Inequality with Applications in Statistical Mechanics*

COLIN J. THOMPSON†
Department of Physics, University of California, San Diego, La Jolla, California
(Received 18 March 1965)

We prove for Hermitian matrices (or more generally for completely continuous self-adjoint linear operators in Hilbert space) A and B that $\mathrm{Tr}\,(e^{A+B}) \leq \mathrm{Tr}\,(e^Ae^B)$. The inequality is shown to be sharper than the convexity property $(0 \leq \alpha \leq 1)$ $\mathrm{Tr}\,(e^{\alpha A+(1-\alpha)B}) \leq [\mathrm{Tr}\,(e^A)]^\alpha[\mathrm{Tr}\,(e^B)]^{1-\alpha}$, and its possible use for obtaining upper bounds for the partition function is discussed briefly.

1. SUMMARY

OUR results are summarized in the following theorems.[1]

Theorem I. For two $n \times n$ Hermitian matrices A and B

$$\mathrm{Tr}\,(e^{A+B}) \leq \mathrm{Tr}\,(e^Ae^B). \tag{1}$$

Theorem II. For two $n \times n$ positive-definite matrices A and B, and $0 \leq \alpha \leq 1$,

$$\mathrm{Tr}\,(A^\alpha B^{1-\alpha}) \leq [\mathrm{Tr}\,(A)]^\alpha[\mathrm{Tr}\,(B)]^{1-\alpha}. \tag{2}$$

Proofs of these theorems (which carry over to completely continuous self adjoint linear operators in Hilbert space) are given in the following two sections.

A consequence of Theorems I and II is[2]:

Corollary: For two $n \times n$ Hermitian matrices A and B, and $0 \leq \alpha \leq 1$,

$$\mathrm{Tr}\,(e^{\alpha A+(1-\alpha)B}) \leq [\mathrm{Tr}\,(e^A)]^\alpha[\mathrm{Tr}\,(e^B)]^{1-\alpha}. \tag{3}$$

The convexity property (3) has been used[3] to obtain an upper bound for the partition function (in the usual notation) $Z = \mathrm{Tr}\,(e^{-\beta H})$ of an antiferromagnetic chain. Equation (1) can also be used to obtain upper bounds for the partition function if we separate the Hamiltonian in a way that enables us to compute the upper bound. In view of (2), the inequality (1) is sharper than (3), so that in general, (1) will probably give us better bounds than (3). Work along these lines is at present in progress.

2. PROOF OF THEOREM I

The proof rests on the following two Lemmas.

Lemma 1. For an $n \times n$ matrix X,

$$|\mathrm{Tr}\,(X)^{2m}| \leq \mathrm{Tr}\,(XX^\dagger)^m, \tag{4}$$

where m is a positive integer and † denotes Hermitian conjugate.

Lemma 2.[1] For two $n \times n$ Hermitian matrices A and B,

$$|\mathrm{Tr}\,(AB)^{2^k}| \leq \mathrm{Tr}\,(A^{2^k}B^{2^k}), \tag{5}$$

where k is a positive integer.

Lemma 1 is a special case of a theorem due to Weyl.[4]

To prove Lemma 2, we first note that with $X = AB$, $X\dagger = BA$ in Lemma 1, we have

$$|\mathrm{Tr}\,(AB)^{2m}| \leq \mathrm{Tr}\,(ABBA)^m = \mathrm{Tr}\,(A^2B^2)^m, \tag{6}$$

where the last equality follows from the cyclic property of the trace. We now proceed by induction. The case $k = 1$ of (5) is just the case $m = 1$ of (6). And if we assume (5) to be true for $k = l$, we have from (6)

$$|\mathrm{Tr}\,(AB)^{2^{l+1}}| = |\mathrm{Tr}\,(AB)^{2(2^l)}| \leq |\mathrm{Tr}\,(A^2B^2)^{2^l}|.$$

The result follows if we then use our inductive assumption with A^2 and B^2 in place of A and B.

The theorem is proved from Lemma 2 by taking $I + 2^{-k}A$ and $I + 2^{-k}B$ in place of A and B, respectively, and proceeding to the limit $k \to \infty$.

We remark that the obvious generalization of (5), namely,

$$|\mathrm{Tr}\,(ABC)^{2^k}| \leq \mathrm{Tr}\,(A^{2^k}B^{2^k}C^{2^k}) \tag{7}$$

is not true, so that Theorem I has no obvious generalization. A counter example to (7) (for $k = 1$) is

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},\ B = \begin{pmatrix} 2 & 0 & -2 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix},\ C = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix}, \tag{8}$$

† Present address: The Rockefeller Institute, New York, N.Y. 10021.

* Supported by the U. S. Air Force Office of Scientific Research, under AF Grant No. AF-AFOSR-610-64, Theory of Solids.

[1] Theorem I, and Lemma 2 in Sec. 2 (for positive-definite matrices only), have recently been proved independently by S. Golden, Phys. Rev. **137**, B1127 (1965).

[2] D. Ruelle, Helv. Phys. Acta **36**, 789 (1963).

[3] R. B. Griffiths, Phys. Rev. **136**, A751 (1964).

[4] H. Weyl, Proc. Natl. Acad. Sci. U. S. **35**, 408 (1949); see also G. Polya, *ibid.* **36**, 49 (1950).

for which $\mathrm{Tr}\,(A^2B^2C^2) = \mathrm{Tr}\,(B^2A^2C^2) = 0$ and $\mathrm{Tr}\,(ABC)^2 = 9$.

3. PROOF OF THEOREM II

We order the eigenvalues a_i of A, and b_i of B in decreasing order, $a_1 \geq a_2 \geq \cdots \geq a_n \geq 0$, $b_1 \geq b_2 \geq \cdots \geq b_n \geq 0$, and use Fan's result[5]

$$\sum_{i=1}^{k} (\varphi_i, B^{1-\alpha}\varphi_i) \leq \sum_{i=1}^{k} b_i^{1-\alpha}, \qquad k = 1, 2, \cdots, n,$$

which holds for an arbitrary orthonormal set of vectors $\{\varphi_i\}$. Choosing the φ_i to be eigenvectors of A and summing by parts gives us

$$\mathrm{Tr}\,(A^\alpha B^{1-\alpha}) = \sum_{i=1}^{n} a_i^\alpha(\varphi_i, B^{1-\alpha}\varphi_i)$$

$$\leq \sum_{i=1}^{n} a_i^\alpha b_i^{1-\alpha}$$

$$\leq \left(\sum_{i=1}^{n} a_i\right)^\alpha \left(\sum_{i=1}^{n} b_i\right)^{1-\alpha}$$

$$= [\mathrm{Tr}\,(A)]^\alpha [\mathrm{Tr}\,(B)]^{1-\alpha},$$

where the last inequality is just Hölder's inequality for positive real numbers.

[5] K. Fan, Proc. Natl. Acad. Sci. U. S. **35**, 652 (1949).

ACKNOWLEDGMENTS

We would like to express our sincere thanks to Professor F. J. Dyson for suggesting a proof of Theorem I, on which ours is based, and for several helpful discussions and comments. The author is also grateful to Professor H. Suhl for the hospitality extended to him at the University of California, San Diego.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 11 NOVEMBER 1965

Oblique Incidence on Plane Boundary between Two General Gyrotropic Plasma Media*

H. Unz†
Antenna Laboratory, The Ohio State University, Columbus, Ohio
(Received 12 March 1965; final manuscript received 22 April 1965)

The problem of a characteristic electromagnetic wave incident obliquely on a plane boundary between two different gyrotropic plasma media is solved. Two characteristic transmitted waves and two characteristic reflected waves will result. The corresponding reflection and transmission coefficients have been evaluated. The particular degenerate cases of free space–gyrotropic medium, gyrotropic medium–free space, and gyrotropic medium–perfect conductor are solved in the appendices.

I. INTRODUCTION

Over the past half-century, many attacks have been made on selected portions of the problem of oblique incidence of electromagnetic waves from free space on a sharply bounded ionosphere. The isotropic case has been discussed in detail by Stratton[1] and by Budden,[2] and their results are in agreement with the ones found originally by Snell (1591–1626), Fresnel (1788–1827), and Brewster (1781–1868). Booker[3–5] treated the obliquely incident wave in the anisotropic ionosphere and derived the well-known "Booker quartic equation" for the refractive index, and his results are given in detail by Budden.[2] Bremmer[6] gave an expression for the reflection coefficients, applicable to the lossless case. Yabroff[7] gave curves showing reflection coefficients as a function of the angle of incidence for various directions of the earth's magnetic field. Additional

* The research reported in this paper was sponsored in part by the Air Force Cambridge Research Laboratories, Office of Aerospace Research, U. S. Air Force, Bedford, Massachusetts, under Contract No. AF 19(604)-7270, at the Antenna Laboratory, The Ohio State University, Columbus, Ohio.

† Consultant to the Antenna Laboratory, Department of Electrical Engineering, The Ohio State University, Columbus, Ohio; Electrical Engineering Department, The University of Kansas, Lawrence, Kansas.

[1] J. A. Stratton, *Electromagnetic Theory* (McGraw-Hill Book Company, Inc., New York, 1941).

[2] K. G. Budden, *Radio Waves in the Ionosphere* (Cambridge University Press, Cambridge, England, 1961).

[3] H. G. Booker, Proc. Roy. Soc. (London) **A155**, 235 (1936).

[4] H. G. Booker, Phil. Trans. Roy. Soc. London **A237**, 411 (1939).

[5] H. G. Booker, J. Geophys. Res. **54**, 243 (1949).

[6] H. Bremmer, *Terrestrial Radio Waves* (Elsevier Publishing Company, New York, 1949).

[7] I. W. Yabroff, Proc. IRE **45**, 750 (1957).