KERNFORSCHUNGSZENTRUM

KARLSRUHE

NOVEMBER/DEZEMBER 1961 KFK 91

INSTITUT FÜR NEUTRONENPHYSIK UND REAKTORTECHNIK

ISOPERIMETRIC AND OTHER INEQUALITIES IN THE THEORY OF NEUTRON TRANSPORT

LAWRENCE DRESNER

KERNREAKTOR
BAU- UND BETRIEBS-GESELLSCHAFT M.B.H.
KARLSRUHE

Reprinted from the JOURNAL OF MATHEMATICAL PHYSICS, Vol. 2, No. 6, 829–847, November–December, 1961
Printed in U. S. A.

Isoperimetric and Other Inequalities in the Theory of Neutron Transport

LAWRENCE DRESNER*
Kernforschungszentrum Karlsruhe, Karlsruhe, Germany
(Received April 28, 1961)

Some isoperimetric and other inequalities occurring in the one-velocity theory of neutron transport are derived. The quantities involved in these inequalities all refer to bare solids with isotropic scattering and are: the critical multiplication, the first-collision probability, the non-escape probability, and the buckling. The inequalities proved provide upper and lower bounds for the quantities considered, and numerous examples of the estimation of these quantities in cases not readily amenable to direct calculation are given.

1. INTRODUCTION

1.1. In a mathematically complicated subject like the theory of neutron transport, simple, exact, and general formulas are usually not obtainable. In order to calculate quantities of interest, recourse must generally be had either to numerical calculation or to the introduction of simplifying but untrue assumptions. The introduction of such assumptions usually provides explicit and easily evaluated formulas but also usually results in errors of indeterminate sign and magnitude. It must be fairly said that situations in which these errors are small can generally be recognized when one understands the "physical" content of the theory, but the intuitive nature of this approach nevertheless involves an unavoidable, and furthermore itself uncertain, extent of error.

An elegant and also useful way out of this dilemma consists of enlarging the class of acceptable results to include inequalities. Thereby is one often provided with relationships involving the quantities of interest which are again simple and general, and which are furthermore exact at least in the sense of involving no mutilation of the theory. Two such inequalities providing, respectively, an upper and lower bound will furthermore yield estimates *whose maximum possible error is known*. If these bounds are close ones, as often happens, the numerical accuracy of the estimates may suffice for practical purposes.

1.2. The quantities of interest which we shall consider in this paper are all set functions which arise in the theory of neutron transport and which refer to bare, homogeneous, convex solids with isotropic scattering.

They are: the critical multiplication, the non-escape (absorption) probability of neutrons from a uniform isotropic source inside the solid, the corresponding first-collision probability, the buckling, and the diffusion-theoretic non-escape probability. The first three of these set-functions belong to strict one-velocity transport theory; the fourth and fifth to the simpler diffusion theory.

The transport-theoretic quantities have been calculated accurately in terms of simple formulas or as the result of not prohibitively great numerical labor only for the simplest geometrical shapes. The critical multiplication[1] has been calculated only for slabs and spheres, the non-escape probability[2] only for slabs, while the first-collision probability[3] has been calculated only for slabs, spheres, infinite right circular cylinders, hemispheres, and some oblate spheroids. For such a simple solid as a cube, however, no exact values for any of these quantities are available. Even the diffusion-theoretic quantities can only be calculated easily for spheres, rectangular parallelepipeds, and finite and infinite right circular cylinders. For more exotic shapes than those just mentioned, straightforward calculation can be very tedious. To avoid this tedium we can try to bound the quantities of interest using the inequalities developed in the body of this paper, and hence estimate them for solids of irregular shape.

1.3. Inequalities for set functions can be derived in several ways. The first and simplest way is just to compare, when possible, the values of the same set function for two solids, one of which can be totally included in the other. A second and more subtle way is to compare the values of the same set function for two solids which are related to each other by some process of symmetrization. (Symmetrization is the name given to a class of geometric transformations by which a solid is transformed into another which in some sense (depending on the precise nature of the transformation) is more symmetrical than its ancestor.) The first process of this kind was invented in 1836 by J. Steiner[4] who showed that this symmetrization leaves the volume of the solid unchanged while diminishing its surface area. Since constant reapplication of Steiner's symmetrization reduces all finite solids to spheres, Steiner was able to prove the classical isoperimetric theorem: Of all solids of a given volume, the sphere has minimum surface

* On leave from Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A., operated by Union Carbide Corporation for the U. S. Atomic Energy Commission.

[1] E. Inönü, Nuclear Sci. and Eng. 5, 248 (1959); M. H. L. Pryce, MSP-2A (declassified 1947), H.M. Stationery Office, London; E. Inönü, USAEC Report ORNL-2842, p. 134, 1959.

[2] N. C. Francis, J. C. Stewart, L. S. Bohl, and T. J. Krieger, *Proceedings of the Second United Nations International Conference on the Peaceful Use of Atomic Energy*, Vol. 16, p. 517, 1958.

[3] K. M. Case, G. Placzek, and F. de Hoffmann, *Introduction to the Theory of Neutron Diffusion* (U. S. Government Printing Office, Washington, D. C., 1953), Vol. I.

[4] G. Pólya and G. Szegö, *Isoperimetric Inequalities in Mathematical Physics* (Princeton University Press, Princeton, New Jersey, 1951).

area.[5] It follows from this theorem that $S^3 \geqq 36\pi V^2$ for any arbitrary solid of surface S and volume V; this "isoperimetric" inequality can now be used to bound the surface of any solid from below. Similarly to the surface area many other set functions, including those which interest us here, vary monotonically under symmetrizing transformations. Thus for each an isoperimetric inequality holds from which a bound may be derived.

Another rich source of inequalities are the variational expressions which exist for many set functions. These expressions, *when they are either of the maximum or minimum type*, can be used directly to obtain bounds by the appropriate choice of trial functions. Indirectly, they can be used as very convenient starting points for the derivation of the inclusion and isoperimetric inequalities mentioned above.

A third source of inequalities arises from the application of what may be termed the "standard" inequalities of analysis to the sum or integral representations of the quantities of interest. In particular, the law of the mean, the inequality connecting the geometric and arithmetic means of a function, some more general inequalities involving convex functions, and the classical inequality of Schwarz are all used later in just this connection.

1.4. Research of the type described above has had a very long history. The isoperimetric theorems connecting the perimeter and area of a circle and the surface area and volume of a sphere were known to the Greeks. The powerful concept of symmetrization, by whose use many more isoperimetric theorems can be proven, was invented by Steiner more than a century ago, and only shortly thereafter a number of interesting isoperimetric inequalities concerning certain physical rather than purely geometric quantities were announced. In 1856, B. Saint Venant conjectured an isoperimetric inequality involving the torsional rigidity of elastic prisms on inductive grounds. In 1877, several isoperimetric theorems concerning the principal frequency of vibration of plates and membranes were stated without proof by Lord Rayleigh, who also developed the variational method of obtaining bounds to a high degree of refinement. In 1903 a famous isoperimetric theorem regarding the electrical capacity of solids was stated by Poincaré, but accompanied by an incomplete proof.

In the years between about 1900 and the present, effort was given to the proof and elaboration of these conjectures by T. Carleman, G. Faber, E. Krahn, R. Courant, G. Szegö, G. Pólya, and others. These workers confined their attention largely to inequalities involving purely geometric quantities and those physical quantities arising from Laplace's, Helmoltz's, or related equations (i.e., arising in electrostatics, the study of vibrations of plates and membranes, hydrodynamics, the theory of elasticity, the theory of heat conduction, etc.). In 1951, G. Szegö and G. Pólya published a book[4] in which all the old results and many new ones are systematically described, thus elevating this research, which in these authors' words "moves somewhat outside the usual channels," to the level of a discipline.

The mathematics of neutron diffusion theory is extremely similar to the mathematics of the studies mentioned parenthetically in the last paragraph. The methods described by Pólya and Szegö can thus be systematically applied to diffusion theory. Indeed, in some cases all that is required is a simple reinterpretation of Pólya and Szegö's results. Regrettably, only those quantities can be really effectively treated which admit of a variational representation of the maximum or minimum type; this limitation is probably a fundamental one. A single paper pointed in this direction has already been written by Ackroyd and Ball[6] who studied the effect of Steiner symmetrization on critical mass in diffusion theory.

The mathematics of strict transport theory is, however, essentially different from that involved in electrostatics, hydrodynamics, etc., since it is governed not by second-order partial differential equations but rather by integral (or integro-differential) equations. To obtain inequalities from these equations the author previously developed several new techniques[7] whose application is extended in the present work.

1.5. The arrangement of this paper is as follows: In the next short section, the quantities of interest are precisely defined, and in the following section, the process of symmetrization is defined. Following that, in the fourth section, the subsequently used representations of the quantities of interest are derived. In the fifth section, the various theorems are stated and proved. In the sixth section, a discussion and some examples are given. The reader who wishes to avoid the laborious details of the proofs may read Secs. 2 and 3, the *statements* of the theorems in Sec. 5, and Sec. 6 without difficulty.

2. DEFINITIONS

2.1. In one-velocity transport theory the criticality of a bare, homogeneous reactor with isotropic scattering is governed by the integral equation

$$\phi(\mathbf{r}) = c\int_V K(|\mathbf{r}-\mathbf{r}'|)\phi(\mathbf{r}')d^3r', \qquad \text{(1a)}$$

where

$$K(r) = e^{-r}/(4\pi r^2). \qquad \text{(1b)}$$

[5] The word "isoperimetric" is actually a misnomer since the solids have the same volume (area) not the same surface area (perimeter). However, the theorem stated is a trivial deduction from the truly isoperimetric theorem: Of all solids of a given surface area, the sphere has maximum volume.

[6] R. T. Ackroyd and J. M. Ball, "On the conjecture that Steiner symmetrization reduces critical mass," UKAEA, Risley, Declassified Reprint WHC-(C)P-36, No. 8135, 1955.

[7] L. Dresner, Nuclear Sci. and Eng. 6, 63 (1959); 7, 260 (1960); 9, 151 (1961).

Here $\phi(\mathbf{r})$ is the flux of neutrons at $\mathbf{r}$, defined as the product of the neutron density at $\mathbf{r}$ and the (single) neutron velocity, c is an eigenvalue whose physical significance is that it is the average number of secondary neutrons emerging from each collision which will make the reactor just critical, and $K(|\mathbf{r}-\mathbf{r}'|)$ is an integral kernel which represents the probability density that a neutron originating at $\mathbf{r}'$ will have its next collision in a differential volume element at $\mathbf{r}$. V finally is the reactor volume. *Here, as in the rest of the paper, the mean free path* (m.f.p.) *in the reactor has been chosen as the unit of length.* The spectrum of eigenvalues of (1) is discrete and characterized by the fact that there is a lowest one c_0.[8] To this lowest eigenvalue corresponds an eigenfunction (flux) which is positive everywhere in the reactor interior, while to all other higher eigenvalues correspond eigenfunctions which change sign somewhere in the reactor interior. Thus the lowest eigenvalue alone corresponds to a physically realizable persisting state, and henceforth, only it will be called the critical multiplication. Its reciprocal P, which will prove more convenient to consider in what follows, is just the average first-collision probability of neutrons spatially distributed in the persisting flux mode. It will therefore simply be called the critical first-collision probability.

2.2. Another average first-collision probability of interest is that of the neutrons originating from a uniform, isotropic source inside V. It will henceforth just be called the first-collision probability and will be denoted by P_c.

2.3. If V is filled with a non-multiplying medium capable of scattering and absorption only, one can consider a second probability referring to a uniform, isotropic source, viz., the average absorption or non-escape probability P_a. This quantity is defined as the average probability that a neutron will be absorbed in V, rather than leak out of it, irrespective of how many scattering collisions it has had.

2.4. The diffusion theoretic calculation of criticality is governed not by (1) but rather by the much simpler second-order partial differential equation

$$D\nabla^2\phi(\mathbf{r})+(c-1)\phi(\mathbf{r})=0, \tag{2}$$

where D is the diffusion constant and c and ϕ are defined as before. The most common boundary condition used with (2) is that the flux ϕ shall vanish on some prescribed surface, usually lying just outside the actual reactor surface. For the considerations of this paper no formal distinction exists between this extrapolated surface and the actual reactor surface. Hence, to (2) we shall add the condition that ϕ vanish on the reactor surface S.

[8] A. M. Weinberg and E. P. Wigner, *The Physical Theory of Neutron Chain Reactors* (University of Chicago Press, Chicago, Illinois, 1958), pp. 406–10. See also the remarks by B. Davison, *Neutron Transport Theory* (Oxford University Press, London, 1957), pp. 195–6.

The eigenvalue problem

$$\nabla^2\phi(\mathbf{r})+B^2\phi(\mathbf{r})=0 \quad \text{in } V, \tag{3a}$$

$$\phi(\mathbf{r})=0 \quad \text{on } S, \tag{3b}$$

which arises from (2) has infinitely many discrete eigenvalues B_n^2, of which there is a lowest, B_0^2.[8] This lowest eigenvalue alone corresponds to a flux ϕ which does not change sign inside V. We call it the buckling. It is a purely geometric quantity depending only on the size and shape of V. In terms of it the criticality condition may be expressed as $c=1+DB_0^2$.

2.5. It is not possible to define an average first-collision probability in pure diffusion theory since the individual flights of the neutrons do not appear in the theory. On the other hand, the average absorption probability can be defined simply as the ratio of the total absorption rate in V to the total source rate in V. When the source is a uniform, isotropic one, we shall denote the corresponding diffusion-theoretic average absorption probability by P_{ad}.

3. SYMMETRIZATION

3.1. The process of Steiner symmetrization can be succinctly defined as follows: Symmetrization with respect to a plane Q changes the solid V into a solid V^* such that:

(i) V^* is reflection symmetric with respect to Q.

(ii) Any straight line perpendicular to Q that intersects one of the solids V and V^* intersects the other also. Both intersections have the same length.

(iii) The intersection with V^* consists of just one line segment. The plane Q is called the plane of symmetrization.

A simple picture of the process of symmetrization is this: The solid is broken into paraxial differential cylinders, all of which are perpendicular to Q. These cylinders are then slid parallel to their axes until their midpoints all lie in Q. In case any of the cylinders consists of several pieces these are slid together and then the resulting single cylinder is slid so that its midpoint lies in Q.

3.2. It is clear from the definition of Steiner symmetrization that it leaves the volume of the solid unchanged. The surface area, on the other hand, is either decreased or remains the same. This last result is not at all obvious; it was first proved by Steiner. A little thought will convince the reader that repeated Steiner symmetrization with respect to a suitably chosen infinitude of planes will change any finite solid into a sphere of equal volume. Furthermore, repeated symmetrization in a suitable infinitude of planes all containing a common line L will reduce any infinite cylinder to a right circular cylinder with axis L. From these last two statements isoperimetric theorems follow for any quantities which never increase (decrease) under Steiner symmetrization.

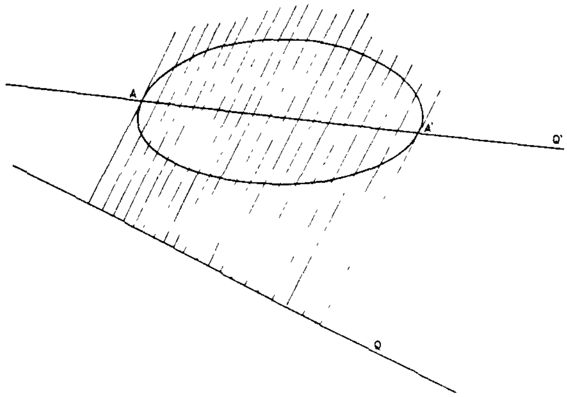

FIG. 1. Steiner symmetrization of a right elliptical cylinder.

3.3. Symmetrization of a right elliptical cylinder with respect to a plane containing its axis produces another right elliptical cylinder of the same base area but with the same or a smaller eccentricity. This fact can be used to show that any quantity which never increases (decreases) under Steiner symmetrization is a monotone increasing (decreasing) function of eccentricity, the base area being held fixed. It can be proven as follows: The midpoints of the differential cylindrical elements already lie in a plane Q' by a well-known property of the ellipse (see Fig. 1). Steiner symmetrization is then equivalent to rotating the plane Q' around the center of the ellipse until it is parallel to Q, all points sliding on lines perpendicular to Q, as though they were beads sliding on wires. This transformation is affine, and hence carries the original ellipse into another (in this case also of the same area). From the fact that one of the *new* axes AA' cannot be larger than the old major axis nor smaller than the old minor axis, but can be chosen arbitrarily in between, the desired conclusion follows.

3.4. Similar conclusions hold for spheroids but to prove them we must introduce the notion of Schwarz symmetrization. A solid V and a *solid of revolution* V^* can be related by Schwarz symmetrization as follows:

(i) Any plane perpendicular to the axis of revolution of V^* which intersects one of the solids V and V^* also intersects the other.

(ii) Both intersections have the same area.

Clearly, Schwarz symmetrization leaves the volume invariant. As it happens, the solid V^* which results from Schwarz symmetrization can also be obtained by an appropriately chosen infinitude of Steiner symmetrizations. We choose this infinitude as follows: All the planes of symmetrization contain the axis of the Schwarz symmetrization, but are otherwise distributed randomly in azimuth. This series of symmetrizations reduces all cross sections perpendicular to the common axis to circles, and thus has the same effect as Schwarz symmetrization.

If we first Steiner symmetrize a prolate spheroid we get, in general, an ellipsoid of equal volume whose largest principal axis lies in magnitude between the principal axes of the initial spheroid (again by an affine transformation). Schwarz symmetrization of the resulting ellipsoid with respect to largest principal axis gives another prolate spheroid of the same volume, but smaller eccentricity. If the original spheroid was oblate, the only difference is that the Schwarz symmetrization must be carried out around the smallest principal axis of the resulting ellipsoid. Since the largest (smallest) principal axis of the ellipsoid can be made as close to that of the original prolate (oblate) spheroid as desired, the eccentricity of the resulting spheroid can be made arbitrarily close to that of the original spheroid, from which the desired conclusion follows: Any quantity which never increases (decreases) under Steiner symmetrization is a monotone increasing (decreasing) function of the spheroid eccentricity, *the volume being held fixed.* (The spheroids being compared, however, must either be all prolate or all oblate.)

3.5. Steiner symmetrization of a hemisphere with respect to its diametral plane gives a volume equivalent oblate spheroid with a ratio of principal axes of $1:1:\frac{1}{2}$. Schwarz symmetrization with respect to a diameter gives a volume equivalent prolate spheroid with a ratio of principal axes of $1:\sqrt{2}/2:\sqrt{2}/2$.

Further discussion of Steiner symmetrization can be found in Pólya and Szegö's book.[4]

4. REPRESENTATIONS OF QUANTITIES OF INTEREST

4.1. A variational representation of the lowest eigenvalue c_0 of (1) is given by the Rayleigh quotient

$$P=\frac{1}{c_0}\geqq\frac{\int_V d^3r\int_V d^3r'\phi(\mathbf{r})K(|\mathbf{r}-\mathbf{r}'|)\phi(\mathbf{r}')}{\int_V \phi^2(\mathbf{r})d^3r}, \tag{4}$$

where $\phi(\mathbf{r})$ is any function. Equality in (4) occurs if and only if $\phi(\mathbf{r})=\phi^*(\mathbf{r})$, the true solution of (1). The sense of the inequality in (4) is related to the nature of the eigenvalue spectrum of (1), which we prove following the method of Davison[9]: Let $\phi_n(\mathbf{r})$ be the normalized, orthogonal eigenfunctions of (1) corresponding to the eigenvalues c_n. In terms of them, Davison writes the kernel $K(|\mathbf{r}-\mathbf{r}'|)$ in a bilinear Hilbert-Schmidt series[10]

$$K(|\mathbf{r}-\mathbf{r}'|)=\sum_{n=0}^{\infty}\frac{\phi_n(\mathbf{r})\phi_n(\mathbf{r}')}{c_n}, \tag{5}$$

[9] B. Davison, reference 8.
[10] See, for example, S. G. Mikhlin, *Integral Equations* (Pergamon Press, London, 1957), Chap. II, especially pp. 88–92.

indicating that (5) holds irrespective of whether the $\phi_n(\mathbf{r})$ form a complete set or not. Furthermore, $\phi(\mathbf{r})$ may be written

$$\phi(\mathbf{r})=\sum_{n=0}^{\infty} a_n\phi_n(\mathbf{r})+p(\mathbf{r}), \tag{6}$$

where $p(\mathbf{r})$ is orthogonal to every $\phi_n(\mathbf{r})$ and vanishes if the latter are complete. Then the right-hand side of (4) is given by

$$\sum_{n=0}^{\infty}(a_n^2/c_n)\Big/\left(\sum_{n=0}^{\infty}a_n^2+\int_V p^2(\mathbf{r})d^3r\right)\leqq c_0^{-1}, \tag{7}$$

since $c_0^{-1}\geqq c_1^{-1}\geqq c_2^{-1}\geqq\cdots$.

4.2. A useful and obvious representation of P_c is

$$P_c=V^{-1}\int_V d^3r\int_V d^3r'K(|\mathbf{r}-\mathbf{r}'|). \tag{8}$$

Another useful representation for P_c, whose derivation is outside the scope of this paper, is

$$P_c=1-\bar{l}^{-1}\int(1-e^{-l})f(l)dl, \tag{9}$$

where $f(l)$ is a certain normalized distribution of chord lengths l, whose mean $\bar{l}$ is equal to $4V/S$, i.e., to four times the volume-to-surface ratio.[3]

4.3. For P_a the only representation we shall use is a variational one. To derive it we must proceed as follows[11]: First consider the equation

$$H\psi^*=S, \tag{10}$$

where H is a positive, hermitian operator, and the star denotes the true solution of (10). A variational expression for the inner product (S,ψ^*) can be obtained by noting that for any ψ

$$|(S,\psi)|^2=|(H\psi^*,\psi)|^2 \leqq (H\psi^*,\psi^*)(\psi,H\psi)=(S,\psi^*)(\psi,H\psi). \tag{11}$$

Here, the inequality has been obtained from an obvious generalization of the Schwarz inequality in which (f,Hg) plays the role of the inner product between f and g. That this is possible depends on the positiveness of the operator H; for this property of H ensures that all norms (f,Hf) are non-negative. The proof that the operators H to which (11) is applied in this paper are positive as well as a proof of the generalized Schwarz inequality are to be found in the Appendix. From (11) it follows that

$$(S,\psi^*)\geqq|(S,\psi)|^2/(\psi,H\psi) \tag{12}$$

for any ψ, with equality if and only if $\psi=\psi^*$.

Now, when a uniform, isotropic source of unit total strength exists in a *non-multiplying* medium V, the flux is determined by the inhomogeneous equation

$$\phi^*(\mathbf{r})=c\int_V K(|\mathbf{r}-\mathbf{r}'|)\phi^*(\mathbf{r}')d^3r' +V^{-1}\int_V K(|\mathbf{r}-\mathbf{r}'|)d^3r', \tag{13}$$

where c is now just the ratio of scattering to total cross section in V. The terms on the rhs are contributions to ϕ^* from collided and uncollided neutrons. The total absorption rate in V is given by

$$(1-c)\int_V\phi^*(\mathbf{r})d^3r$$
$$=(1-c)\left[c\int_V d^3r\int_V d^3r'K(|\mathbf{r}-\mathbf{r}'|)\phi^*(\mathbf{r}') +V^{-1}\int_V d^3r\int_V d^3r'K(|\mathbf{r}-\mathbf{r}'|)\right], \tag{14a}$$
$$=(1-c)\left[cV\int_V S(\mathbf{r})\phi^*(\mathbf{r})d^3r+P_c\right]. \tag{14b}$$

The second line follows from (8) and the identification of $S(\mathbf{r})$ with the last term on the rhs of (13). Furthermore, H must then be given by

$$H=\int_V\cdots[\delta(\mathbf{r}-\mathbf{r}')-cK(|\mathbf{r}-\mathbf{r}'|)]d^3r'. \tag{15}$$

Now, applying (12) to the first term on the rhs of (14b), we have that for any function ϕ

$$P_a=(1-c)\int_V\phi^*(\mathbf{r})d^3r\geqq(1-c)\left\{\frac{c\left\{V^{-1}\int_V d^3r\int_V d^3r'\phi(\mathbf{r})K(|\mathbf{r}-\mathbf{r}'|)\right\}^2}{V^{-1}\int_V\phi^2(\mathbf{r})d^3r-cV^{-1}\int_V d^3r\int_V d^3r'\phi(\mathbf{r})K(|\mathbf{r}-\mathbf{r}'|)\phi(\mathbf{r}')}+P_c\right\}. \tag{16}$$

Since the source in V is of unit total strength, the lhs of (16) is equal to P_a and has been so denoted. Equality again occurs if and only if $\phi=\phi^*$.

[11] T. Kahan, G. Rideau, and P. Roussopoulos, *Memorial des Sciences Mathematiques, Fascicule CXXXIV* (Gauthier-Villars, Paris, 1956); N. C. Francis *et al.*, reference 2.

4.4. For B_0^2 we also employ a variational representation based on a Rayleigh quotient, viz.[12]:

$$B_0^2 \leqq -\int_V \phi\nabla^2\phi d^3r \Big/ \int_V \phi^2 d^3r$$

$$= \int_V |\nabla\phi|^2 d^3r \Big/ \int_V \phi^2 d^3r \quad (17)$$

for any (suitably well behaved) function ϕ which vanishes on S, the surface of V. The sense of the inequality follows from the eigenvalue structure of (3), that is from the fact that B_0^2 is the *lowest* eigenvalue. Equality again occurs if and only if ϕ equals the exact flux given by (3).

4.5. For eigenvalue problems with Hermitian operators another variational representation of the eigenvalue due to Weinstein[13] exists which gives both upper and lower bounds. However, the most forceful application of Weinstein's method unfortunately involves a number of assumptions which render it fundamentally unsound.

Let us begin by considering the quantity

$$M = (\{H-R\}\phi, \{H-a\}\phi)/(\phi,\phi), \quad (18)$$

where H is an hermitian operator, R is the Rayleigh quotient associated with ϕ, i.e., $(\phi,H\phi)/(\phi,\phi)$, and a is any number. Multiplying the numerator out shows that

$$M = W - R^2, \quad (19)$$

where

$$W = (H\phi,H\phi)/(\phi,\phi). \quad (20)$$

If now we set $\phi = \sum_{n=0}^{\infty} a_n\phi_n$, where ϕ_n are the eigenfunctions of H corresponding to eigenvalues λ_n *and are now assumed to form a complete set*, it can easily be shown that

$$M = \sum_{n=0}^{\infty} |a_n|^2(\lambda_n - R)(\lambda_n - a) \Big/ \sum_{n=0}^{\infty} |a_n|^2. \quad (21)$$

Now we designate by λ_m the eigenvalue to which R lies closest. Furthermore we choose $a=R$. In this case, it follows from (21) that

$$M \geqq (\lambda_m - R)^2, \quad (22)$$

from which it follows that

$$R+\sqrt{M} \geqq \lambda_m \geqq R-\sqrt{M}. \quad (23)$$

Now if in fact $\lambda_m=\lambda_0$, the lowest eigenvalue, (23) will give bounds for it. The Rayleigh quotient R itself is clearly a better upper bound than $R+\sqrt{M}$, but for the lower bound $R-\sqrt{M}$, there is as yet no competitor. Two remarks need to be made: First, of the identification $\lambda_m=\lambda_0$ we cannot in general be sure. Second, if the trial function ϕ differs from the true lowest eigenfunction of H by a small quantity of order ζ, λ_0 will differ from R by a quantity of order ζ^2 as is well known, but it will differ as we shall see presently from $R-\sqrt{M}$ by a quantity of order $|\zeta|$. Thus, (23) will provide only very rough bounds. This situation can be improved if we choose $a=\lambda_m+\lambda_{m+1}-R$ or $a=\lambda_m+\lambda_{m-1}-R$ according as R is $>$ or $<\lambda_m$. In these cases, respectively,

$$M \geqq (\lambda_{m+1}-R)(R-\lambda_m), \quad R>\lambda_m, \quad (24a)$$

$$M \geqq (R-\lambda_{m-1})(\lambda_m-R), \quad R<\lambda_m. \quad (24b)$$

If $\lambda_m=\lambda_0$, we can use (24a) and obtain

$$\lambda_0 \geqq R-(W-R^2)(\lambda_1-R)^{-1}. \quad (25)$$

In case some simple estimate of λ_1 can be made, (25) may provide a much sharper estimate of λ_0 than (23). To see how this may happen let us consider a trial function ϕ, which differs from the true lowest eigenfunction ϕ_0, by a quantity of order ϵ. Then, for small ϵ, it can easily be shown that $W-R^2$ is of order ϵ^2. Since $R-\lambda_0$ is also of order ϵ^2, it follows that the rhs of (25) differs from λ_0 by a quantity of the order of ϵ^2 at most. In the case of Eq. (23), however, the rhs and lhs both differ from λ_0 by a quantity of order ϵ.

In principle, Weinstein's method may be used to bound B_0^2; we shall say more of this application later.

4.6. Finally, we derive a variational representation for P_{ad} using (12) exactly as we did in treating P_a. This we do as follows: When a uniform, isotropic source of unit total strength exists in V, the flux ϕ is given according to diffusion theory by

$$-\nabla^2\phi^*(\mathbf{r})+\kappa^2\phi^*(\mathbf{r}) = (VD)^{-1} \quad \text{in } V, \quad (26a)$$

$$\phi^*(\mathbf{r}) = 0 \quad \text{on } S, \quad (26b)$$

where κ^2 is the inverse squared diffusion length and is given by $(1-c)/D$. The total absorption rate, equal here to P_{ad}, is given by

$$P_{ad} = (1-c)\int_V \phi^*(\mathbf{r})d^3r$$

$$= (1-c)VD\int_V S(\mathbf{r})\phi^*(\mathbf{r})d^3r, \quad (27)$$

where here $S(\mathbf{r})=(VD)^{-1}$. Using (12) plus the identification $H=-\nabla^2+\kappa^2$, we have that for any function ϕ

$$P_{ad} \geqq \frac{(1-c)VD\left(\int_V (VD)^{-1}\phi(\mathbf{r})d^3r\right)^2}{\int_V \phi(\mathbf{r})(-\nabla^2+\kappa^2)\phi(\mathbf{r})d^3r}, \quad (28a)$$

[12] R. Courant and D. Hilbert, *Methoden der Mathematischen Physik*, Erster Band (Springer Verlag, Berlin, 1931), sechstes Kapitel.

[13] D. H. Weinstein, Proc. Nat. Acad. Sci. **20**, 529 (1934); G. Goertzel and N. Tralli, *Some Mathematical Methods of Physics* (McGraw-Hill Book Company, Inc., New York, 1960), pp. 213–15.

or

$$P_{ad} \geqq \frac{1}{V} \frac{\left(\int_V \phi(\mathbf{r}) d^3r\right)^2}{\kappa^{-2}\int_V |\nabla\phi|^2 d^3r + \int_V \phi^2 d^3r}, \tag{28b}$$

if we use Green's theorem. Equality occurs if and only if $\phi = \phi^*$.

5. THEOREMS AND PROOFS

5.1. A versatile class of results which will prove extremely useful in estimating all of the quantities P, P_a, P_c, B_0^2, and P_{ad} is the class of inclusion theorems. Our results along this line are expressed in the following theorem:

Theorem 1. If V_1 can be entirely included in V_2, then $P(V_2) \geqq P(V_1)$, $B_0^2(V_2) \leqq B_0^2(V_1)$, $V_2P_{ad}(V_2) \geqq V_1P_{ad}(V_1)$, $V_2P_c(V_2) \geqq V_1P_c(V_1)$, and $V_2P_a(V_2) \geqq V_1P_a(V_1)$.

Proof. The proofs of all parts of the theorem follow the same general rationale and are based on the expressions (4), (17), (28b), (8), and (16). The details for the first three parts of the theorem are very similar and we shall only carry them through for P: Let $\phi_1^*(\mathbf{r})$ be the exact flux in V_1, i.e., the exact solution of (1) in V_1. Let us define a trial function $\phi_2(\mathbf{r})$ in V_2 by the stipulation: $\phi_2(\mathbf{r}) = \phi_1^*(\mathbf{r})$ in V_1, $\phi_2(\mathbf{r}) = 0$ otherwise. Then

$$P(V_1) = \frac{\int_{V_1} d^3r \int_{V_1} d^3r' \phi_1^*(\mathbf{r}) K(|\mathbf{r}-\mathbf{r}'|) \phi_1^*(\mathbf{r}')}{\int_{V_1} [\phi_1^*(\mathbf{r})]^2 d^3r}$$

$$= \frac{\int_{V_2} d^3r \int_{V_2} d^3r' \phi_2(\mathbf{r}) K(|\mathbf{r}-\mathbf{r}'|) \phi_2(\mathbf{r}')}{\int_{V_2} \phi_2^2(\mathbf{r}) d^3r}$$

$$\leqq P(V_2), \tag{29}$$

Q.E.D. The fourth part of the theorem, that for P_c, follows trivially from (8), the last part for P_a results from proving an inclusion theorem for the quantity $V(P_a - P_c)$ from (16), and then using the already proven result for P_c.

5.2. Another class of results arising from the comparison of different solids is expressed by:

Theorem 2. P, P_a, P_c, and P_{ad} all increase under Steiner symmetrization. B_0^2 decreases under Steiner symmetrization.

Proof. The proof of this theorem must be accomplished by two separate techniques. The first suffices to prove the theorem for the transport-theoretic quantities P, P_c, and P_a, while the second is reserved for the diffusion-theoretic quantities B_0^2 and P_{ad}. Let us begin with P_c in the form (8). Let us break the solid V up into paraxial differential cylindrical volume elements all of which are perpendicular to the plane of symmetrization, which for simplicity we take to be the xy plane. Let two of these cylindrical volume elements lie at x_1y_1 and x_2y_2 and have base areas dx_1dy_1 and dx_2dy_2, respectively. Let them intersect the surface of the (convex) solid V in z_1' and z_1'', z_2', and z_2'', respectively. The contribution of this pair to the multiple integral in (8) is

$$dP_c = \frac{dx_1dy_1dx_2dy_2}{V} \int_{z_1'}^{z_1''} dz_1 \int_{z_2'}^{z_2''} dz_2 \times K([(x_1-x_2)^2+(y_1-y_2)^2+(z_1-z_2)^2]^{\frac{1}{2}}). \tag{30a}$$

This can be rewritten as

$$dP_c = \frac{dx_1dy_1dx_2dy_2}{V} \int_{-\infty}^{+\infty} dz_1 \int_{-\infty}^{+\infty} dz_2 f_1(z_1) f_2(z_2) \times K([(x_1-x_2)^2+(y_1-y_2)^2+(z_1-z_2)^2]^{\frac{1}{2}}), \tag{30b}$$

where $f_1(z_1) = 1$ for $z_1' \leqq z_1 \leqq z_1''$ and zero otherwise, and similarly for f_2. Now let us rearrange the functions f_1 and f_2 in symmetrical decreasing order, i.e., let us replace $f_1(z_1)$ by a new function $\tilde{f}_1(z_1)$ which (i) is symmetric around $z_1 = 0$; (ii) is monotone decreasing; and (iii) has values between w and $w+dw$ over a set of the same measure as that for which f_1 has values between w and $w+dw$; and similarly with f_2. Since in (30b) K is a monotone decreasing function of $|z_1 - z_2|$, by theorem 380 of Hardy *et al.*,[14] this rearrangement increases the integral of (30b). The result of this rearrangement can also be seen to be just the integral

$$d\tilde{P}_c = \frac{dx_1dy_1dx_2dy_2}{V} \int_{-a/2}^{+a/2} dz_1 \int_{-b/2}^{+b/2} dz_2 \times K([(x_1-x_2)^2+(y_1-y_2)^2+(z_1-z_2)^2]^{\frac{1}{2}}), \tag{31}$$

where $a = z_1'' - z_1'$ and $b = z_2'' - z_2'$. But this is precisely the result of Steiner symmetrization, since the infinitesimal cylinders now have their midpoints in the plane of symmetrization. Thus, P_c increases under Steiner symmetrization.

The proofs for P and P_a follow similar lines. For P this is the procedure: If $\phi^*(\mathbf{r})$ is the exact solution of (1) in V, then

$$P(V) = \frac{\int_V d^3r \int_V d^3r' \phi^*(\mathbf{r}) K(|\mathbf{r}-\mathbf{r}'|) \phi^*(\mathbf{r}')}{\int_V [\phi^*(\mathbf{r})]^2 d^3r}. \tag{32}$$

[14] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities* (Cambridge University Press, London and New York, 1934).

On the other hand, for any arbitrary function $\phi^\dagger(\mathbf{r})$ in $V^\dagger$, the Steiner symmetrized solid

$$P(V^\dagger)\geqq\frac{\int_{V^\dagger}d^3r\int_{V^\dagger}d^3r'\phi^\dagger(\mathbf{r})K(|\mathbf{r}-\mathbf{r}'|)\phi^\dagger(\mathbf{r}')}{\int_{V^\dagger}[\phi^\dagger(\mathbf{r})]^2d^3r}. \quad (33)$$

To obtain $\phi^\dagger(\mathbf{r})$ we again divide the solid V into infinitesimal cylinders whose axes are perpendicular to the plane of symmetrization Q. $\phi^\dagger(\mathbf{r})$ is obtained by sliding each of these cylinders with the value of $\phi(\mathbf{r})$ fixed in the cylinder until all the midpoints lie in Q, and then rearranging $\phi^*(\mathbf{r})$ along each of these cylinders in symmetrical decreasing order with the midpoints of the cylinders as the centers of symmetry. Since rearranging of a function in symmetrical decreasing order does not alter the measure of the set over which it lies between specified values, the denominators of the rhs's of (32) and (33) are clearly equal. However (since ϕ^* and $\phi^\dagger$ are $\geqq 0$), by a repetition of the argument given above in connection with P_c, it can be shown that the numerator of the rhs of (33) exceeds that of (32). Thus $P(V^\dagger)\geqq P(V)$, Q.E.D.

A very similar method applied to (16), together with the already proven result for P_c, yields the announced result for P_a.

For B_0^2 and P_{ad} the approach is roughly similar but does not involve the notion of "symmetrical rearrangement in decreasing order." Instead a function $\phi^\dagger$ in V is used in (17) and (28b) which is obtained from ϕ^*, the exact flux in V, by symmetrizing its level surfaces. That is to say, if $\phi^*=C$ on the surface S_1 of V_1 then $\phi^\dagger=C$ on the surface $\tilde{S}_1$ of the symmetrized solid $V_1^\dagger$. From this definition it easily follows that

$$\int_V G(\phi^*(\mathbf{r}))d^3r=\int_{V^\dagger}G(\phi^\dagger(\mathbf{r}))d^3r, \quad (34)$$

where G is any function. However, as we shall presently see

$$\int_V|\nabla\phi^*|^2d^3r\geqq\int_{V^\dagger}|\nabla\phi^\dagger|^2d^3r. \quad (35)$$

From these last two equations the desired results for B_0^2 and P_{ad} easily follow.

Rather than (35) we shall prove a more general theorem, due to Pólya and Szegö,[4] whose method we follow without change: Let $F(x)$ be a concave-upwards, monotone increasing function of x. Then, with $\phi^\dagger$ and ϕ^* related as above,

$$I\equiv\int_V F(|\nabla\phi^*|)d^3r\geqq\int_{V^\dagger}F(|\nabla\phi^\dagger|)d^3r\equiv\tilde{I}. \quad (36)$$

To prove (36) we proceed as follows: Let S be a level surface of ϕ^* on which $\phi^*=C$; on $\tilde{S}$ let $\phi^\dagger=C$. Let an infinitesimal cylinder perpendicular to the plane of symmetrization (now chosen as the xy plane) and with base area $dA=dxdy$ intersect S at z_1 and z_2, and $\tilde{S}$ at $\pm z_0$. Let us compare the contributions to I and $\tilde{I}$ from the respective volumes lying inside $dxdy$ and corresponding, respectively, to values of ϕ^* and $\phi^\dagger$ between C and $C+dC$. In V there are two such volumes, one at z_1 of volume $dV=dAdC|dz_1/dC|$, and one at z_2 of volume $dV=dAdC|dz_2/dC|$. At z_1 the value of $|\nabla\phi^*|$ is

$$\left|\frac{dC}{dz_1}\right|\frac{1}{n_{z1}},$$

where n_{z1} is the z component of the outward normal to S at (x,y,z_1), and similarly at z_2. The contribution of the two volumes to I is then just

$$dI=\left[\left|\frac{dz_1}{dC}\right|F\left(\left|\frac{dC}{dz_1}\right|\frac{1}{n_{z1}}\right)+\left|\frac{dz_2}{dC}\right|F\left(\left|\frac{dC}{dz_2}\right|\frac{1}{n_{z2}}\right)\right]dAdC. \quad (37)$$

Since F is concave upwards, we may write according to theorem 204 of Hardy *et al.*,[14]

$$dI\geqq F\left(\left[\frac{1}{n_{z1}}+\frac{1}{n_{z2}}\right]\Big/\left[\left|\frac{dz_1}{dC}\right|+\left|\frac{dz_2}{dC}\right|\right]\right)\times\left(\left|\frac{dz_1}{dC}\right|+\left|\frac{dz_2}{dC}\right|\right)dAdC. \quad (38)$$

Next we note that $2z_0=z_2-z_1$. Hence,

$$2\left|\frac{dz_0}{dC}\right|=\left|\frac{dz_2}{dC}-\frac{dz_1}{dC}\right|. \quad (39a)$$

Since dz_2/dC and dz_1/dC must have *opposite* signs, (39a) can be rewritten

$$2\left|\frac{dz_0}{dC}\right|=\left|\frac{dz_2}{dC}\right|+\left|\frac{dz_1}{dC}\right|. \quad (39b)$$

Furthermore, since

$$2\frac{\partial z_0}{\partial x}=\frac{\partial z_2}{\partial x}-\frac{\partial z_1}{\partial x}, \quad (40a)$$

$$2\frac{\partial z_0}{\partial y}=\frac{\partial z_2}{\partial y}-\frac{\partial z_1}{\partial y}, \quad (40b)$$

we have by application of Minkowski's inequality

$$2\left(1+\left(\frac{\partial z_0}{\partial y}\right)^2+\left(\frac{\partial z_0}{\partial y}\right)^2\right)^{\frac{1}{2}}$$

$$=\left(4+\left(\frac{\partial z_2}{\partial x}-\frac{\partial z_1}{\partial x}\right)^2+\left(\frac{\partial z_2}{\partial y}-\frac{\partial z_1}{\partial y}\right)^2\right)^{\frac{1}{2}} \quad (41a)$$

$$\leqq\left(1+\left(\frac{\partial z_2}{\partial x}\right)^2+\left(\frac{\partial z_2}{\partial y}\right)^2\right)^{\frac{1}{2}}$$

$$+\left(1+\left(\frac{\partial z_1}{\partial x}\right)^2+\left(\frac{\partial z_1}{\partial y}\right)^2\right)^{\frac{1}{2}}. \quad (41b)$$

But (41b) is simply the equation

$$\frac{2}{n_{z0}}\leqq\frac{1}{n_{z1}}+\frac{1}{n_{z2}}, \quad (42)$$

since for any surface

$$(n_z)^{-1}=\left(1+\left(\frac{\partial z}{\partial x}\right)^2+\left(\frac{\partial z}{\partial y}\right)^2\right)^{\frac{1}{2}}.$$

Substituting (39b) and (42) in (38), and using the monotonicity of F, we obtain

$$dI\geqq 2F\left(\left|\frac{dC}{dz_0}\right|\frac{1}{n_{z0}}\right)\left|\frac{dz_0}{dC}\right|dAdC. \quad (43)$$

Comparing the rhs of (43) with that of (37), we see it is just the contribution to dI from the volume dV lying inside dA for which $C\leqq\phi^\dagger\leqq C+dC$. Thus, $dI\geqq d\tilde{I}$ and $I\geqq\tilde{I}$, Q.E.D.

It is worth noting that (41) and (42) express the essential step in showing that the surface is decreased by Steiner symmetrization; for further discussion of this as well as the preceding proof the reader is referred to reference 4.

5.3. In cylindrical and rectangular coordinates, among others, the diffusion equation is separable, so that solutions to problems involving right cylinders and rectangular parallelepipeds may be expressed in terms of results applicable to slabs and infinite cylinders. For two of the transport-theoretic quantities considered in this paper, viz., P and P_c, results are available which have to some extent the same effect. The first of these is:

Theorem 3. If V is the volume common to (i) two perpendicular slabs S_1 and S_2, or (ii) three mutually perpendicular slabs S_1, S_2, and S_3, or (iii) an infinite right cylinder C and a slab S perpendicular to it, then

(i) *$P_c(S_1)$ and $P_c(S_2)\geqq P_c(V)\geqq P_c(S_1)P_c(S_2)$.*

(ii) *$P_c(S_1)$ and $P_c(S_2)$ and $P_c(S_3)\geqq P_c(V)\geqq P_c(S_1)\times P_c(S_2)P_c(S_3)$.*

(iii) *$P_c(C)$ and $P_c(S)\geqq P_c(V)\geqq P_c(C)P_c(S)$; where $P_c(S_1)$ is the value of P_c for the slab S_1, etc. The same results also hold for P.*

Proof. We shall first prove the theorem for P_c: Let us introduce the characteristic function $f(\mathbf{r})$ of V defined by

$$f(\mathbf{r})=1 \quad \mathbf{r} \text{ in } V$$
$$=0 \quad \text{otherwise} \quad (44)$$

and its Fourier transform

$$f(\mathbf{k})=\int_V e^{i\mathbf{k}\cdot\mathbf{r}}d^3r;\quad f(\mathbf{r})=(2\pi)^{-3}\int_\infty f(\mathbf{k})e^{-i\mathbf{k}\cdot\mathbf{r}}d^3k. \quad (45a)$$

Introducing these into (8) one can show after some simple manipulation that

$$VP_c=(2\pi)^{-3}\int_\infty |f(\mathbf{k})|^2\frac{\arctan k}{k}d^3k. \quad (46a)$$

Here, use has been made of the fact that

$$K(\mathbf{k})=\int_\infty \frac{e^{-r}}{4\pi r^2}e^{i\mathbf{k}\cdot\mathbf{r}}d^3r=\frac{\arctan k}{k}. \quad (45b)$$

If in (8) one replaces $K(\mathbf{r})$ of $\delta(\mathbf{r})$, the Dirac delta function, (46a) becomes

$$V=(2\pi)^{-3}\int_\infty |f(\mathbf{k})|^2d^3k. \quad (46b)$$

Let us first consider part (i) of the theorem. Let V be a rectangular parallelepiped of sides b_1, b_2, and b_3. Then,

$$f(\mathbf{k})=\prod_{j=1}^{3}\int_{-b_j/2}^{+b_j/2}e^{ik_jx_j}dx_j=\prod_{j=1}^{3}f_j(k_j), \quad (47)$$

where x_j are the cartesian coordinates of $\mathbf{r}$, and k_j are the cartesian coordinates of $\mathbf{k}$. Now if b_2 and b_3 become very large, $f(\mathbf{k})$ is only appreciable when k_2 and k_3 are near zero. Hence, for large b_2 and b_3, $k\approx k_1$, and

$$(2\pi)^3VP_c=\prod_{j=1}^{3}\int_{-\infty}^{+\infty}|f_j(k_j)|^2\frac{\arctan k_1}{k_1}dk_j \quad (48a)$$

$$=(2\pi)b_2(2\pi)b_3\int_{-\infty}^{+\infty}|f_1(k_1)|^2\frac{\arctan k_1}{k_1}dk_1. \quad (48b)$$

The second equality follows from the one-dimensional analog of (46b). If we let b_2 and b_3 approach infinity, we then have

$$(2\pi)b_1P_c(b_1)=\int_{-\infty}^{+\infty}|f_1(k)|^2\frac{\arctan k}{k}dk, \quad (49a)$$

where $P_c(b_1)$ is the value of P_c for a slab of thickness b_1. On the other hand, if only b_3 becomes infinite,

$$(2\pi)^2b_1b_2P_c=\prod_{j=1}^{2}\int_{-\infty}^{+\infty}|f_j(k_j)|^2\frac{\arctan k}{k}dk_j, \quad (49b)$$

where $k^2=k_1^2+k_2^2$ and P_c is appropriate to the volume common to two perpendicular slabs. With k so defined it is true that

$$\frac{\arctan k_1}{k_1} \text{ and } \frac{\arctan k_2}{k_2} \geqq$$

$$\frac{\arctan k}{k} \geqq \frac{\arctan k_1}{k_1} \cdot \frac{\arctan k_2}{k_2}. \tag{50}$$

Substitution of (50) in (49b) and use of (49a) yields the conclusion $P_c(b_1)$ and $P_c(b_2) \geqq P_c \geqq P_c(b_1) \cdot P_c(b_2)$, Q.E.D. Parts (ii) and (iii) of the theorem for P_c are treated similarly.

For P we proceed as follows: If we introduce the Fourier transforms

$$\phi(\mathbf{k}) = \int_V \phi(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}} d^3r;$$
$$\phi(\mathbf{r}) = (2\pi)^{-3} \int_\infty \phi(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{r}} d^3k \tag{51}$$

of any trial function $\phi(\mathbf{r})$ *which vanishes outside of* V, then it follows from (4) that

$$P \geqq \int_\infty |\phi(\mathbf{k})|^2 \frac{\arctan k}{k} d^3k \Big/ \int_\infty |\phi(\mathbf{k})|^2 d^3k. \tag{52}$$

Let us choose V to be the same rectangular parallelepiped as before. As a trial function, $\phi(\mathbf{r})$, let us choose $\prod_{j=1}^3 \phi_j(x_j)$ where $\phi_j(x_j)$ is the *exact* solution of (1) in a slab of thickness b_j, and therefore vanishes for $|x_j| > \frac{1}{2}b_j$. Then

$$\phi(\mathbf{k}) = \prod_{j=1}^{3} \phi_j(k_j). \tag{53}$$

It now follows from (52) and (53) by reasoning quite similar to that used previously in connection with P_c that

$$P \geqq P(b_1)P(b_2)P(b_3), \tag{54}$$

where $P(b_j)$ is the critical first-collision probability for a slab of thickness b_j. This is the second inequality of part (ii) for P. Since $P(b_3)$ approaches unity as b_3 becomes infinite, (54) also gives the second inequality of part (i). The second inequality of part (iii) follows similarly. For P the first inequality follows from theorem 1.

5.4. The potentialities of this method are not yet exhausted and a slightly more subtle application of it yields the following very beautiful and powerful theorem for P_c:

Theorem 4. Part 1. Consider a convex solid V and an arbitrary line L in space. Let the position of a point on L be measured by a coordinate z. Let the intersection of V and a plane Q perpendicular to L at z be denoted by $A(z)$. Let $P_c(z)$ be the collision probability of an infinite right cylinder with $A(z)$ as base and L as axis. Then

$$P_c(V) \leqq V^{-1} \int_a^b dz A(z) P_c(z),$$

where a and b are the limits determined by the planes tangent to V and perpendicular to L.

Theorem 4. Part 2. Consider a convex solid V and an arbitrary plane Q in space. Let the position of a point on Q be measured by a coordinate two-vector $\boldsymbol{\varrho}$. Let the intersection of V and a normal L to Q at $\boldsymbol{\varrho}$ be a line segment of length $t(\boldsymbol{\varrho})$. Let $P_c(\boldsymbol{\varrho})$ be the collision probability of a slab of thickness $t(\boldsymbol{\varrho})$. Then

$$P_c(V) \leqq V^{-1} \int_A d\boldsymbol{\varrho}\, t(\boldsymbol{\varrho}) P_c(\boldsymbol{\varrho}),$$

where A is the projected area of V on Q.

Proof. To prove this theorem we again employ the method of characteristic functions introduced above. For the first part of the theorem let us write the Fourier transform $f(\mathbf{k})$ of the characteristic function of V as

$$f(\mathbf{k}) = \int_a^b e^{ik_3 z} dz \int_{A(z)} d\boldsymbol{\varrho} \exp(i\mathbf{k}_\rho \cdot \boldsymbol{\varrho}), \tag{55a}$$

where $\boldsymbol{\varrho}$ is the coordinate two-vector in the plane Q. Substituting (55a) in (46a) and rearranging the order of integration we obtain

$$(2\pi)^3 V P_c(V) = \int_{-\infty}^{+\infty} dk_3 \int_a^b e^{ik_3 z} dz \int_a^b e^{-ik_3 z'} dz' \int_\infty d\mathbf{k}_\rho \times \int_{A(z)} d\boldsymbol{\varrho} \exp(i\mathbf{k}_\rho \cdot \boldsymbol{\varrho}) \times \int_{A(z')} d\boldsymbol{\varrho}' \exp(-i\mathbf{k}_\rho \cdot \boldsymbol{\varrho}') K(\mathbf{k}). \tag{56}$$

Since the integrand with respect to k in (56) is positive for all $\mathbf{k}$ we can insert $K(\mathbf{k}_\rho) \geqq K(\mathbf{k})$ on the right-hand side, perform the k_3 and z' integrations in that order, and obtain

$$(2\pi)^3 V P_c(V) \leqq (2\pi) \int_a^b dz \left[\int_\infty d\mathbf{k}_\rho \int_{A(z)} d\boldsymbol{\varrho} \exp(i\mathbf{k}_\rho \cdot \boldsymbol{\varrho}) \times \int_{A(z)} d\boldsymbol{\varrho}' \exp(-i\mathbf{k}_\rho \cdot \boldsymbol{\varrho}') K(\mathbf{k}_\rho) \right]. \tag{57}$$

If we let a approach $-\infty$ and b approach $+\infty$ and imagine $A(z)$ is a fixed area, we obtain, by now familiar reasoning, the result that the square bracket in (57) is just $(2\pi)^2 A(z)P_c(z)$. But then

$$VP_c(V) \leqq \int_a^b dz A(z)P_c(z) \tag{58}$$

Q.E.D. The second part of theorem 3 is proven in an exactly similar manner, except we write $f(\mathbf{k})$ as

$$f(\mathbf{k}) = \int_A \exp(i\mathbf{k}_\rho \cdot \boldsymbol{\varrho}) d\boldsymbol{\varrho} \int_{a(\rho)}^{b(\rho)} e^{ik_3 z} dz, \tag{55b}$$

where $a(\boldsymbol{\varrho})$ and $b(\boldsymbol{\varrho})$ are the intersections of the normal L to Q at $\boldsymbol{\varrho}$ with V, interchange the order of integration so that $\mathbf{k}_\rho$, $\boldsymbol{\varrho}$ and, $\boldsymbol{\varrho}'$ are last, and substitute $K(k_3) \geqq K(\mathbf{k})$ for $K(\mathbf{k})$.

5.5. With a somewhat different use of Fourier transforms one can prove the following theorem:

Theorem 5. $P \geqq B_0^{-1} \arctan B_0$, *where* B_0^2 *defined by (3) is the buckling of the solid* V *to which* P *refers.*

This theorem is related to the so-called "second fundamental theorem of reactor physics"[15]; more will be said about this connection in Sec. 6.

Proof. If the angular integrals in (52) are performed it becomes

$$P \geqq \int_0^\infty W(k) \frac{\arctan k}{k} dk, \tag{59}$$

where $W(k)$ is a positive normalized weighting function of k only. Now $k^{-1} \arctan k$ is a convex-downwards decreasing function of k^2; thus by theorem 204 of Hardy *et al.*[14]

$$\int_0^\infty W(k) \frac{\arctan k}{k} dk \geqq \frac{\arctan k_0}{k_0}, \tag{60}$$

where

$$k_0^2 = \int_0^\infty W(k) k^2 dk \tag{61a}$$

$$= \int_\infty |\phi(\mathbf{k})|^2 k^2 d^3k \Big/ \int_\infty |\phi(\mathbf{k})|^2 d^3k. \tag{61b}$$

If we invert the transforms in Eq. (61b), noting that

$$i\nabla\phi(\mathbf{r}) = (2\pi)^{-3} \int_\infty \mathbf{k}\phi(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{r}} d^3k, \tag{62}$$

we find

$$k_0^2 = \int_V [\nabla\phi(\mathbf{r})]^2 d^3r \Big/ \int_V \phi^2(\mathbf{r}) d^3r. \tag{63}$$

Equation (62) can only converge if $\phi(\mathbf{r}) = 0$ on S the surface of V; for, otherwise, since $\phi(\mathbf{r}) = 0$ outside of V [see (51)], $\nabla\phi(\mathbf{r})$ will have an infinite singularity on S. Now, since $k^{-1} \arctan k$ is a monotone decreasing function, the best value for k_0^2 will be the smallest possible. But it follows from (63) and (17) that $(k_0^2)_{\min} = B_0^2$. Thus combining this result, (59), and (60) we have $P \geqq B_0^{-1} \arctan B_0$, Q.E.D.

5.6. As noted in the introduction, a rich source of inequalities are the variational representations of the different quantities; e.g., by the simple choice of a constant trial function $\phi = 1$ in (4) and (16) one obtains the following two theorems immediately[16]:

Theorem 6. $P \geqq P_c$.

Theorem 7. $P_a \geqq (1-c)P_c/(1-cP_c)$.

By combining the variational technique with an application of Schwarz's inequality, one can furthermore prove:

Theorem 8. $P_a \leqq (1-c)P_c/(1-cP)$.

Proof. Let $\phi = \phi^*$, the exact solution of (13). Then

$$P_a = (1-c) \left[\frac{c\left\{ V^{-1} \int_V d^3r \int_V d^3r' \phi^*(\mathbf{r}) K(|\mathbf{r}-\mathbf{r}'|) \right\}^2}{V^{-1} \int_V [\phi^*(\mathbf{r})]^2 d^3r - cV^{-1} \int_V d^3r \int_V d^3r' \phi^*(\mathbf{r}) K(|\mathbf{r}-\mathbf{r}'|) \phi^*(\mathbf{r}')} + P_c \right] \tag{64a}$$

$$\leqq (1-c) \left[\frac{cV^{-2} \left\{ \int_V d^3r \int_V d^3r' K(|\mathbf{r}-\mathbf{r}'|) \right\} \left\{ \int_V d^3r \int_V d^3r' \phi^*(\mathbf{r}) K(|\mathbf{r}-\mathbf{r}'|) \phi^*(\mathbf{r}') \right\}}{V^{-1} \int_V [\phi^*(\mathbf{r})]^2 d^3r - cV^{-1} \int_V d^3r \int_V d^3r' \phi^*(\mathbf{r}) K(|\mathbf{r}-\mathbf{r}'|) \phi^*(\mathbf{r}')} + P_c \right]. \tag{64b}$$

The application of Schwarz's inequality here is made in the same way as in (11). Dividing the numerator and denominator in (64b) by

$$cV^{-1} \int_V d^3r \int_V d^3r' \phi^*(\mathbf{r}) K(|\mathbf{r}-\mathbf{r}'|) \phi^*(\mathbf{r}')$$

and using (4) and (8) one finds $P_a \leqq (1-c)P_c/(1-cP)$, Q.E.D. By exactly the same technique as above applied to (28b) one can prove:

[15] A. M. Weinberg and E. P. Wigner, reference 8, pp. 397–406.

[16] Theorem 6 is due to P. A. M. Dirac, "Approximate rate of neutron multiplication for a solid of arbitrary shape and uniform density," declassified British Report MS-D-5, Part I, 1943. Theorem 7 is originally due to H. Hurwitz, Jr., according to N. C. Francis *et al.* (reference 2); see also: G. W. Stuart, Nuclear Sci. and Eng. 2, 617 (1957).

Theorem 9. $P_{ad} \leqq (1+B_0^2/\kappa^2)^{-1}$.

Noting that in diffusion theory the critical first-collision probability

$$P_d = (1+DB_0^2)^{-1} = (1+(1-c)(B_0^2/\kappa^2))^{-1},$$

the last theorem can be rewritten as

$$P_{ad} \leqq (1-c)P_d/(1-cP_d).$$

(Here c means only the fraction of scattering per collision.) In a sense this statement is weaker than that of theorem 8, since there P is replaced by $P_c \leqq P$ in the numerator resulting in a lower upper-bound. Also somewhat weaker than the theorems already proved and a consequence of them are the physically obvious inequalities $P \geqq P_c \geqq P_a$ and $P_d \geqq P_{ad}$.

5.7. For the strictly transport-theoretic quantities which can be variationally represented, viz., P and P_a, use of more complicated trial functions than those mentioned above leads to great difficulties [although theorem 5 results in a manner of speaking from the use of a diffusion-theoretic trial function in (4)]. For the diffusion-theoretic quantities the case is otherwise, and some elegant and useful results can be obtained by suitable choice of trial fluxes. These results can most easily be expressed in terms of a certain "effective radius" of a solid R_0, which is defined by

$$R_0^{-2} = \frac{1}{3V}\int\int (\mathbf{r}\cdot\mathbf{n})^{-1} dS, \tag{65}$$

where V is the volume of the solid, $\mathbf{r}$ is the radius vector from some fixed point O in the interior of V to any point Q on the surface, $\mathbf{n}$ is the outward normal at Q, and dS is the infinitesimal element of surface at Q. When V has a center of symmetry it will be chosen as O, otherwise the choice is left open and R_0 will be a function of O. In terms of this "radius" one can prove:

Theorem 10. For any finite solid V, $B_0^2 \leqq \pi^2/R_0^2$.

Proof. We use the method of prescribed level surfaces described by Pólya and Szegö[4]: With O as origin let the equation of the surface S of V be $r = R(\boldsymbol{\omega})$, where $\boldsymbol{\omega}$ is a unit vector giving the direction of $\mathbf{r}$, and r is the latter's magnitude. Let us choose the level surfaces of the trial flux $\phi(\mathbf{r})$ to be the surfaces $r = uR(\boldsymbol{\omega})$ where $0 \leqq u \leqq 1$. (The point $u=0$ is O, the surface $u=1$ is S.) Furthermore, let us set $\phi(uR(\boldsymbol{\omega})) = f(u)$, where $f(1)=0$ and $f(u)$ is as yet otherwise undetermined.

Now the volume dV between the surfaces u and $u+du$ and lying inside an infinitesimal cone whose apex is at O and whose intersection with S is dS, is given by $dV = u^2du(\mathbf{r}\cdot\mathbf{n})dS$. Furthermore, at Q, $|\nabla\phi|$ is given by $|df/du|(\mathbf{r}\cdot\mathbf{n})^{-1}$. Using these relations in (17) gives

$$B_0^2 \leqq \frac{\int_0^1 (df/du)^2 u^2 du \int\int (\mathbf{r}\cdot\mathbf{n})^{-1} dS}{\int_0^1 f^2u^2du \int\int (\mathbf{r}\cdot\mathbf{n}) dS} = \frac{\int_0^1 (df/du)^2 u^2 du}{\int_0^1 f^2 u^2 du} \cdot R_0^{-2} \tag{66}$$

since

$$V = \int_0^1 u^2 du \int\int (\mathbf{r}\cdot\mathbf{n}) dS = \tfrac{1}{3}\int\int (\mathbf{r}\cdot\mathbf{n}) dS.$$

The best choice of $f(u)$ is that function which will make the rhs of (66) a minimum. We can formulate the requirements on $f(u)$ conveniently through the variational equations

$$\delta\int_0^1 (df/du)^2 u^2 du = 0, \tag{67a}$$

$$\int_0^1 u^2 f^2 du = 1, \tag{67b}$$

$$f(1) = 0. \tag{67c}$$

The corresponding Euler-Lagrange differential equation for $f(u)$ is

$$\frac{d^2f}{du^2} + \frac{2}{u}\frac{df}{du} + \gamma^2 f = 0, \tag{68}$$

where γ is an undetermined Lagrange multiplier. The regular solution of (68) is

$$f(u) = u^{-\frac{1}{2}} J_{\frac{1}{2}}(\gamma u) \propto u^{-1}\sin\gamma u. \tag{69}$$

To satisfy the requirement $f(1)=0$, γ must be chosen as π. Furthermore, by a partial integration the ratio $\int_0^1 (df/du)^2 u^2 du / \int_0^1 f^2 u^2 du$ can be shown to equal $\gamma^2 = \pi^2$ if $f(u)$ satisfies (68). Substituting this value in (66) gives the desired result. Finally, equality occurs when V is a sphere. By an exact repetition of the foregoing argument one can prove:

Theorem 11. For any infinite right cylinder,

$$B_0^2 = \alpha^2/(R_0')^2,$$

where α is the first root of the Bessel function J_0 ($=2.405$), and R_0' is defined by

$$(R_0')^{-2} = (2A)^{-1}\int (\mathbf{r}\cdot\mathbf{n})^{-1} ds. \tag{70}$$

Here A is the base area of the cylinder, $\mathbf{r}$ is the two-dimensional radius vector from some arbitrary fixed point O in the interior of A to a point Q on the perimeter of A, $\mathbf{n}$ is the outward normal at Q, and ds is the infinitesimal element of perimeter at Q. Equality occurs here for right circular cylinders.

5.8. With a very similar technique one can prove the theorems:

Theorem 12. For any finite solid,

$$P_{ad}\geqq 1-\frac{3}{\kappa R_0}\left\{\coth\kappa R_0-\frac{1}{\kappa R_0}\right\}.$$

Theorem 13. For any infinite right cylinder,

$$P_{ad}\geqq 1-\frac{2}{\kappa R_0'}\frac{I_1(\kappa R_0')}{I_0(\kappa R_0')}.$$

Proof. By using the same level lines as in the last section (5.7) in (28b) we can rewrite it in the case of a finite solid as

$$P_{ad}\geqq\frac{3\left(\int_0^1 fu^2du\right)^2}{\int_0^1 f^2u^2du+(\kappa R_0)^{-2}\int_0^1(df/du)^2u^2du},\qquad(71)$$

where $f(u)$ is again undetermined, save $f(1)=0$. The requirement that $f(u)$ be so chosen as to make the rhs of (71) a maximum leads to the Euler-Lagrange equation

$$\frac{d^2f}{du^2}+\frac{2}{u}\frac{df}{du}-(f-\gamma)\kappa^2R_0^2=0;\quad f(1)=0,\qquad(72)$$

where γ is again an undetermined constant. It enters $f(u)$ however, only as a multiplicative factor and hence does not affect the rhs of (71). Indeed,

$$f(u)=\gamma\left\{1-\frac{\sinh(\kappa R_0u)}{u\sinh(\kappa R_0)}\right\}.\qquad(73)$$

With the help of this expression and a partial integration, the rhs of (71) can be evaluated and yields the theorem as stated. Equality occurs again for spheres. When the solid is an infinite right cylinder the proof is similar. In theorem 13 equality occurs for right circular cylinders.

5.9. For the buckling, the variational treatment can be extended by application of the method of Weinstein. To carry this method through we must not only calculate the Rayleigh quotient R, as is done in the last section, but also the quantity W of (20). Choosing ϕ exactly as in Sec. 5.7 [i.e., choosing $f(u)$ given by (69) in the case of finite solids, etc.], we find that for finite solids and infinite right cylinders, respectively,

$$W=\pi^4/R_1^4\quad\text{(finite solids)},\qquad(74a)$$

$$W=\alpha^4/(R_1')^4\quad\text{(infinite right cylinders)},\qquad(74b)$$

where

$$R_1^{-4}=\tfrac{1}{3}V^{-1}\int\int(\mathbf{r}\cdot\mathbf{n})^{-3}dS\qquad(75a)$$

and

$$(R_1')^{-4}=\tfrac{1}{2}A^{-1}\int(\mathbf{r}\cdot\mathbf{n})^{-3}dS\qquad(75b)$$

We have not stated these results in the form of a theorem because of the somewhat uncertain nature of our application of Weinstein's method. The proof of (74) is as follows:

Proof. We consider only finite solids; the proof for cylinders is very similar. Since $H=-\nabla^2$, we need an expression for ∇^2 in terms of the variable u. This we obtain by noting first that, from our previous expressions for dV and $|\nabla\phi|$, it follows that

$$\int_V\nabla\psi\cdot\nabla\phi dV=\int\int(\mathbf{r}\cdot\mathbf{n})^{-1}dS\int_0^1\frac{df}{du}\frac{dg}{du}u^2du,\qquad(76)$$

where $\phi(\mathbf{r})=f(u)$ and $\psi(\mathbf{r})=g(u)$. An integration by parts in both sides of (76) gives [since both $\phi(\mathbf{r})$ and $\psi(\mathbf{r})$ vanish on S]:

$$-\int_V\psi\nabla^2\phi dV$$
$$=-\int\int dS(\mathbf{r}\cdot\mathbf{n})^{-1}\int_0^1 g(u)\left\{\frac{d^2f}{du^2}+\frac{2}{u}\frac{df}{du}\right\}u^2du\qquad(77a)$$
$$=-\int dS(\mathbf{r}\cdot\mathbf{n})u^2du\cdot g(u)\cdot\left\{\frac{d^2f}{du^2}+\frac{2}{u}\frac{df}{du}\right\}(\mathbf{r}\cdot\mathbf{n})^{-2}.\qquad(77b)$$

Since the first factor in the integrand on the rhs of (77b) is dV, and the second factor is ψ, *which is arbitrary*, it must be that

$$\nabla^2\phi=\left\{\frac{d^2f}{du^2}+\frac{2}{u}\frac{df}{du}\right\}(\mathbf{r}\cdot\mathbf{n})^{-2}.\qquad(78)$$

Then

$$W=\frac{\int_0^1\left\{\frac{d^2f}{du^2}+\frac{2}{u}\frac{df}{du}\right\}^2u^2du\int\int(\mathbf{r}\cdot\mathbf{n})^{-3}dS}{\int_0^1u^2f^2du\int\int(\mathbf{r}\cdot\mathbf{n})dS}.\qquad(79)$$

If we furthermore require f to satisfy (68) with $\gamma=\pi$, we get

$$W=\pi^4\int\int(\mathbf{r}\cdot\mathbf{n})^{-3}dS\Big/3V,\qquad(80)$$

which is identical with (74a), Q.E.D.

5.10. The variational theorems of the last three sections explicitly state relations between B_0^2 and P_{ad} and certain effective "radii." A similar theorem for P_c which has an origin quite different from a variational principle is:

Theorem 14. $P_c \geqq 1-\bar{l}^{-1}(1-\exp(-\bar{l}))$; $\bar{l}-4V/S$.

Proof. The proof is based on the use of the following inequality of Hardy *et al.*[14] (theorem 184) in (9):

$$\int f(l)e^{-l}dl \geqq \exp\left[-\int f(l)l\,dl\right] \tag{81}$$

if $f(l)$ is a normalized probability density.

6. DISCUSSION AND EXAMPLES

6.1. A number of remarks will be given below concerning the question of when a particular theorem can be expected to yield a limit close to the actual value of the quantity being estimated and when not. All of these remarks, however, only apply in extreme cases and, in general, the limits supplied by the various theorems must be compared to see which are best.

Theorem 1. This "inclusion" theorem works best when the volumes of the solids being compared are not too different. Thus inscribing a sphere in a cube may give fairly good limits while inscribing one in a long, thin cylinder should give rather bad limits.

Theorem 2. When using the isoperimetric corollary to theorem 2, viz., "Of all finite solids of a given volume the sphere has maximum P, P_a, P_c, P_{ad}, and minimum B_0^2," the best results will be obtained with equilateral or "sphere-like" solids. Thus cubes, cylinders with height and diameter equal, or ellipsoids of low eccentricity are all suitable for the application of theorem 2, while solids which are much longer in some directions than in others are quite unsuitable. Similar remarks refer to the isoperimetric corollary for cylinders.

Theorem 3. In discussing this theorem let us consider for the sake of argument situation (iii) of the hypothesis, viz., the perpendicular intersection of a cylinder C and a slab S. Furthermore, although we only discuss P_c in what follows, similar remarks apply to P. If the radius of the cylinder C is large, then $P_c(S)$ will be a very good upper limit and $P_c(S)P_c(C)$ a very good lower limit for P_c of the intersection solid. This is simply because $P_c(C)$ does not differ very much from unity, and thus the upper and lower limits do not differ very much from each other. Quite a similar conclusion holds if the slab is very thick. On the other hand, if the thickness and radius are both small, then it immediately follows that $P_c(C)\cdot P_c(S)$ will be a very bad lower limit since the value of the product of the two P_c's falls much more rapidly with decreasing size than the P_c of the intersection solid. One expects that the upper limit in this latter case will also not be very close to the exact value for the following reason: In the intersection solid neutrons born at any point and with any direction of velocity are within a short flight of the edge. In the bounding solids, i.e., cylinder or slab, however, neutrons whose velocity is nearly parallel to the elements of the surface are removed by a long flight from the edge. In sum, theorem 3 will work best for large solids.

Theorem 4. This theorem essentially generalizes the *upper* limits obtainable from theorem 3, and roughly similar remarks apply to it.

Theorem 5. This theorem is based essentially on the choice of a diffusion-theoretic trial function in the variational expression for P. Thus it ought to be a close underestimate for relatively large reactors where diffusion theory is approximately valid. This tendency is reinforced by the fact that for large reactors both P and B_0^{-1} arctan B_0 approach unity.

Theorem 5 is related to the so-called "second fundamental theorem of reactor theory"[15] which, for a one-velocity, bare reactor with isotropic scattering, equates P and B_0^{-1} arctan B_0, but which permits adjustment of B_0 through the introduction of an extrapolated surface. The requirement that the flux vanish on an extrapolated surface has the effect of decreasing B_0 and raising the value of B_0^{-1} arctan B_0. This will, in general, prolong agreement of this latter formula with P to much smaller sizes than otherwise, but render the sign of the error uncertain. Moreover the choice of an extrapolated surface is arbitrary although quite reasonable procedures can be worked out based on the extrapolation distance one obtains in Milne's problem. This arbitrariness renders the existence of any general inequality involving P and an extrapolated buckling unlikely, so that theorem 5 seems the strongest theorem we can prove in this direction.

Theorem 6. Theorem 6 is based on the choice of a flat trial flux and therefore should be best for small solids, in which the curvature of the true flux is not too large. Furthermore, since both P and P_c must both approach unity for large solids, this theorem may even provide useful estimates for large solids.

Theorems 7, 8. These two theorems are discussed together because: If P and P_c are close to one another, then *ceteris paribus* the upper and lower limits provided by these theorems should also be close. This will occur particularly for small solids as mentioned in the last paragraph although it should be pointed out, for example, that for slabs of any thickness P and P_c *never* differ by more than 3%. Trouble can develop, however, when c, the scattering fraction, is near unity and the solid is large, so that P is near unity, too. Then the upper limit provided by theorem 8 may grow inconveniently large.

Theorem 9. Theorem 9 has the same meaning in diffusion theory as theorem 8 has in the strict transport theory. One expects therefore, that for small solids the two sides of the inequality are not widely different in analogy with the discussion above. This can be directly supported as follows: Aside from the use of Schwarz's inequality, the chief step in the derivation of theorem 9 (or for that matter 8 too) is the use of the flux originating from a uniform isotropic source as a trial value for the critical flux (i.e., as a trial flux in the variational principle for B_0^2 or P). Since the first of these fluxes is concave upwards and the second concave downwards,

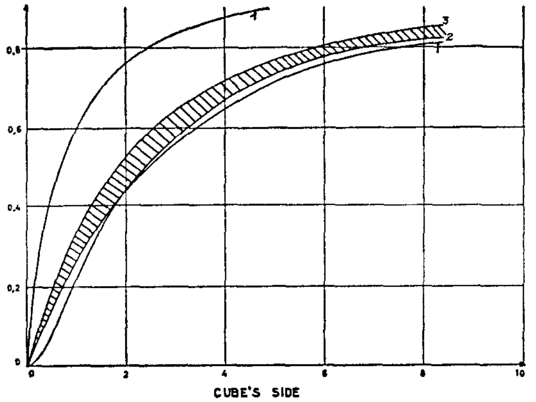

FIG. 2. Limits for P_c of cubes as a function of the length of the side. The limits shown are: 1, the lower limit from theorem 14; 2, the cube of P_c of the circumscribing slab; 3, P_c of the volume equivalent sphere; and 4, P_c of the circumscribing slab. The true value must lie in the cross-hatched area.

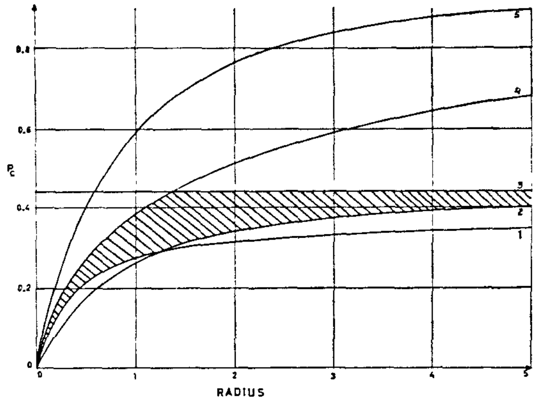

FIG. 3. Limits for P_c of $\frac{1}{2}$ m.f.p. thick disks of various radii. The limits shown are: 1, the lower limit from theorem 14; 2, the product of the P_c's of the circumscribing slab and cylinder; 3, P_c of the circumscribing slab; 4, P_c of the volume equivalent sphere; 5, P_c of the circumscribing cylinder. The true value must lie in the cross-hatched area.

one can be a good trial value for the other only when the curvature of both is negligible. This happens, however, only in small solids. Furthermore, when the flux is essentially flat, the use made of Schwarz's inequality also entails little error.

Theorems 10–13. Since in theorems 10 and 12 equality holds for spheres, these theorems should give very close limits for "sphere-like" solids. However, these need not be the only solids for which they give close limits, since the trial fluxes we have used are quite reasonable for many solids. Similar remarks apply to theorems 11 and 13.

Theorem 14. Regrettably little else can be said about when to expect close estimates from this theorem, save that it has the right values for very large and very small bodies.

6.2. As our first example let us consider the estimation of P_c for cubes. According to theorem 2 the value of P_c for the volume equivalent sphere is an upper limit for P_c of a cube. Another upper limit is P_c for a circumscribing slab, that is a slab of thickness equal to the cube's side. This follows from theorem 3. Theorem 3 also gives a lower limit, namely the cube of P_c for the circumscribing slab. Finally, theorem 14 gives a lower limit. These limits are plotted in Fig. 2 as a function of the cube's side; the true value of P_c for a cube must lie in the shaded region.

A number of observations concerning this figure are relevant. In the first place, for cubes, P_c of the circumscribing slab is a very bad overestimate as one might originally have expected. Indeed, in the range of sides from 0.2 to 8.0 m.f.p., this upper limit is much larger than that given by the volume equivalent sphere. On the other hand, for cubes one expects P_c for the volume equivalent sphere to be a fairly close over-estimate and this is borne out in the case at hand by its nearness to the lower limits in Fig. 2. For large cubes the underestimate provided by the cube of P_c of the circumscribing slab is the better of the two considered; however, for small cubes for which P_c of the circumscribing slab becomes small, its cube becomes extremely small and provides a rather useless limit. Thus, for cubes whose side is less than 2.0 m.f.p., the better lower limit is that of theorem 14.

For solids which are not "sphere-like" P_c of the volume equivalent sphere is usually a gross overestimate. This can be clearly seen in Fig. 3 where limits for P_c of disks $\frac{1}{2}$ m.f.p. thick and of various radii are plotted. These limits are: P_c for a slab $\frac{1}{2}$ m.f.p. thick (an upper limit by theorem 3); P_c for an infinite cylinder of radius equal to the disk radius (an upper limit by theorem 3); the product of these two numbers (a lower limit by theorem 3); P_c for the volume equivalent sphere (an upper limit by theorem 2); and the lower limit given by theorem 14. For disks for which the radius is very much greater than the thickness, P_c for the volume equivalent sphere is much larger than P_c for the circumscribing slab, whereas when the thickness and radii are comparable this situation is reversed. Not surprising is the further fact that the lower limit from theorem 14 is better than that from theorem 3 when the disk radius is small and worse when the disk radius is large.

Cubes and disks belong to that special class of solids which can be formed by the orthogonal intersection of slabs and cylinders. When we consider solids not belonging to this class we can no longer use theorem 3; however, we can use theorem 4 instead. The latter, however, supplies only an upper limit; hence, we have for the upper limit but two choices: the one just mentioned (theorem 4) and P_c of the volume equivalent sphere. For a lower limit we can use only theorem 14 in general.

Oblate spheroids are an excellent example of solids not belonging to this special class. A simple and useful upper limit for P_c for them can be obtained from theorem 4, part 2 by choosing the plane Q perpendicular

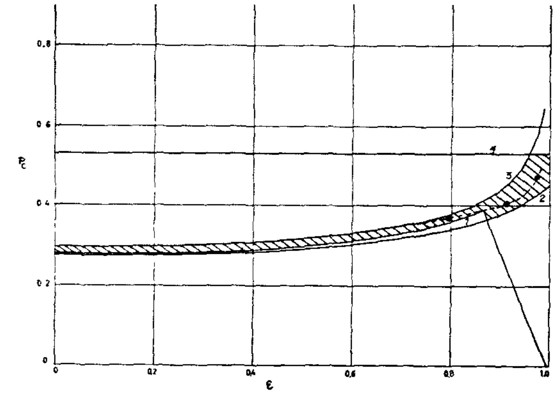

FIG. 4. Limits for P_c of oblate spheroids whose minor axes are 1 m.f.p. long as a function of eccentricity. The limits shown are: 1, a lower limit based on the P_c of certain hemispheres related to the spheroids by Steiner symmetrization (see text); 2, the lower limit from theorem 14; 3, P_c of the volume equivalent sphere; 4, an upper limit from the factorization theorem 4. The true value must lie in the cross-hatched area. Given also are three exact values available from reference 3.

to the minor axis of the spheroid. If the half-length of the minor axis is b it can then be shown from theorem 4 that

$$P_c \leqq 3\int_0^1 u^2 P_{cs}(2bu)du, \tag{82}$$

where P_{cs} is the value of P_c for a slab of thickness $2bu$. Interestingly enough, this limit depends only on the length $2b$ of the minor axis and not at all on the eccentricity of the spheroid!

In Fig. 4 the limit (82), the value of P_c for the volume equivalent sphere, and the limit from theorem 14 have been plotted as functions of the eccentricity ϵ for oblate spheroids with $2b=1$ m.f.p. ϵ is defined by

$$\epsilon^2 = 1 - b^2/a^2, \tag{83}$$

where a is the semi-major axis of the spheroid. The limit

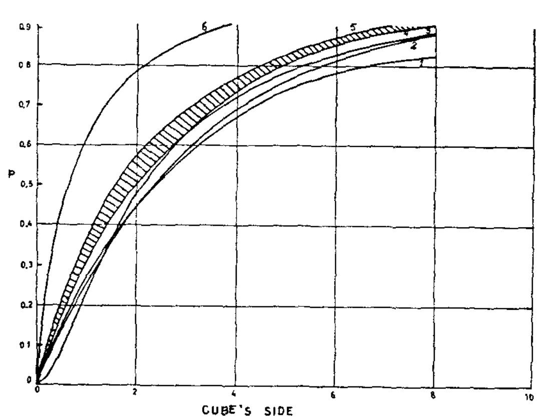

FIG. 5. Limits for P of cubes as a function of the length of the side. The limits shown are: 1, the lower limit for P_c taken from Fig. 2; 2, B_0^{-1} arctan B_0; 3, P of the inscribed sphere; 4, the cube of P of the circumscribing slab; 5, P of the volume equivalent sphere; and 6, P of the circumscribing slab. The true value must lie in the cross-hatched area.

of theorem 14 has been calculated with the aid of the relations[3]:

$$\bar{l} = (8/3)b/F(\epsilon), \tag{84a}$$

$$F(\epsilon) = 1 + \epsilon^{-1}(1-\epsilon^2)\tanh^{-1}\epsilon. \tag{84b}$$

In addition to these limits one other can be derived which is applicable only to spheroids. If $\bar{P}_c(V)$ is the first-collision probability for a *hemisphere* of volume V, we can write for oblate spheroids:

$$P_c(b=\tfrac{1}{2};\epsilon) = P_c(V;\epsilon) \geqq P_c(V;\epsilon=\sqrt{3}/2) \geqq \bar{P}_c(V) \quad \text{if } \epsilon \leqq \sqrt{3}/2. \tag{85}$$

The first inequality follows from the monotonic decreasing behavior of the first-collision probability for spheroids with eccentricity proved in Sec. 3.4; the second from the fact that an oblate spheroid of eccentricity $\sqrt{3}/2$ results from Steiner symmetrizing a hemisphere in its diametral plane (Sec. 3.5). When $\epsilon \geqq \sqrt{3}/2$

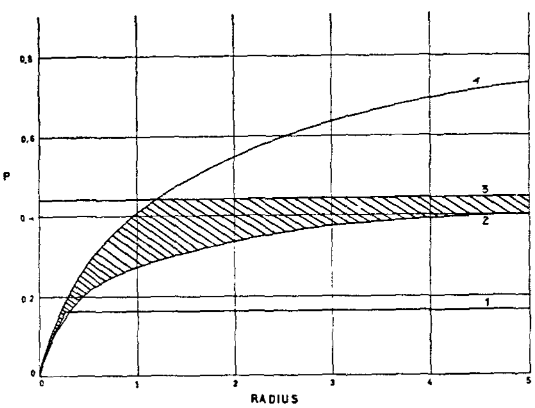

FIG. 6. Limits for P of $\frac{1}{2}$ m.f.p. thick disks of various radii. The limits shown are: 1, P of the inscribed sphere; 2, the lower limit for P_c taken from Fig. 3; P of the volume equivalent sphere; and P of the circumscribing slab. The true value must lie in the cross-hatched area.

we can furthermore write

$$\frac{P_c(b=\frac{1}{2};\epsilon)}{4(1-\epsilon^2)} \geqq P_c(b=\tfrac{1}{2};\sqrt{3}/2) \geqq \bar{P}_c(V') \quad \text{if } \epsilon \geqq \sqrt{3}/2, \tag{86}$$

where V' is the volume of an oblate spheroid with $b=\frac{1}{2}$ and $\epsilon=\sqrt{3}/2$. Here, the first inequality comes from the inclusion theorem (theorem 1), and the second again from the Steiner symmetrization process. Since $\bar{P}_c$ is tabulated,[3] this limit can be realized and is also plotted in Fig. 4.

Included in the diagram are three exact values of P_c corresponding to ratios a/b equal to 5/3, 5/2, and 5 which have been taken from the work of Case *et al.*[3] They indicate that for values of $\epsilon \lesssim 0{,}7$ at least the value of P_c is very close to that of the volume equivalent sphere. The upper limit of theorem 4 under these circumstances (i.e. $\epsilon \lesssim 0{,}7$) is much too high. However, when the eccentricity approaches 1 with the minor axis remaining fixed, the volume increases rapidly and P_c for the volume equivalent sphere rapidly approaches 1. Finally for the most eccentric spheroids the upper limit of theorem 4 becomes applicable. Since when $1-\epsilon \ll 1$, $F(\epsilon)$ is very close to unity in this case the upper and lower limits depend only on b.

6.3. In the estimation of P, somewhat similar indications apply. Shown in Fig. 5 are the following limits for the P of cubes: (i) P of the circumscribing slab (upper limit by theorem 3 or theorem 1); (ii) the cube of P of the circumscribing slab (lower limit by theorem 3); (iii) P of the volume equivalent sphere (upper limit by theorem 2); (iv) B_0^{-1} arctan B_0 (lower limit by theorem 5); (v) the lower limit given by theorems 6 and 14; and (vi) P for the inscribed sphere (lower limit by theorem 1). (i) and (iii) are upper limits of which (iii) is much the lower of the two due to the equilateral nature of the cube. For large cubes (ii) gives the best lower limit; for small ones it is a gross underestimate

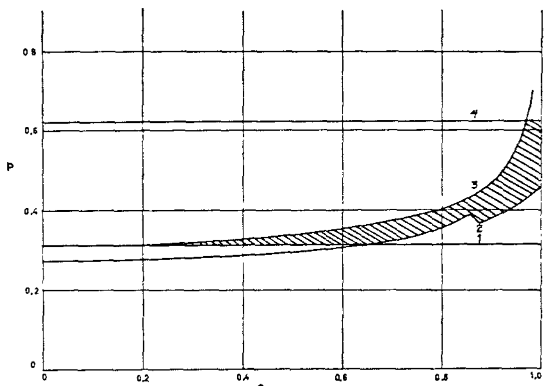

FIG. 7. Limits for P of oblate spheroids whose minor axis is 1 m.f.p. long as a function of eccentricity. The limits shown are: 1, P of the inscribed sphere; 2, the lower limit for P_c taken from Fig. 4; 3, P of the volume equivalent sphere; and 4, P of the circumscribing slab. The true value must lie in the cross-hatched area.

an (vi) is the best lower limit. (ii) and (vi) together cover the range plotted and for the best estimate in this case it is not necessary to use (iv) and (v).

In Figs. 6 and 7 are shown the following limits for P of the disks and spheroids we discussed in the last section: (i) P of the circumscribing slab, (ii) P of the volume equivalent sphere, (iii) P of the inscribed sphere, and (iv) the previously calculated lower limit to P_c. The first two are upper limits, the second two are lower limits.

In Fig. 6, both of the limits (i) and (ii) are used, (i) for the larger radii and (ii) for the smaller exactly as in Fig. 3. The lower limit consists mainly of (iv) except for the smallest cylinders where (iii) was used. In Fig. 7, the situation is quite similar to that of Fig. 4. For small eccentricities, (ii) was used for the upper limit; while for large eccentricities (i), which is the

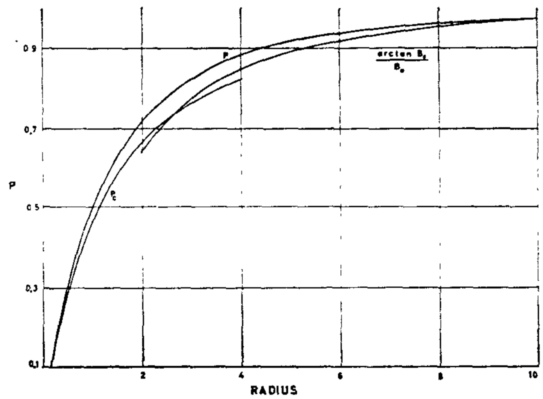

FIG. 8. A comparison of P with P_c and B_0^{-1} arctan B_0 for spheres.

analog of the upper limit in Fig. 4 from theorem 4, was used. For large eccentricities, (iv) was used for a lower limit; while for small eccentricities, (iii) was used. Figure 7 shows clearly the extremely slow variation of P with eccentricity for small eccentricity.

In neither of these last two figures was B_0^{-1} arctan B_0 used as a lower limit for P. In the case of the oblate spheroids, this is because the calculation of B_0 presents difficulties, and indeed the estimation of B_0^2 for spheroids forms the subject of one of the later paragraphs of this paper. In the case of the disks, however, B_0^{-1} arctan B_0 was calculated and found always to be less than limit (iv) above. This is due to the fact that theorem 5 is always unsuitable for a disk of thickness $\frac{1}{2}$ m.f.p., since such small dimensions preclude the use of diffusion theory.

The expectation that P_c should be the closer lower limit to P for small solids and B_0^{-1} arctan B_0 the closer lower limit to P for large solids has already been alluded to in the discussion of paragraph 6.1. Presented in Figs. 8 and 9 are comparisons of these two limits with the exact values of P for spheres and slabs, at once confirming this expectation and showing the rather good accuracy attainable with these variationally derived limits.

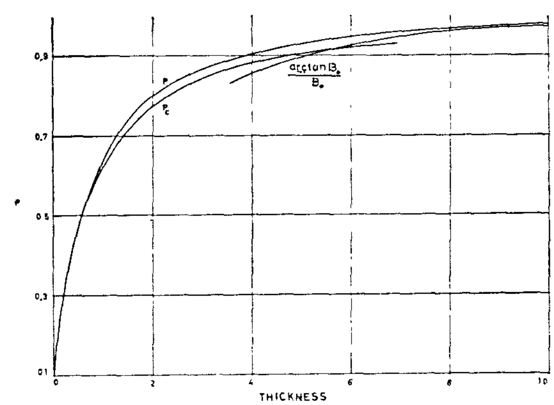

FIG. 9. A comparison of P with P_c and B_0^{-1} arctan B_0 for slabs.

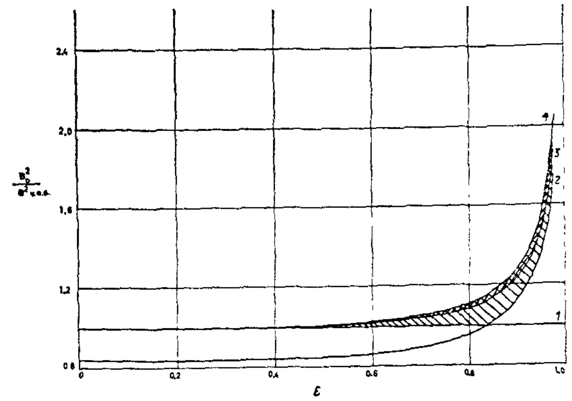

Fig. 10. Limits for the buckling of prolate spheroids as a function of eccentricity. Plotted as ordinate is the ratio of the buckling of the spheroid to that of the volume equivalent sphere. The limits shown are: 1, that derived from the volume equivalent sphere, i.e., unity; 2, that derived from the circumscribed cylinder; 3, that derived from Weinstein's method, i.e., from (25) and (87); and 4, that derived from theorem 10. The true value must lie in the cross-hatched area; if curve 3 is admitted as a lower limit the true value must then lie in the smaller doubly cross-hatched area.

6.4. We shall study but one more example, this one chosen to illustrate the technique of estimating the diffusion-theoretic quantities. We consider estimating the buckling of a prolate spheroid: From theorem 10 and (65) a short calculation shows that

$$\frac{B_0^2}{B_{\text{ves}}^2} \leqq \frac{1-\frac{1}{3}\epsilon^2}{(1-\epsilon^2)^{\frac{1}{3}}}, \tag{87}$$

where B_{ves} is the buckling of the volume equivalent sphere and ϵ is the eccentricity again defined by (83). From theorem 2 it follows that $B_0 \geqq B_{\text{ves}}$, so that the lhs of (87) is always greater than or equal to unity. This last limit we expect to be a good one near $\epsilon=0$, but to become quite useless for highly eccentric spheroids. The only remedy we have for this situation must be found in theorem 1, the inclusion theorem, since no other one can be directly applied to the estimation of B_0. For eccentric prolate spheroids a suitable solid for comparison is the circumscribed cylinder, i.e., that one with the semi-minor axis as radius and the major axis as height. It follows from the properties of this solid and theorem 1 that

$$\frac{B_0^2}{B_{\text{ves}}^2} > \frac{4\alpha^2/\pi^2+(1-\epsilon^2)}{4(1-\epsilon^2)^{\frac{1}{3}}}. \tag{88}$$

The limits (87) and (88) for the ratio B_0/B_{ves} have been plotted in Fig. 10 as functions of the eccentricity. Also plotted is the lower limit unity. Finally, a lower limit based on the method of Weinstein is plotted. This curve was obtained by using (25) with the further assumption that $\lambda_1=4R$ (*correct when* $\epsilon=0$). W, which is given by (74) and (75), can easily be evaluated explicitly and is given by

$$\frac{W}{B_{\text{ves}}^4} = \frac{1-\frac{2}{3}\epsilon^2+\frac{1}{5}\epsilon^4}{(1-\epsilon^2)^{\frac{2}{3}}}. \tag{89}$$

Under these circumstances the estimate given by (25) differs from B_0^2 by a quantity of order ϵ^4. This high accuracy is reflected in the fact that for small ϵ the curve based on (89) and that on theorem 10 (which also differs from B_0^2 by order ϵ^4) nearly coincide. Of course, the limit based on (88) is not a proven lower limit because of the inexact value of λ_1 used to obtain it, and its inclusion in Fig. 10 is to some extent contrary to the spirit of the rest of the paper.

6.5. What remains to be done? Very much indeed the author believes, so that when it is all mentioned the present paper will appear, as it properly should only as a beginning.

In the first place, the notion of seeking inequalities rather than equalities, and the rather exotic techniques (at least for physicists) this notion brings with it, such as Steiner symmetrization or rearrangement of a function in symmetrical decreasing order, have only been very slightly applied to physical problems. It is doubtless true that this enlargement of the conventional point of view will be a very fruitful one.

In the second place, even if we confine ourselves to the framework of neutron transport phenomena, the present paper is little more than a start. For example, our considerations here have been based on the presupposition that the scattering process is isotropic. But surely it is true that Steiner symmetrization decreases the critical multiplication even in a solid in which scattering is anisotropic. And quite probably there is some inequality similar to theorem 5 in media with anisotropic scattering, too.

Not only must a generalization to anisotropic scattering be made, but reflected media must be considered as well. Indeed, a start in this direction has already been made by Ackroyd and Ball,[6] who essentially consider the effect of Steiner symmetrization on critical multiplication for reflected systems.

Finally, even within the restricted milieu of bare, one-velocity reactors with isotropic scattering there are a number of open questions. For example: Does a factorization theorem like theorem 3 hold for P_a or not? Are the multiple collision probabilities from a uniform, isotropic source increased by Steiner symmetrization? (The answer here seems intuitively clear; the basic difficulty is generalizing theorem 380 of Hardy *et al.*[14]) Does a factorization theorem hold for these multiple collision probabilities or not?

ACKNOWLEDGMENTS

The author wishes to thank Dr. F. Stummel and Dr. J. Merkwitz for many illuminating discussions.

APPENDIX

By a positive operator is meant one for which $(f,Hf)\geqq 0$ for all f. If H is both positive and Hermitian then for *any* λ

$$0\leqq (f+\lambda g, H\{f+\lambda g\}) = (f,Hf)+2\ \mathrm{Re}[\lambda(f,Hg)]+|\lambda|^2(g,Hg). \quad (A1)$$

If we now choose $\arg\lambda=-\arg(f,Hg)$, it is easy to verify that

$$2\ \mathrm{Re}[\lambda(f,Hg)]=2|\lambda||(f,Hg)|. \quad (A2)$$

Combining (A1) and (A2) gives the identity in $|\lambda|$:

$$|\lambda|^2(g,Hg)+2|(f,Hg)||\lambda|+(f,Hf)\geqq 0. \quad (A3)$$

For (A3) to hold for all values of the modulus $|\lambda|$, the discriminant must never be positive, i.e.,

$$|(f,Hg)|^2-(f,Hf)(g,Hg)\leqq 0, \quad (A4)$$

which is a generalization of the usual Schwarz inequality.

The first of the operators for which (A4) is to be applied is (15). To prove it is positive we first expand $f(\mathbf{r})$ as

$$f(\mathbf{r})=\sum_{n=0}^{\infty} a_n\phi_n(\mathbf{r})+p(\mathbf{r}), \quad (A5)$$

where $p(\mathbf{r})$ is orthogonal to all the $\phi_n(\mathbf{r})$. Then we use the Hilbert-Schmidt series (5) for $K(|\mathbf{r}-\mathbf{r}'|)$ to obtain

$$(f,Hf)=\int_V |f(\mathbf{r})|^2d^3r-c\sum_{n=0}^{\infty}|a_n|^2/c_n \geqq \int_V |f(\mathbf{r})|^2d^3r-\sum_{n=0}^{\infty}|a_n|^2 = \int_V |p(\mathbf{r})|^2d^3r\geqq 0. \quad (A6)$$

Here use has been made of the fact that c, the fraction of scattering, is by definition less than one, while c_n by virtue of its definition as a critical multiplication, must be larger than one.

The second operator to which (A4) is to be applied is $K(|\mathbf{r}-\mathbf{r}'|)$ itself; from (5) and (A5) is trivially follows that

$$(f,Kf)=\sum_{n=0}^{\infty}|a_n|^2/c_n\geqq 0. \quad (A7)$$

The third operator to which (A4) is to be applied is the operator $-\nabla^2+\kappa^2$ in the volume V, with vanishing boundary condition on the surface S of V. Then by a simple application of Green's theorem,

$$(f,Hf)=\int_V f^*(\mathbf{r})\{-\nabla^2+\kappa^2\}f(\mathbf{r})d^3r = \int_V\{|\nabla f(\mathbf{r})|^2+\kappa^2|f(\mathbf{r})|^2\}d^3r\geqq 0. \quad (A8)$$