_D)^2 + (y - y_D)^2 + (z - z_D)^2} \quad , \quad (32)$$

where

x_D, y_D, z_D = some specified point (which may be a function of the particle coordinates at the collision point), and
 $\lambda = \pm 1$.

These forms closely resemble g forms already introduced in (4), (10) and (12). They imply, respectively,

$$\vec{\omega} \cdot \nabla f = \vec{\omega} \cdot \vec{e} \quad (33)$$

and

$$\vec{\omega} \cdot \nabla f = \frac{\lambda^2 \vec{\omega} \cdot \vec{t}}{f} = \vec{\omega} \cdot \left(\frac{\lambda \vec{t}}{|\vec{t}|} \right) \quad (34)$$

$$\vec{t} = (x - x_D) \vec{i} + (y - y_D) \vec{j} + (z - z_D) \vec{k} \quad (35)$$

$$f = \lambda |\vec{t}| \quad . \quad (36)$$

Form (31) emphasizes transport in the fixed direction $\vec{e}$. Form (32) emphasizes transport toward (or away from, depending on the sign of λ) some point (x_D, y_D, z_D) . Note that (x_D, y_D, z_D) is not required to be a fixed point but is permitted to be a function of the collision point coordinates. In taking the gradient of f , however, we consider (x_D, y_D, z_D) a fixed point and introduce its dependence upon collision point coordinates only after differentiation. Hence, a variety of situations may be served by the one form (32). A few are shown in Table 1.

The factor ν governs the magnitude of the transformation. For the time being we shall require that

$$0 \leq \nu < 1 \quad . \quad (37)$$

(This restriction can be relaxed later.) $\nu = 0$ corresponds to no transform

Table 1

Choice of f-Coordinates Leading to Certain Biases

Bias	x_D	y_D	z_D	λ
Toward a point (x_0, y_0, z_0)	x_0	y_0	z_0	-1
Away from a point (x_0, y_0, z_0)	x_0	y_0	z_0	+1
Toward a spherical shell of radius R centered on origin*	$\frac{R}{r}x$	$\frac{R}{r}y$	$\frac{R}{r}z$	-1
Away from the above shell*	$\frac{R}{r}x$	$\frac{R}{r}y$	$\frac{R}{r}z$	+1
Toward a cylindrical shell* of radius ρ centered on the Z-axis	$\frac{\rho}{q}x$	$\frac{\rho}{q}y$	z	-1
Away from the above shell**	$\frac{\rho}{q}x$	$\frac{\rho}{q}y$	z	+1
Toward a ring specified by $z = z_0, x^2 + y^2 = D^2$ †	$\frac{D}{q}x$	$\frac{D}{q}y$	z_0	-1
Away from the above ring†	$\frac{D}{q}x$	$\frac{D}{q}y$	z_0	+1

$$r^2 = x^2 + y^2 + z^2$$

$$q^2 = x^2 + y^2$$

$$* \text{ at } (x, y, z) = (0, 0, 0) \quad \vec{w} \cdot \nabla f = -\lambda$$

$$** \text{ at } (0, 0, z) \quad \vec{w} \cdot \nabla f = -\lambda$$

$$\dagger \text{ at } x^2 + y^2 = D^2, z = z_0; \quad \vec{w} \cdot \nabla f = -\lambda$$

($g = 1$, $\varphi = N$); while $\nu \rightarrow 1$ implies a very substantial amount of importance sampling.

The general effect of the transform, especially in form (23), on track selection is now clear. ∇f is a field of unit vectors which, at each point of the domain, represents the preferred direction of transport. $\vec{\omega}$ is the unit vector in the direction of neutron motion. We therefore have

$$-1 \leq \vec{\omega} \cdot \nabla f \leq 1 \quad . \quad (38)$$

Form (25) shows very clearly that the maximum cross-section reduction, $\nu\sigma$, occurs when the neutron flight is in the preferred direction ($\vec{\omega} \cdot \nabla f = 1$); the maximum increase, $\nu\sigma$, when the neutron flight is opposed to the preferred direction ($\vec{\omega} \cdot \nabla f = -1$); and that intermediate adjustments occur for intermediate inclinations of the flight to the preferred direction.

It is apparent that for flight length selection ∇f , not f , is the fundamental quantity. Only in the relatively less important biasing of the source term (19) does f enter directly.

V. TRACKING

In Appendix A it is shown that, in equations of the form of (1), independent of the sign or form of the quantity $\sigma^*(\vec{r}, E)$, distributions of track lengths between collision points, from source point to first collision point or from an intermediate point on the track to the next collision point, are governed by distributions of the form

$$L(s) = \frac{\left| \sigma^*(\vec{r}_0 + \vec{\omega}s, E) \right| e^{-\int_0^s \sigma^*(\vec{r}_0 + \vec{\omega}u, E) du}}{\int_0^T \left| \sigma^*(\vec{r}_0 + \vec{\omega}s, E) \right| e^{-\int_0^s \sigma^*(\vec{r}_0 + \vec{\omega}u, E) du} ds} \quad , \quad (39)$$

$$0 \leq s \leq T \quad (40)$$

provided that with restriction (40) the denominator of (39) has a value.

When σ^* does not change sign it, is instructive to set

$$y(s) = \int_0^s \sigma^*(\vec{r}_0 + \vec{u}u, E) du \quad (41)$$

$$L^*(y) = \frac{e^{-y}}{\int_0^T e^{-y(s)} ds} \quad (42)$$

A value of the (dimensionless) parameter y is selected randomly from the distribution (42) and is interpreted as a length through the solution of (41) for s .

It very commonly happens that σ^* is constant except at boundaries of discontinuity. In such cases one often determines a set of y_i , $i = 1, 2 \dots n$ such that

$$y = \sum_{i=1}^n y_i + y' \quad , \quad (43)$$

with

y_1 = distance along track from last collision to first boundary,

y_j = distance along track from $(j - 1)^{st}$ boundary to j^{th} boundary.

$y_i = 0$ if track does not reach i^{th} boundary,

y' = excess of track length remaining after crossing n^{th} boundary (y' is insufficient to reach $n+1^{st}$ boundary).

All y's are in dimensionless units--number of mean free paths. One then replaces (41) with

$$y' = \int_0^{s'} \sigma_{i+1} du = \sigma_{i+1} s' \quad (44)$$

and easily finds s' , the track length in length units into the region in which the next collision occurs. The same simple procedure can be used in the exponential integral solution with σ^* replacing σ . The treatment of boundaries of discontinuity in g is considered in Appendix B.

VI. SOME DIFFICULTIES ASSOCIATED WITH THE TRANSFORM

The problem which has most persistently troubled users of the exponential transform is the possibility that g will have been so chosen that for some values of σ form (24) will be negative or zero, i.e.,

$$\sigma^* = \sigma - \vec{w} \cdot \nabla g \leq 0 \quad . \quad (45)$$

Some users have restricted the choice of g , independent of σ , so that form (24) is positive for the minimum value of σ and therefore for all σ .^{4,9,10} This device places undue restriction on one's freedom in choosing g .

In a single-medium system one can with no complications use the restricted form (23) for g which leads to form (25) for σ^* . There is no difficulty in maintaining this form positive definite. Use of forms (23) and (25) will make it easy to avoid the problem of negative cross sections in a multimediuim system also. However, the discontinuities in σ^* lead to complications in tracking discussed in the previous section and in Appendix B. It is nonetheless often worthwhile to use forms (23) and (24) and deal with tracking complications as they arise.

It is possible to define a cross section, σ_N , for the process of doing nothing at all, i.e.,

$$\sigma_N(\vec{r}, E) N(\vec{r}, \vec{\omega}, E) = \int \int \sigma_N(\vec{r}, E') \delta(E' - E) \delta(\vec{\omega}' - \vec{\omega}) N(\vec{r}, \vec{\omega}', E') dE' d\vec{\omega}' . \quad (46)$$

The solution to (1) is unchanged if we add (46) to (1). Now, if this is done, for whatever g might be chosen σ_N can be adjusted so that σ^* is everywhere positive.² In fact, σ_N can be so chosen that at each energy σ^* is a constant in $\vec{r}$.^{15,16} This device avoids negative cross sections and reduces the problem of tracking to the next collision point to the simple problem it assumes in a one-medium system. The reworking of cross sections involved might be very awkward in an existing Monte Carlo program, but should not cause excessive work in a program newly constructed. The cross section, σ_N , causes an increase of collision points in low-cross-section media. Whether or not this increase is beneficial depends upon the problem at hand.

Leimdorfer has dealt with the negative cross section in two distinct ways. His first suggestion⁵ is that wherever $\sigma^* < 0$ redefine

$$\sigma^* = \sigma' > 0 , \quad (47)$$

with σ' chosen arbitrarily by the user. [In fact, Leimdorfer suggests a definite form for σ' . It is, however, not an ideal choice, and the essence of this palliative is the arbitrary replacement, not a particular form.] The more elegant suggestion of Leimdorfer and the one that he has used in most of his work is that when $\sigma^* \leq 0$ one simply accepts it and selects from the appropriate distribution (A-3).⁶ This is always possible if $y(s)$ [see (A-2)] is finite for every s allowed by (A-3a). In practical terms this almost always means the system must have an external boundary finitely placed in the direction of tracking.

One can, in substantial approximation, use this method even if $y(s)$ can become infinite. We simply allow σ^* to be negative for $0 \leq s \leq T_1$ and set

$$\sigma^* = \sigma' > 0 \quad (48)$$

for $s > T_1$ (with $\sigma' = \sigma$ a logical choice). The dynamics of (48) are dealt

with in Appendix D [see (D-9) through (D-17)]. T_1 can be specified directly as a length or indirectly as the distance to some specified location.

The notions implicit in (48) can be applied with profit often when $y(s)$ remains finite. Ordinarily in such cases evaluation of $y(T)$ and normalizing (A-3) implies tracking to an external boundary and not permitting escapes. If there should be a particularly important region close to an external boundary and far from the source it would very likely be profitable to go through the mechanics of non-escape for tracks from collision points in such a region. For other tracks the benefits are not apt to justify the labor. For such other tracks we can again define a T_1 (which may be region-dependent); use $\sigma^* < 0$ out to T_1 and (48) beyond.

Let us note that both (47) and (48) represent departures from the exponential transform. That is, when we modify our sampling rule as required by (47) or (48), the rule is no longer that dictated by the exponential transform. We shall consider these and other departures from the strict formalism of the transform.

It has been observed that rather large variances are associated with particles scattered in the "wrong" direction for certain values of g . With form (23), for example, in a system of reasonably large diameter (say three mean free paths or more), the values $\nu \approx 1$, $\vec{\omega} \cdot \nabla f \approx -1$ lead to this condition. Or, more generally, for **any** form

$$\sigma^* \gtrsim 2 \sigma \quad (49)$$

generates a group of large variance particles. This is readily seen as follows. The unmodified distribution function is

$$p_1(x) = \sigma e^{-\sigma x} \quad (50)$$

The modified function is

$$p_2(x) = \beta \sigma e^{-\beta \sigma x} \quad (51)$$

with associated weight correction factor

$$w_2(x) = \frac{1}{\beta} e^{-\sigma x(1 - \beta)} \quad . \quad (52)$$

The variance of $w_2(x)$ in the distribution $p_2(x)$ is

$$\sigma_{w_2}^2 = \int_0^D (w_2(x) - \bar{w}_2)^2 p_2(x) dx = \frac{\sigma}{\beta} \int_0^D e^{-\sigma x(2 - \beta)} dx - 1 \quad , \quad (53)$$

which becomes very large for $\beta \geq 2$ and D large, D being the maximum track length in the direction chosen. A large value of σ_w^2 is indicative of a large statistical error. A small value has no significance whatever; for note that $w_1(x)$ [associated with $p_1(x)$] is one.

If a value of $\beta \approx 2$ is applied to a generally unimportant class of histories--say those which backscatter in a deep penetration problem--it may have little effect on the whole answer. We can accept a large percentage error in a small part of the answer. If, however, one attempts to confine important particles to an important region by applying to them a scheme equivalent to $\beta \geq 2$, then trouble will almost surely follow.

We have noticed that the tracking procedure involves solving (41) for s in terms of y --often using (43). There is little difficulty in this when the spatial part of $g(\vec{r}, \vec{u}, E)$ is given by form (31). More elaborate forms, including the very useful form (32), present difficulties in the solution of (41) explicitly for s . Such difficulties can be faced by providing correct values for the integrals in question by brute force methods. Much simpler approximate approaches appear to offer more promise, and, on account of their nature, we shall discuss them in the following section.

(Note Added in Proof)

A particularly artful approach to the negative effective total-cross-section problem involves the definition of an artificial "fission" process and appears to be attributable to Goertzel and Kalos.² Let us write the total cross section as the sum of an absorption and a scattering part $\sigma_s + \sigma_a$ and let us add to the effective total cross section, σ^* , the quantity $(v + 1)\gamma$ such that the adjusted effective total cross section σ_1^* is

$$\sigma_1^* = \sigma_s + \sigma_a - \vec{\omega} \cdot \nabla g + (v + 1)\gamma$$

$$v > 0$$

$$\gamma > 0$$

$$\sigma_a - |\nabla g| + \gamma > 0 \quad .$$

We now require that the scattering processes governed by σ_s occur with probability

$$\frac{\sigma_s}{\sigma_s + \sigma_a - \vec{\omega} \cdot \nabla g + (v + 1)\gamma} \quad ,$$

the absorption process with probability

$$\frac{\sigma_a - \vec{\omega} \cdot \nabla g + \gamma}{\sigma_s + \sigma_a - \vec{\omega} \cdot \nabla g + (v + 1)\gamma} \quad ,$$

and the newly defined fission process with probability

$$\frac{v\gamma}{\sigma_s + \sigma_a - \vec{\omega} \cdot \nabla g + (v + 1)\gamma} \quad ,$$

and further that the number of particles emerging from fission be $\frac{v+1}{v}$ (or that the weight emerging from fission be $\frac{v+1}{v}$ times the weight inducing fission), with the emergent particles having the same phase space coordinates as the particle inducing fission.

VII. VARIATIONS FROM THE EXPONENTIAL TRANSFORM

A number of references^{3,13,14} quoted in Section II do not deal with the exponential transform as such but rather apply to track length distribution importance-sampling schemes which may closely resemble those suggested by the transform. In the preceding section we have noted that two of the remedies offered for the negative cross section problem take one, at certain stages of the calculation, outside the formalism of the transform. We shall not hesitate to consider other procedures which may be suggested by the form of the transform or as simplifications of some of its more complex forms, even though those procedures may not fit within the narrow confines of the transform. This does not mean that we are introducing error or approximation into the calculation of the flux density N . As long as condition (22) is met, N retains the same expectation value. We simply are using a different variance reduction system when we choose from an altered distribution.

We have noted in Section III that we essentially satisfy the requirements of the exponential transform if we use (19) and (20) as the altered source and track length distributions, respectively. Choice of an altered source distribution other than (19) may sometimes be indicated for reasons of convenience or more effective variance reduction. The fact that such a distribution is not consistent with (3) should not prevent one from employing it.

We have seen in Sections V and VI that tracking problems can arise from two sources: discontinuities in g and difficulties in evaluating s as an explicit function of y . If we approximate y by

$$y = \int_0^s \sigma^*(\vec{r}_0 + u\vec{w}) du \approx \sigma^*(\vec{r}_0) s, \quad (54)$$

leading to

$$L'(y) = \frac{| \sigma^* | e^{-\sigma^* s}}{\int_0^T | \sigma^* | e^{-\sigma^* s} ds} , \quad (55)$$

both these problems are overcome.

The form of the transform is violated at boundary crossings and in regions where ∇g departs significantly from a parallel field over the length of the track. If we are using forms (23) and (32), the latter deviation will occur significantly in the neighborhood of (x_D, y_D, z_D) . But giving ν a small magnitude in a reasonable neighborhood of (x_D, y_D, z_D) will avoid this problem in large measure and is the appropriate thing to do if $\lambda = -1$. (If $\lambda = +1$ the problem should be studied in detail to determine what, if anything, ought to be done.)

The variations on the transform so far suggested are all in the nature of simplifications or time savers. They are themselves, therefore, variance reducers (if properly chosen) on account of the equivalence of time and variance.

We now consider a strikingly different variation. It has already been noted that the altered distributions (19) and (20) can be dealt with independently of each other. Distribution (20) does not contain g explicitly--only ∇g . The track length distribution scheme (20) has a more substantial effect on a deep penetration calculation than does the source distribution scheme (19). Form (23), along with (31) or (32), brings out rather forcefully the significance of

$$\nabla g = \nu \sigma \nabla f . \quad (56)$$

ν measures the magnitude of the alteration introduced by the transform. Ordinarily in form (23) it is restricted to $0 \leq \nu \leq 1$; however, with Leimdorfer's scheme of permitting negative and zero values of σ^* (i.e., selecting from a positive exponential distribution for negative values and from a uniform distribution for zero) the upper limit is removed.

∇f is a field of unit vectors which indicates at each point the a priori estimate of the direction of transport which offers the greatest probability of reaching some selected point or region (e.g., a detector). Note that in an inhomogeneous medium this direction will not in general indicate a straight-line trajectory to the desired point.

In using Monte Carlo one never knows the ideal scheme of importance sampling, for then the problem could be directly solved by a quadrature; however, schemes which only roughly approximate the ideal can lead to very substantial savings of effort.² Given a complex geometric configuration in which to compute radiation transmission, an experienced person should be able to estimate roughly the preferred direction (for reaching the detector) from each section of the system. (Generally, it will be the line of fewest mean free paths except possibly when close to the detector.) In this manner a crude ∇f of simple form can be sketched. Experience, trial and error, or examination of results like those presented in the next section should provide enough help to choose ν -- possibly different ν 's in different regions. One can therefore construct the functions required by form (23) in quite complex situations and without very great difficulty.

Now if it should happen that the function $(\sigma \nu \nabla f)$ so constructed is the gradient of a scalar field--possibly having discontinuities at certain surfaces--then the function satisfies the requirements of the exponential transform, and the scalar field is the function g . But even if the function is not derivable from a scalar field so that there exists no function g which when cast into (2) generates this function as ∇g in form (24), it is still just as useful a function and serves the same purposes when cast in the role indicated by form (25). That is to say, even though we do not use the exponential transform formally in this case we use a form suggested by it and expect to derive similar benefits.

VIII. SOME RESULTS

We are fortunate in having available results reported by a number of users of the exponential transform. The estimated relative error reported in each case appears to be defined as the deviation of the mean divided by the estimate of the mean.

The first case we deal with⁵ is gamma-ray penetrations of a concrete slab (thickness d , density 2.35 gm/cm^3 ordinary concrete). Sources are monoenergetic (7 MeV), isotropic, and spatially distributed according to the rule e^{-kz} . The transform used was given by

$$g = \vec{\omega} \cdot \vec{\epsilon} C, \quad (57)$$

in which $\vec{\epsilon}$ = unit vector normal to slab face.

Relative errors are given in Fig. 1 for various cases plotted as a function of c/Σ . (Σ = total cross section at the source energy.) Negative σ^* are dealt with by use of Leimdorfer's first scheme--the replacement of the negative σ^* by an arbitrarily chosen positive σ .

Figure 2 deals with the same problem as Fig. 1 but with $d = 100 \text{ cm}$.⁶ The notation is the same as in Fig. 1. Curves marked I are replicas of some of those appearing in Fig. 1. For the curves marked II negative cross sections were dealt with by use of Leimdorfer's second scheme--selection from a positive truncated exponential.

The important conclusions to draw from Figs. 1 and 2 are that C/Σ (which is essentially equivalent to the quantity we call ν) when taken equal to ~ 1 in the first scheme leads to a poor statistical error, but in the second scheme gives very nearly the best results. That is, C/Σ (or ν) ≈ 1 seems to be a good choice for these problems, but only with the truncated positive exponential scheme can one properly exploit that choice.

Figure 3⁸ relates to neutron calculations. Again the medium is concrete. The source is a point isotropic fission spectrum at the face of the slab. The same form of g is used as in the preceding cases. Two concrete thicknesses are considered, 30 cm and 210 cm. The errors quoted for the latter slab are rather high, but it is notable that each point on both curves was computed with the same time investment. The 30-cm slab shows best statistics for $C/\Sigma_{tr} \approx 1.1$. (Note that the transport cross section has replaced the total cross section in this unit. It is no longer equivalent to ν .)

Figure 4 deals with neutrons also.⁷ More important, it is a problem in reflections. The medium is hydrogen and the number current albedo for

Fig. 1. Relative Error for Various Choices of Biasing Parameter (Gamma-Ray Penetration of Concrete, Form (4) of Effective Cross Sections, Leimdorfer's First Scheme for Negative Cross Sections, Source Given by e^{-kz}).

Fig. 2. Relative Error for Various Choices of Biasing Parameter (Gamma-Ray Penetration of Concrete, Form (4) of Effective Cross Sections, Source Given by e^{-kz} . I. Leimdorfer's First Scheme for Negative Cross Sections. II. Leimdorfer's Second Scheme for Negative Cross Sections).

Fig. 3. Relative Error for Various Choices of Biasing Parameter (Neutron Penetration of Concrete Point Fission Spectrum, Form (4) of Effective Cross Sections, Leimdorfer's Second Method for Negative Cross Sections).

Fig. 4. Relative Error for Various Choices of Biasing Parameter
(Neutron Reflection from Hydrogen, Form (4) of Effective Cross Sections).

3-MeV neutrons is evaluated. The form of g is again as above, but negative C values are chosen. The curve of relative error is quite flat and approximately minimal in the range

$$-1 > C > -3 ,$$

roughly equivalent to

$$-\frac{\sigma_{tr}}{\sigma_{tr}} > C > -\frac{\sigma_{total}}{\sigma_{tr}}$$

There is a significant improvement over the untransformed ($C = 0$) case. The transform, contrary to what might have been expected, has useful application in this, a reflection problem.

Figure 5 deals with calculations of the gamma-ray energy buildup factor in lead for normally incident 6-MeV gamma rays.¹² The form of g is still (4) with

$$C = F \sigma_{min} ;$$

σ_{min} is the minimum value assumed by the total cross section and the buildup factors are plotted versus F . In each case the value plotted for $F = .65$ comes closest to agreement with the value interpolated from moments methods calculations.¹⁶ This value of F corresponds to C values for the preceding cases considered much less than 1. There is at least the possibility of disagreement, but it must be borne in mind that the cases are not strictly comparable.

Figure 6 deals with penetration of an aluminum slab by ^{60}Co gamma rays (1.25 MeV) normally incident.⁴ The form of g is again (4). Results are plotted against C as an independent variable. μ_0 , the total cross section at the source energy, is shown on the chart, and it is clear that best results are obtained for C close to μ_0 . Figures 7 and 8 in the same general problem relate to 2-MeV gamma rays incident upon iron and 6-MeV

ORNL-DWG 65-9510

Fig. 5. Relative Error for Various Choices of Biasing Parameter (Buildup Factor for 6-MeV Gamma Rays Normally Incident on Various Thicknesses of Lead, Form (4) of Effective Cross Sections).

ORNL-DWG 65-9511

Fig. 6. Relative Error for Various Choices of Biasing Parameter (Gamma-Ray Penetration of Aluminum, 1.25-MeV Source Energy, Form (58) of Effective Cross Sections).

Fig. 7. Relative Error for Various Choices of Biasing Parameters (Gamma-Ray Penetration of Iron, 2-MeV Source Energy, ○ - Form (59) of Effective Cross Sections, — - Form (58) of Effective Cross Sections).

ORNL-DWG 65-9513

Fig. 8. Relative Error for Various Choices of Biasing Parameters (Gamma-Ray Penetration of Lead, 6-MeV Source Energy, O - Form (59) of Effective Cross Sections, — - Form (58) of Effective Cross Sections).

gamma rays incident upon lead, respectively.⁴ Two different forms of the track length distribution were run comparatively in each of these cases, specialized versions of forms (4) and (23):

$$\sigma^* = \sigma - C \vec{\omega} \cdot \vec{\epsilon} \quad (58)$$

$$\sigma^* = \sigma(1 - M \vec{\omega} \cdot \vec{\epsilon}) \quad (59)$$

Results obtained with (59) are plotted as circles and results obtained with (58) are indicated by the curves in Figs. 7 and 8. No significant difference is disclosed in the comparison. According to the users,⁴ however, form (59) was much more convenient on account of the ease of cross-section preparation (since $M = C/\mu_0$).

Another comparison was run in a lithium hydride system approximately two feet thick with a plane disk isotropic distribution of fission-like neutrons at one surface.¹⁸ Three cases were run:

1. Scheme 1 - no track length biasing
2. Scheme 2 - form (25) with ν region-dependent ranging from $2/3$ far from the detector to zero close to the detector,
3. Scheme 3 - track length distribution

$$1/k e^{-x/k}$$

with

$$k = \frac{1}{1 - \nu} \quad .$$

The same source selection scheme was used in all three cases, and it was devised independently of the form of any of the track length distributions. Results obtained were as follows:

Scheme No.	Relative Error
1	.32
2	.109
3	.109

The relative error of scheme 1 is so great that little meaning can be ascribed to the results. Schemes 2 and 3, it will be noted, give essentially identical relative errors. This is probably because the biasing schemes are very much alike for forward scattering, and particles which have back-scattered are not apt to be important in the result.

The following problem was devised to provide a test for some of the concepts introduced in this report.¹⁹ There were two media: $Z > 0$, air; $Z \leq 0$, water. A point, monoenergetic (6 MeV) source was located at $(0,0,-10)$. The detector was a plane surface given by $x^2 + y^2 > 25000^2$, $Z = -10$. One case was run without track length selection biasing. The other case was run using scheme (25) with the following choices:

Region I $Z > 0$

f is form (32) $\nu = .3$

$$x_D = \frac{D}{q} x, \quad y_D = \frac{D}{q} y, \quad z_D = 0,$$

$$\lambda = -1, \quad q^2 = x^2 + y^2, \quad D = 25020$$

Region II $Z \leq 0 \quad x^2 + y^2 < 6.2001 \times 10^8$

f is form (31) $\nu = .42$

$$\vec{\epsilon} = (0, 0, 1)$$

Region III $Z \leq 0 \quad 6.2001 \times 10^8 \leq (x^2 + y^2) < 25000^2$

f is form (32) $\nu = .42$

$$x_D = \frac{D}{q} x, \quad y_D = \frac{D}{q} y, \quad z_D = -15,$$

$$\lambda = -1, \quad q^2 = x^2 + y^2, \quad D = 25000$$

Region IV $-15 < Z \leq 0 \quad x^2 + y^2 \geq 25000^2$

f is form (31) $\nu = 0$

$$\vec{\epsilon} = (0, 0, -1)$$

Region V $Z \leq -15 \quad x^2 + y^2 \geq 25000^2$

f is form (31) $\nu = 0$

$$\vec{\epsilon} = (0, 0, +1)$$

The relative error in the unbiased case was .275, and in the biased case .195. Machine running times were substantially the same (within 3%). This comparison was chosen to illustrate the simple construction of the vector field ∇f for a fairly complicated geometry. No attempt was made to optimize or improve the choices of the parameters used in biasing.

ACKNOWLEDGEMENT

We are indebted to many people for their help in preparing this report. Keith Penny and D. C. Irving made thorough reviews with suggestions and corrections. The staff of RSIC sought out many of the references used. Martin Leimdorfer gave us the benefit of numerous instructive conversations. Nancy Betz pulled us back from the brink of error and provided helpful suggestions besides. Francis Kam provided us with some special comparison runs. R. R. Coveyou shed needed light on boundary surfaces.

REFERENCES

1. H. Kahn, Nucleonics, 6(5), 27-37 and 6(6), 60-65 (1950).
2. G. Goertzel and M. Kalos, Progress in Nuclear Energy, Series I, Vol. II (1958).
3. R. A. Liedtke and H. A. Steinberg, A Monte Carlo Code for Gamma-Ray Transmission through Laminated Slab Shields, WADC-TR 58-80 (AD 203641) (1958).
4. J. F. Perkins and M. O. Burrell, Efficiency of Some Biased Sampling Schemes in Monte Carlo Calculations, Nuclear Engineering and Science Conference April 6-9, 1959, published by Engineers Joint Council.
5. M. Leimdorfer, Nukleonik, 6(2), 58-65 (March 1964).
6. M. Leimdorfer, "On the Transformation of the Transport Equation for Solving Deep Penetrating Problems by the Monte Carlo Method," Trans. Chalmers University of Tech., Gothenberg, Sweden, No. 286 (1964).
7. M. Leimdorfer, "The Backscattering of Fast Neutrons from Plane and Spherical Reflectors," Trans. Chalmers University of Tech., Gothenberg, Sweden, No. 288 (1964).
8. M. Leimdorfer, "On the Use of Monte Carlo Methods for Calculating the Deep Penetration of Neutrons in Shields," Trans. Chalmers University of Tech., Gothenberg, Sweden, No. 287 (1964).
9. M. Kalos, Methods in Monte Carlo Solution of the Radiation Transport Equation, NDL-TR-36 (May 1962).
10. B. Eriksson, On the Use of Importance Sampling in Particle Transport Problems, AE-190 (June 1965).
11. L. A. Beach, R. B. Theus, J. D. Plawchan, and W. R. Faust, Stochastic Estimates of γ -Ray Spectral Intensities for Shallow and Deep Penetration, NRL-4412 (1954).
12. A. G. Pieper and L. A. Beach, Stochastic Estimates of the Penetration of Gamma Rays Through Slab Shields, NRL-5981 (1963).
13. J. J. Loechler and J. E. MacDonald, Flexible Monte Carlo Programs, FMC-N and FMC-G, TID-4500 (1961).
14. M. B. Wells, Monte Carlo Multi-Layer Slab Geometry Shielding, Code C-18, FZK-134-3 (1961).
15. J. B. Parker and E. R. Woodcock, Monte Carlo Criticality Calculations, Progress in Nuclear Energy, Series IV, Vol. 4, Pergamon Press, Oxford (1961).

16. E. R. Woodcock, T. Murphy, P. J. Hemmings, and T. C. Longworth, "Techniques Used in the GEM Code for Monte Carlo Neutronics Calculations in Reactors and Other Systems of Complex Geometries," Conference on Applications of Computing Methods to Reactor Problems, Argonne National Laboratory, May 1965, ANL-7050
17. H. Goldstein and J. E. Wilkins, Jr., Calculations of the Penetration of Gamma Rays, NYO-3075 (1954).
18. Private communication, F. B. K. Kam.
19. Private communication, N. A. Betz.

APPENDIX A. SPACE INTEGRAL OF THE BOLTZMANN EQUATION

Let us integrate an equation having the form of (1) or (3) from $\vec{r}_0$ to $\vec{r}_0 + \vec{w}s$, with $\vec{w}$ fixed and without reference to energy dependence. This leads readily to

$$\begin{aligned}
 F(\vec{r}_0 + \vec{w}s, \vec{w}) &= M_1 F(\vec{r}_0, \vec{w}) \frac{e^{-\int_0^s \sigma^*(u) du}}{M_1} \\
 &+ \int_0^s dy \frac{e^{-\int_y^s \sigma^*(u) du}}{M_2} \sum_{i=1}^I \int \int \alpha_i \sigma_{si}(\vec{r}_0 + \vec{w}y) K_i(\vec{w}' - \vec{w}) \\
 &M_2 F(\vec{r}_0 + \vec{w}y, \vec{w}') d\vec{w}' + \int_0^s dy \frac{e^{-\int_y^s \sigma^*(u) du}}{M_3} M_3 S^*(\vec{r}_0 + \vec{w}y, \vec{w}) . (A-1)
 \end{aligned}$$

The first term on the right represents the portion of the particle density directed toward $\vec{w}$ at $\vec{r}_0$ which reaches $\vec{r}_0 + \vec{w}s$; the second term represents that portion scattered into the $\vec{w}$ beam along the track from $\vec{r}_0$ to $\vec{r}_0 + \vec{w}s$ and surviving to reach $\vec{r}_0 + \vec{w}s$; the third term represents source neutrons born on the trajectory, directed toward and surviving to $\vec{r}_0 + \vec{w}s$. The factors M_1 , M_2 and M_3 are normalizing factors on the track length distributions when they appear in denominators, and they are weight compensation factors when they appear in numerators.

It is apparent for every one of these classes that, when $\sigma^* \neq 0$, the distribution along the track length is a negative exponential in the quantity

$$y(s) = \int_0^s \sigma^*(u) du \quad . \quad (A-2)$$

This statement holds independently of the form or the sign of σ . The distribution can be normalized only if $\exp [-y(s)]$ is bounded for all values of s . Hence, negative values of σ^* (corresponding to net production over absorption) require, in general, bounded s . Normalizability is, of course, essential to transforming the distribution into a probability distribution. The probability distributions implied by (A-1) are of the form

$$L(s) = \frac{|\sigma^*(s)| e^{-\int_0^s \sigma^*(u) du}}{\int_0^T |\sigma^*(s)| e^{-\int_0^s \sigma^*(u) du} ds} \quad (A-3)$$

$$0 \leq s \leq T \quad . \quad (A-3a)$$

It is convenient, though not essential, to incorporate the absolute values symbol in (A-3).

The special case $\sigma^* = 0$ implies that

$$L(s) = \text{constant} = \frac{1}{T} \quad (A-4)$$

$$0 \leq s \leq T \quad . \quad (A-4a)$$

APPENDIX B. SPATIAL DISCONTINUITIES IN g

The meaning of the term $\vec{w} \cdot \nabla g$ is not defined at a surface across which g has a finite discontinuity. [This is the only kind of nondifferentiable behavior we shall allow in the spatial dependence of g .] The particle flux density, N , in (1) is everywhere continuous. It therefore follows from (2) that φ must be discontinuous at the boundaries at discontinuity in g . Let us use superscripts (+) and (-) to designate the two values of each of the discontinuous quantities at the boundary. Then

$$\varphi^{(-)} e^{g^{(-)}} = \varphi^{(+)} e^{g^{(+)}} \quad . \quad (B-1)$$

Hence, in Monte Carlo solutions of (3), the φ -particle weight is multiplied by $e^{g^{(-)} - g^{(+)}}$ when the particle crosses the boundary in the direction $(-) \rightarrow (+)$, and by the factor $e^{g^{(+)} - g^{(-)}}$ when the particle crosses the boundary in the direction $(+) \rightarrow (-)$.

In the solution of Eq. (1) by the importance-sampling techniques described in Section III, the same factors, $(\pm)(e^{g^{(+)} - g^{(-)}})$, occur explicitly in the weight correction ratio of (14) to (20) when the line integral in these expressions crosses a boundary of discontinuity.

APPENDIX C. EQUIVALENCE OF THE IMPORTANCE SAMPLED BOLTZMANN SOLUTION AND THE TRANSFORMED SOLUTION

Suppose Eqs. (1) and (3) are solved after the manner described in Section III with the same set of random numbers in both cases. [The importance-sampling technique described in III is used in the solution of (1).] Then, since in both cases the same source, track length, and angle-energy distributions are sampled, it follows that identical sets of phase coordinates are generated at every collision in every history. It remains only to investigate weights.

First, we shall deal with the case of continuous g and the effective cross section everywhere positive.

Let us use the subscript notation b,j and a,j to denote respectively the value of a quantity immediately before and immediately after a collision at the point $\vec{r}_j$. Let us also note that equivalent solutions of (1) and (3) must have relative source normalizations in the ratio

$$\frac{\int S(\vec{r},E) d\vec{r} dE d\vec{\omega}}{\int S(\vec{r},E) e^{g(\vec{r},\vec{\omega},E)} d\vec{r} d\vec{\omega} dE} \quad (C-1)$$

[see (13) and (19)].

The weight adjustment applied at each collision point in the solution of (3) is given by (21) so that the weight just before the i^{th} collision in a history is

$$\prod_{j=1}^{i-1} \frac{\sigma_s(\vec{r}_j, E'_j) e^{g(\vec{r}_j, \vec{\omega}_j, E_j) - g(\vec{r}_j, \vec{\omega}'_j, E'_j)}}{\sigma(\vec{r}_j, E'_j) - \omega_j \cdot \nabla g(\vec{r}_j, \vec{\omega}'_j, E'_j)} \alpha^j, \quad (C-2)$$

where α^j here means the multiplicity factor at the j^{th} collision [see (18)].

In the solution of (1) the initial source weight is adjusted by the ratio of (13) to (19) and at each collision point the weight factor, the

ratio of (14) to (20), is applied. This leads to a weight at the i^{th} collision point of

$$\frac{\int S(\vec{r}', E') e^{g(\vec{r}', \vec{\omega}', E')} d\vec{r}' d\omega' dE'}{e^{g(\vec{r}_0, \vec{\omega}_0, E_0)} \int S(\vec{r}', E') d\vec{r}' dE' d\vec{\omega}}$$

$$\prod_{j=2}^i \left(\frac{\sigma}{\sigma - \frac{dg}{ds}} \right)_{a,j-1} e^{-g_b(\vec{r}_j) + g_a(\vec{r}_{j-1})} \alpha^{j-1}$$

$$= \frac{\int S(\vec{r}', E') e^{g(\vec{r}', \vec{\omega}', E')} d\vec{r}' d\vec{\omega}' dE'}{\int S(\vec{r}', E') d\vec{r}' dE'} e^{-g_b(\vec{r}_i)}$$

$$\prod_{j=1}^{i-1} \left(\frac{\sigma}{\sigma - \frac{dg}{ds}} \right)_{a,j} e^{-g_b(\vec{r}_j) + g_a(\vec{r}_j)} \alpha^j \quad . \quad (C-3)$$

But since

$$e^{-g_b(\vec{r}_j) + g_a(\vec{r}_j)} = e^{g(\vec{r}_j, \vec{\omega}, E) - g(\vec{r}_j, \vec{\omega}', E')} \quad , \quad (C-4)$$

expression (C-3) is equal to (C-2) multiplied by the equivalent source renormalization [see (C-1) and the conversion factor given by (2)]. The solutions are therefore equivalent for this case.

When there are boundaries present, at each boundary crossing in the solution of (3) for ϕ the weight is corrected by the factor $\pm(e^{g(-)} - e^{g(+)})$. In the corresponding history in the importance-sampled solution of (1), the weight adjustment factor [ratio of (14) to (20)] contains the same correction factor (see Appendix B). The equivalence proof therefore holds for discontinuous g .

When an effective cross section is somewhere not positive, or for some other reason, like non-escape truncation, an explicit normalization number ($\neq 1$) must be carried in the track length distribution [see (A-1)], that number occurs as an explicit factor in the ratio of (14) to (20), the biased track length weight correction factor for the importance-sampled solution to (1). We see in Appendix A (A-1), however, that the normalization number also becomes a factor in the scattering kernel weight function (or on the source) in the solution of (3) for φ .

The solutions are therefore equivalent for all cases.

APPENDIX D. CHOOSING FROM THE TRACK LENGTH DISTRIBUTIONS

We have already encountered the general expressions for track length distributions

$$L(s) = \frac{|\sigma^*(s)| e^{-\int_0^s \sigma^*(u) du}}{\int_0^T |\sigma^*(s)| e^{-\int_0^s \sigma^*(u) du} ds} \quad (D-1)$$

and

$$0 \leq s \leq T \quad . \quad (D-2)$$

We have also observed that if σ^* does not change sign it is convenient to set

$$y(s) = \int_0^s \sigma^*(u) du \quad (D-3)$$

and

$$L^*(y) = \frac{e^{-y}}{\int_0^T e^{-y(s)} ds} \quad (D-4)$$

For $\sigma^* > 0$ and $T = \infty$, (D-4) is the familiar exponential distribution for which there are a number of standard selection techniques. For $\sigma^* > 0$ and T finite, (D-4) becomes the truncated exponential distribution. To select from it we need only select from the exponential distribution, divide by $y(T)$, and retain the remainder.

The denominator of (D-4) must, of course, be finite. Hence, for $\sigma^* < 0$, either T must be finite or $\sigma^*(u)$ must tend to zero fast enough so that

$$y(T) = \int_0^{\infty} \sigma^*(u) du \quad (D-5)$$

$y(T)$ = some finite (negative) member.

In either case we may set

$$z(s) = y(T) - y(s) \quad (D-6)$$

and choose z from

$$L'(z) = \frac{e^{z(s)}}{\int_0^T e^{z(s)} ds} \quad , \quad (D-7)$$

which is again the truncated negative exponential since z is negative. Equation (D-6) implies that

$$z(s) = - \int_T^s \sigma^*(u) du \quad . \quad (D-8)$$

That is, to find the point s we lay off z not from the collision point outward along the direction $\vec{w}$ but from the extremity in that direction ($s = T$) inward toward the collision point.

Note the equivalence

$$\frac{\int_0^{T_m} |\sigma^*(s)| e^{-\int_0^s \sigma^*(u) du} ds}{\int_0^{T_m} |\sigma^*(s)| e^{-\int_0^s \sigma^*(u) du}} = \sum_{i=0}^{M-1} \lambda_i W_i(s), \quad (D-9)$$

$$\lambda_i = \frac{\int_{T_i}^{T_{i+1}} |\sigma^*(s)| e^{-\int_0^s \sigma^*(u) du} ds}{\int_0^{T_m} |\sigma^*(s)| e^{-\int_0^s \sigma^*(u) du} ds}, \quad (D-10)$$

$$W_i(s) = \frac{H_i(s) e^{-\int_0^s \sigma^*(u) du} |\sigma^*(s)|}{\int_{T_i}^{T_{i+1}} |\sigma^*(s)| e^{-\int_0^s \sigma^*(u) du} ds}, \quad (D-11)$$

$$\begin{aligned}
 H_i(s) &= 0 & s < T_i, \geq T_{i+1} \\
 &= 1 & T_i \leq s < T_{i+1} \quad ,
 \end{aligned}
 \tag{D-12}$$

$$\sum_{i=0}^{M=1} \lambda_i = 1 \quad ,
 \tag{D-13}$$

$$\int_{T_i}^{T_{i+1}} W_i(s) ds = 1 \quad .
 \tag{D-14}$$

Note further that selection of s from $W_i(s)$ is equivalent to selection of s' from

$$W'(s') = \frac{
 \begin{aligned}
 & - \int_0^{s'} \sigma^*(v + T_i) dv \\
 & |\sigma^*(s' + T_i)| e^0
 \end{aligned}
 }{
 \begin{aligned}
 & \int_0^{T_{i+1}-T_i} |\sigma^*(s' + T_i)| e^0 \\
 & - \int_0^{s'} \sigma^*(v + T_i) dv
 \end{aligned}
 } ds'
 \tag{D-15}$$

and setting

$$s = s' + T_i \quad .$$

In the above we have

$$T_0 = 0$$

(D-17)

$$T_0 < T_1 < \dots < T_M \quad .$$

The T_i 's correspond to boundaries of some sort at which the sign or form of $\sigma^* [= \sigma(u)]$ may change. Hence, this rather complicated case is reducible to a combination of (41) and (D-7).

Distribution

1. E. P. Blizzard
2. F. C. Maienschein
3. W. H. Jordan
4. Arthur Reetz, NASA Headquarters
5. Major William Alfonte, DASA Headquarters
- 6-15. Carmen Cialella, Ballistic Research Laboratories
16. P. B. Hemmig, Div. of Reactor Development and Technology, AEC
- 17-18. R. E. Bowman, DTI
- 19-28. F. H. Clark
29. G. Dessauer (consultant)
30. M. L. Goldberger (consultant)
31. M. H. Kalos (consultant)
32. L. V. Spencer (consultant)
- 33-35. Laboratory Records Department
36. Laboratory Records, RC
37. ORNL Patent Office
- 38-39. Central Research Library
40. Document Reference Section
- 41-190. RSIC Distribution
- 191-640. Division of Technical Information Extension
641. Research and Development Division (ORO)