Pacific Journal of Mathematics

IN THIS ISSUE—

Vol. 11, No. 4 , 1961

PACIFIC JOURNAL OF MATHEMATICS

Mathematical papers intended for publication in the *Pacific Journal of Mathematics* should be typewritten (double spaced), and the author should keep a complete copy. Manuscripts may be sent to any one of the four editors. All other communications to the editors should be addressed to the managing editor, L. J. Paige at the University of California, Los Angeles 24, California.

50 reprints per author of each article are furnished free of charge; additional copies may be obtained at cost in multiples of 50.

The *Pacific Journal of Mathematics* is published quarterly, in March, June, September, and December. The price per volume (4 numbers) is $12.00; single issues, $3.50. Back numbers are available. Special price to individual faculty members of supporting institutions and to individual members of the American Mathematical Society: $4.00 per volume; single issues, $1.25.

Subscriptions, orders for back numbers, and changes of address should be sent to Pacific Journal of Mathematics, 103 Highland Boulevard, Berkeley 8, California.

Printed at Kokusai Bunken Insatsusha (International Academic Printing Co., Ltd.), No. 6, 2-chome, Fujimi-cho, Chiyoda-ku, Tokyo, Japan.

PUBLISHED BY PACIFIC JOURNAL OF MATHEMATICS, A NON-PROFIT CORPORATION

The Supporting Institutions listed above contribute to the cost of publication of this Journal, but they are not owners or publishers and have no responsibility for its content or policies.

Reprinted 1966 in the United States of America

PREDICTION THEORY FOR MARKOFF PROCESSES

A. V. BALAKRISHNAN

In this paper we consider the least square prediction problem for Markoff processes with stationary transitions. The main result concerns the partial differential equation characterizing the prediction operator, and the conditions for the uniqueness of the solutions.

Introduction. Let $x(t)$ be a Markoff process with stationary transitions. It is well-known that the optimum mean square predictor of $g(x(s+t))$ given $x(\sigma)$ for $\sigma \leqq s$ is given by the conditional expectation:

$$E[g(x(t+s)) \mid x(\sigma) \leqq s] \ .$$

For a Markoff process this becomes

$$E[g(x(t+s)) \mid x(s)] \tag{1.1}$$

and further, if the transitions are stationary, we need only to consider:

$$E[g(x(t)) \mid x(0)] \tag{1.2}$$

Let $p(t, \xi \mid x)$ be the distribution function (suitably normalized) of the conditional or transition probability of transition from x to ξ in time t. Then, of course, (1.2) becomes

$$\int g(\xi) d_\xi p(t, \xi \mid x) \ . \tag{1.3}$$

Now if $g(.)$ is in $C[\alpha, \beta]$, where $-\infty \leq \alpha < \beta \leq +\infty$ is the interval over which the transition probabilities are defined, we obtain a semigroup of linear operators over $C[\alpha, \beta]$ defined through (1.3). If now we know the infinitesimal generator of this semigroup, we obtain an abstract differential equation for (1.3):

$$\frac{du(t, g)}{dt} = \Delta u(t, g) \tag{1.4}$$

where $u(t, g)$ represents (1.3) and Δ is the infinitesimal generator, provided $g(.)$ is in the domain of Δ. If we know the representation of Δ, and if in particular, it turns out to be a partial differential operator, (1.4) offers an alternate way of determing the prediction functions (1.2) provided uniqueness of the solution can be proved. In what follows, we shall be concerned primarily with situations where such a reduction is possible, and the associated conditions for uniqueness.

Received September 1960.

Main Results:

2. Markoff processes of the diffusion type. A well-known set of sufficient conditions under which the reduction to a parabolic partial differential equation is possible are the Lindberg-Levy conditions which we state here in their weakest form due to Feller [3, 4]. Let

$$\text{(i)}\quad \frac{1}{t}\int_{|\xi-x|>\varepsilon} d_\xi P(t, \xi \mid x) \to 0 \text{ as } t \to 0+$$

$$L_1\quad \text{(ii)}\quad \frac{1}{t}\int_{|\xi-x|<\varepsilon} (\xi - x) d_\xi P(t, \xi \mid x) \to b(x) \text{ as } t \to 0+$$

$$\text{(iii)}\quad \frac{1}{t}\int_{|\xi-x|<\varepsilon} (\xi - x)^2 d_\xi P(t, \xi \mid x) \to 2a(x) \text{ as } t \to 0+$$

Then for each $g(.)$ in $C[\alpha, \beta]$, if we set:

$$T(t)g(x) = \int_\alpha^\beta g(\xi) d_\xi P(t, \xi \mid x)\ . \tag{2.1}$$

$T(t)$ is a semigroup of linear bounded operators over $[\alpha, \beta]$ and moreover

(i) $\| T(t)g \| g \|$ (contraction semigroup)

(ii) $\| T(t)g - g \| \to 0$ as $t \to 0+$ (strongly continuous)

(iii) $T(t)g$ is non-negative if g is nonnegative. (positivity preserving)

(vi) For $g(\xi) =$ constant $T(t)g = g$.

Properties (i), (iii) and (iv) are obvious from (2.1). That $T(t)g$ again belongs to $C[\alpha, \beta]$ follows from condition (i) of L_1, and so does property (ii). Let Λ be the infinitesimal generator of the semigroup. Then the most important property one would like to deduce from L_1 is that it coincides with a second-order differential operator. Unfortunately, however, this is not always entirely true. For example following Feller [4], suppose we define the transition density kernels,

$$P(t, \xi \mid x) = \frac{1}{\sqrt{2\pi t}}\left[\exp - \frac{(g(\xi) - g(x))^2}{2t}\right] g'(\xi),\ \alpha = -\infty,\ \beta = +\infty$$

where say $g(.)$ is a polynomial which vanishes at the origin, and $g'(\xi) > 0$. Taking $g(\xi) = \xi^3$, we obtain for $\xi \neq 0$

$$b(\xi) = -\frac{2}{9}\xi^{-5}$$

$$a(\xi) = \frac{\xi^{-4}}{9}\ .$$

However, at $\xi = 0$,

$$a(0) = 0$$
$$b(0) = 0 \ .$$

Direct substitution into (2.1) shows that for $f(.)$ in the domain of A, $f(0) = f'(0) = f''(0)$ and that

$$Af(0) = \frac{2}{6!} f^6(0) \ .$$

Although for any $\xi \pm 0$,

$$Af(\xi) = a(\xi)\frac{d^2f}{d\xi^2} + b(\xi)\frac{df}{d\xi} \ . \tag{2.2}$$

Here it may be noted that the exceptional point zero is a point of discontinuity of the functions $a(.)$ and $b(.)$. One might then expect that this may avoided if they are required to be continuous. However, it should be noted even in this case that A may not still coincide entirely with the differential operator on the right in (2.2)—in fact, it may only be a contraction of that operator. With some additional conditions on $a(\xi)$ and $b(\xi)$ we can nevertheless obtain a stronger result.

THEOREM 2.1. *Let $a(\xi)$, $b(\xi)$ given by L_1 be continuously twice differentiable in the open interval (α, β) and $a(\xi) > 0$ therein. Let the limits in L_1 hold uniformly in x in each compact sub-interval. Suppose in addition they satisfy*:

$$\int_0^\alpha q(\xi)w(\xi)d\xi = +\infty = \int_0^\beta q(\xi)w(\xi)d\xi \tag{2.3}$$

where $\alpha < 0 < \beta$

$$q(t) = \int_0^t \frac{dt}{a(t)w(t)}$$
$$w(t) = \exp - \int_0^t \frac{b(\xi)}{a(\xi)} d\xi \ .$$

Then the infinitesimal generator A of the semigroup coincides with the differential operator C

$$C = a(\xi)\frac{d^2}{d\xi^2} + b(\xi)\frac{d}{d\xi} \tag{2.4}$$

where the domain of C consists of functions $f(\xi)$ with first and second derivatives such that

$$a(\xi)\frac{d^2f}{d\xi^2} + b(\xi)\frac{df(\xi)}{d\xi}$$

belongs to $C[\alpha, \beta]$.

Conversely suppose the functions $a(\xi)$ and $b(\xi)$ are given, with $a(\xi)$ positive and continuous and $b(\xi)$ continuous in the open interval $[\alpha, \beta]$, and suppose (2.3) is satisfied. Then C generates a semigroup given by (2.1) where the kernels are Markoff transition probabilities which satisfy the conditions L_1, the limits holding uniformly in x in each compact sub-interval.

Proof. Let us consider the converse statement first. Under the conditions (2.3) on the coefficients $a(\xi)$ and $b(\xi)$, Hille [5] has shown that C is the infinitesimal generator of a strongly continuous positive contraction semigroup. Denoting this semigroup by $S(t)$, we have, for any $f(.)$ in $C[\alpha, \beta]$:

$$S(t)f(x) = \int_{\alpha}^{\beta} f(\xi) d_{\xi} P(t, \xi; x) \tag{2.5}$$

where the $P(t, \xi; x)$ are Markoff transition kernels. Moreover, it is readily shown that the kernels satisfy the conditions L_1, with the necessary uniformity.

Suppose next that we are given transition probabilities satisfying L_1 where $a(.)$ and $b(.)$ satisfy (2.3). We know then (2.1) yields a strongly continuous semigroup, and we have to show that its infinitesimal generator A coincides with C. For this, suppose $f(.)$ is in the domain of C^2. Then $f(.)$ has first and second derivatives. Further, suppose $f'(.)$ vanishes outside a compact sub-interval, say $[r_1, r_2]$. Now because $a(\xi) > 0$ and continuous in $[\alpha, \beta]$, it follows that $f''(.)$ is continuous in compact sub-intervals, and hence in particular in $[r_1, r_2]$. Now for each x in $[a, \beta]$

$$\begin{aligned}\frac{S(t)f(x) - f(x)}{t} &= \frac{1}{t}\int_{|\xi - x| > \varepsilon} [f(x) - f(\xi)] dP(t, \xi \mid x) \\ &+ \frac{f'(x)}{t}\int_{|\xi - x| < \varepsilon} (\xi - x) dP(t, \xi \mid x) \\ &+ \frac{f''(x + \theta\varepsilon)}{2t}\int_{|\xi - x| < \varepsilon} (\xi - x)^2 dP(t, \xi\, x)\end{aligned}$$

where $0 < |\theta| < 1$.

In view of L_1, it follows that

$$\operatorname*{limit}_{t \to 0} \frac{T(t)f(x) - f(x)}{t} = a(x)f''(x) + b(x)f'(x) \tag{2.6}$$

and because of the asserted uniformity of the limits in L_1 and the conditions on $f(.)$, it is clear the limit in (2.6) is uniform in x in $[\alpha, \beta]$.

Hence, for such $f(.)$ it follows that

$$\Delta f = Cf .$$

Moreover, for the same $f(.)$, note that Cf again vanishes outside $[r_1, r_2]$. Also, Cf again belongs to the domain of C and hence has first and second derivatives. Hence the argument above can be repeated to yield that

$$\Delta Cf = C^2 f$$

and, of course

$$C\Delta f = C^2 f$$

or

$$C\Delta f = \Delta Cf .$$

Denoting the semigroup generated by C by $S(t)$ if follows readily that

$$S(t)T(t)f = T(t)S(t)f \qquad t > 0$$

and hence using the Dunford argument [See [7]]:

$$S(t)f - T(t)f = \int_0^t (d/d\sigma)S(\sigma)T(\sigma)f = \int_0^t S(\sigma)T(\sigma)(\Delta f - Cf)d\sigma$$

it follows that

$$S(t) = T(t)f .$$

It only remains to show that the set of such functions $f(.)$ is dense in $C[\alpha, \beta]$. Now the class of functions in $C[\alpha, \beta]$ whose derivatives vanish outside compact subsets is dense in the domain of C. Because of the postulated twice differentiability of the coefficients $a(x)$ and $b(x)$, it follows that this class automatically belongs to the domain of C^2, proving the required denseness. It is quite probable the result holds without demanding differentiability of the functions $a(x)$ and $b(x)$.

This proves the theorem.

We note in passing that the conditions (2.3) do not imply uniqueness of solutions of the forward equation, as Hille [3] has shown. From our point of view, this lack of uniqueness is of no concern to us, thus avoiding problems associated with the duality between the backward and forward equations. In particular, Theorem 2.1 establishes that for $f(.)$ in the domain of C,

$$E[f(x(t)) \mid x(0)] = u(t, x)$$

is the unique solution of the Cauchy problem:

$$\frac{\partial u}{\partial t} = a(x)\frac{\partial^2 u}{\partial x^2} + b(x)\frac{\partial u}{\partial x}$$

with

$$u(0, x) = f(x) .$$

As an example, consider the situation of Gaussian white-noise input to a nonlinear system, the input-output processes being related by [See Doob [2, p. 273] for the notation]

$$dx(t) = \sqrt{2}\,(x(t)^2 + 1)d\zeta(t) \tag{2.7}$$

$\zeta(t)$ being the real Gaussian additive process with

$$E[|\, d\zeta(t)\,|^2] = dt$$

so that the output process is Markoffian and we have for the limits in L_1:

$$a(x) = (x^2 + 1)^2$$
$$b(x) = 0 .$$

These clearly satisfy all the required conditions of Theorem 2.1 and the predication function is the solution of the equation:

$$\frac{\partial u}{\partial t} = (x^2 + 1)^2 \frac{d^2 u}{\partial x^2} \tag{2.8}$$

subject to the initial condition

$$u(0, x) = f(x)$$

where it is assumed that

$$(x^2 + 1)f''(x)$$

belongs to $C[-\infty, +\infty]$.
In this particular case, we can obtain the solution in terms of orthogonal functions:

$$u(t, x) = \sum_0^\infty a_n W_n(x) e^{-n(n+1)t} \tag{2.9}$$

where

$$a_n = \int_{-\infty}^{\infty} \frac{W_n(x)}{(x^2 + 1)^2} f(x) dx$$

$$W_n(x) = \sqrt{\frac{2}{\pi}}\,\sqrt{(x^2 + 1)}\, \sin\,(n + 1)\left(\frac{\pi}{2} - \text{arc tan } x\right) .$$

[See Hille [5]] for this solution. The convergence of the series in (2.4) is uniform in x in $[\alpha, \beta]$.

It may be noted that functions such as

$$f(x) = x$$

and

$$f(x) = \exp\, i\lambda x$$

are not in $C[-\infty, +\infty]$, so that we cannot obtain the prediction as the solution of the partial differential equation directly, in the sense in which we have stated that Cauchy problem. It may, however, be possible to consider a slightly different B-space such as the space of functions $f(x)$ continuous in $(-\infty, \infty)$ and such that

$$\operatorname*{limit}_{|x|\to\infty} f(x) \exp - |x|^{\rho}\,, \quad 0 < \rho < 1$$

exist for some ρ, as Hille [6] does for the heat equation.

It should also be noted that in this example, the transition density kernel has the expansion

$$p(t; \xi \mid x) = (\xi^2 + 1)^{-2} \sum_0^\infty W_n(x) W_n(\xi) e^{-n(n+1)t}\,.$$

As $t \to \infty$, we obtain the density

$$p(\infty; \xi \mid x) = 2/\pi(\xi^2 + 1)^{-2} \tag{2.10}$$

and it should be noted that (2.9) for each t is an orthogonal expansion with respect to this density. Also (2.10) corresponds to the (unique) stationary first order distribution with respect to which the process is ergodic. A sufficient condition for the existence of such an expansion (which automatically also yield the corresponding limiting density) due to Hille [5] is that in addition to (2.3) the following

$$\int_0^\alpha q'(x)dx \int_0^x w(\xi)d\xi < +\infty \text{ and } \int_0^\beta q'(x)dx \int_0^x w(x)dx < \infty \tag{2.11}$$

be also satisfied. In this case, the limiting density is simply

$$\frac{q'(x)}{q(\beta) - q(\alpha)}\,.$$

All transition probabilities are absolutely continuous.

3. Markoff processes not of the diffusion type. We shall next consider the prediction problem not of the diffusion type, i.e., whose transition kernels do not satisfy conditions L_1, but rather an extended version of them, leading to elliptic partial differential equations. Thus, let the transition density kernels satisfy:

(i) $2p(t, \xi \mid y) - p(2t, \xi \mid y) \geqq 0$

(ii) $\dfrac{1}{t^2}\displaystyle\int_{|y-\xi|>\varepsilon} [2p(t, \xi \mid y) - p(2t, \xi \mid y)]d\xi \to 0 \quad \text{as } t \to 0+ \; L_2$

(iii) $\dfrac{1}{t^2}\displaystyle\int_{|y-\xi|<\varepsilon} (\xi - y)[2p(t, \xi \mid y) - p(2t, \xi \mid y)]d\xi \to b(y)$

(iv) $\dfrac{1}{t^2}\displaystyle\int_{|y-\xi|<\varepsilon} (y - \xi)^2[2p(t, \xi \mid y) - p(2t, \xi \mid y)]d\xi \to 2a(y)$.

Then the prediction function satisfies the elliptic partial differential equation:

$$\frac{\partial^2 u}{\partial t^2} + a(x)\frac{\partial^2 u}{\partial x^2} + b(x)\frac{\partial u}{\partial x} = 0 \; . \tag{3.1}$$

As before, the main difficulty is in obtaining uniqueness of the solutions.

THEOREM 3.1. *Suppose $a(y)$, $b(y)$ are twice continously differentiable in $[\alpha, \beta]$ and $a(y) > 0$ therein. Suppose further that the limits in L_1 hold uniformly in y in each compact subinterval. Further, suppose that $a(\xi)$, $b(\xi)$ satisfy* (2.3). *Then for each $f(.)$ in the domain of C,*

$$u(t, x) = E[f(x(t)) \mid x(0) = x]$$

satisfies the partial differetial equation

$$\frac{\partial^2 u}{\partial t^2} + Cu(\mathrm{t}, x) = 0 \tag{3.2}$$

and is the only solution of it satisfying to the conditions:

(a) $\| u(t, .) - f(.) \| \to as \; t \to 0$

(b) $\| 2u(t, .) - u(2t) \| \leqq \| f \|$

(c) $\mathrm{Sup}_t \| u(t, .) \| < \infty$.

Conversely, suppose $a(\xi)$, $b(\xi)$ are given such that they are continuous in $[\alpha, \beta]$ and $a(\xi) > 0$ therein, and such that they satisfy (3.1). *Then the Cauchy problem for* (3.1) *has a unique solution satisfying* (a), (b) *and* (c) *for each $f(.)$ in the domain C, the solution being given by*

$$u(t, x) = \int_\alpha^\beta f(\xi)p(t, \xi \mid x)d\xi$$

where the $p(t, \xi \mid x)$ are Markoff transition densities satisfying L_2, the limits existing uniformly in y in compact sub-intervals.

Proof. For a proof of the converse part [see [6]]. Since $a(.)$, $b(.)$

satisfy these conditions in the forward part as well, let us denote the corresponding semigroup by $T(t)$ with generator B. Then we know that

$$B^2 = -C .$$

For each $f(.)$ in $C[\alpha, \beta]$ let us next let

$$u(t, x) = E[f(x(t)) \mid x(0) = x] .$$

Then the conditions L_2 on the transition kernels imply that $u(t, x)$ satisfies (a), (b) and (c), and moreover setting

$$u(t, x) = S(t)f(x)$$

$S(t)$ is a strongly continuous semigroup over $C[\alpha, \beta]$. Let us denote its generator by Λ. We have now to show that

$$\Lambda^2 = B^2 = -C . \tag{3.2}$$

For this, let $f(.)$ belong to $C[\alpha, \beta]$ and let $f'(.)$ vanish outside a compact sub-interval $[r_1, r_2]$. Then as in the proof of Theorem 2.1, we shall first show that $f(.)$ belongs also to the domain of Λ^2 and that

$$\Lambda^2 f = B^2 f .$$

For this we note that

$$\frac{T(2t) + I - 2T(t)}{t^2} f(x)$$
$$= \frac{1}{t^2}\int_{-\infty}^{\infty}(f(\xi) - f(x))[p(2t, \xi \mid x) - 2p(t, \xi \mid x)]d\xi$$

and as before, as $t \to 0$, by virtue of L_2 this goes to

$$-a(x)f''(x) - b(x)f'(x)$$

and the rest of the arguments go over similarly. Also we readily obtain that:

$$\Lambda^2 B^2 f = B^2 \Lambda^2 f .$$

This is enough to imply that

$$T(t)f = S(t)f$$

and the differentiability properties of $a(x)$ and $b(x)$ again imply that such functions $f(.)$ are dense in the domain of C and hence (3.2) follows. This concludes the proof.

The simplest example of a process with transition kernels satisfying the conditions L_2 is the Cauchy additive process, with the independent increments having a Cauchy distribution:

$$E[\exp\ i\,s(\zeta(t+\Delta)-\zeta(t))] = \exp - |\,s\,|\,\Delta\ .$$

More generally, such a process arises as the output of a first-order system whose imput is the Cauchy additive process:

$$dx(t) = b(x(t)dt + a(x(t))d\zeta(t)$$

in the notation of Doob (loc. cit.), $\zeta(t)$ being the input Cauchy additive process ('non-Gaussian white noise'). Now

$$x(t+\Delta) - x(t) = b[x(t)]\Delta + a[x(t)][\zeta(t+\Delta) - \zeta(t)]$$

where the right-side, for given $x(t)$ is specified in terms of $\zeta(t)$ whose statistics are known. The limits required in L_2 are then established by direct calculation. In the case of (i), we may note that we need only prove it for small t, since the semigroup property will then imply it for all values of t. We omit the details of these calculations. The differential equation is:

$$\frac{\partial^2 u}{\partial t^2} + \frac{a(x)^2}{2}\frac{\partial^2 u}{\partial x^2} + b(x)\frac{\partial u}{\partial x} = 0\ .$$

As an example we may consider the case where: $a(.)$ and $b(.)$ are constants:

$$a(x) = \sqrt{2}$$
$$b(x) = -2x\ .$$

The differential equation then is:

$$\frac{\partial^2 u}{\partial t^2} = -2x\frac{\partial u}{\partial x} + \frac{\partial^2 u}{\partial x^2}\ .$$

The (unique) solution of this is the prescribed type for each initial function $f(.)$ can be expanded in Hermite polynomials [See [6] for a general proof]

$$u(t, x) = \sum_0^\infty a_n H_n(x) \exp - \sqrt{2n}\, t \tag{3.3}$$

where the $H_n(.)$ are the Hermite polynomials orthogonal with respect to the Gaussian density:

$$(1/\sqrt{\pi}) \exp - x^2$$

and

$$a_n = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} f(x)H_n(x) \exp - x^2 dx\ .$$

The series in (3.3) converges to the solution function uniformly in compact subsets of $(-\infty, +\infty)$. The transition kernel density $p(t, \xi \mid x)$ is given by

$$p(t, \xi \mid x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{t}{\sqrt{\sigma^2(1 - \exp - 2\sigma)}} \exp\left[\frac{-[\xi - x \exp - 2\sigma]^2}{(1 - \exp - 2\sigma)} - \frac{t^2}{4\sigma}\right] d\sigma$$

as follows again from the theory in [6]. Alternately, it has the expansion:

$$p(t, \xi \mid x) = \sum_0^\infty H_n(x) H_n(\xi) \exp - \xi^2 - \sqrt{2n}\, t\ .$$

As $t \to \infty$, the limiting density is: Gaussian:

$$p(\xi) = \frac{1}{\sqrt{\pi}} \exp - \xi^2$$

with respect to which (as first-order density) the process becomes strictly stationary. It may be shown that the limiting density is again always given by

$$p'(x)/(q(\beta) - q(\alpha))$$

and is thus completely determined by the system, that is by $a(.)$ and $b(.)$ only. The expansion (3.3) is, of course, in terms of functions orthogonal with respect to this density. Thus, taking the example treated in § 2, with

$$a(x) = (x^2 + 1)\sqrt{2}$$
$$b(x) = 0$$

yielding the differential equation:

$$\frac{\partial^2 u}{\partial t^2} + (x^2 + 1)^2 \frac{\partial^2 u}{2x^2} = 0$$

we have the expansion:

$$u(t, x) = \sum_0^\infty a_n W_n(x) \exp - \sqrt{n(n-1)}\, t \tag{3.4}$$

with $W_n(x)$ and a_n as in (2.3). As before, a sufficient condition for the existence of such expansion, is that (2.3) and (2.11) be satisfied. However, this is not necessary as the previous example (3.3) shows.

Extensions. A generalization of the type of process treated in § 3 is got by replacing the kernels in L_2 by

$$\frac{1}{t^n}\left[\sum_1^n \binom{n}{r}(-1)^{r-1} p(rt, \xi \mid x)\right]$$

$\binom{n}{r}$ being the Binomial coefficients, leading to the equations

$$\frac{\partial^n u}{\partial t^n} = (-1)^{n+1}\left[a(x)\frac{\partial^2 u}{\partial x^2} + b(x)\frac{\partial u}{\partial x}\right].$$

However, we have been unable as yet to establish the conditions for uniqueness of the solutions.

We have only so far considered first-order Markoff processes. The extension to higher order processes is similar in principle although it entails partial differential equations in several space variables [see [8] for example], and the results on the related Cauchy problems are still incomplete to a large degree.

REFERENCES

1. A. V. Balakrishnan, *On a class of elliptic equations in the theory of Markoff processes,* Journal of Mathematics and Mechanics, Vol. 10, No. 4, July 1961.

2. J. L. Doob, *Stochastic Processes,* John Wiley and Sons, New York, 1954.

3. W. Feller, *Zur theorie der stochastischen prozesse* (existence und eindeutigkeitssatze), Math. Annalen, **113** (1936), 113-160.

4. _______, *On positivity preserving semigroups of transformations on* $C[r_1, r_2]$, Annales Polonaize, (1952), 85-94.

5. E. Hille, *The abstract Cauchy problem and Cauchy's problem for parabolic equations,* Journal d'Analyse Mathematique, (1953/54), 81-196.

6. _______, *Quelque remarques sur l'equation de la chaleur,* Rend. Mat. e Appe. (5), **15** (1956), 102-118.

7. E. Hille and R. S. Phillips, *Functional analysis and semigroups,* A.M.S. Colloquium Publications, **31** (1958).

8. N. C. Wang and G. E. Uhlenbeck, *On the theory of Browian motion* II, Reviews of Modern Physics, **17** (1945) 323-342.

UPPER BOUNDS FOR THE EIGENVALUES OF SOME VIBRATING SYSTEMS

DALLAS BANKS

1. Introduction. Let $p(x) \geqq 0$, $x \in [0, a]$, be the density of a string fixed at the points $x = 0$ and $x = a$ under unit tension. The natural frequencies of the string are determined by the eigenvalues of the differential system

$$u'' + \lambda p(x)u = 0,\ u(0) = u(a) = 0\ . \tag{1}$$

We note that these eigenvalues depend on the density function $p(x)$ and denote them accordingly by

$$0 < \lambda_1(p) < \lambda_2(p) < \lambda_3(p) < \cdots\ .$$

M. G. Krein [5] has found the sharp bounds

$$\frac{4Hn^2}{M^2}X\left(\frac{M}{aH}\right) \leqq \lambda_n(p) \leqq \frac{\pi^2 n^2 H}{M^2} \qquad (n = 1, 2, \cdots)$$

where $X(t)$ is the least positive root of the equation

$$\sqrt{X}\tan X = \frac{t}{1-t}$$

and where $p(x)$ is such that $\int_0^a p(x)dx = M$ and $0 \leqq p(x) \leqq H$.

Sharp lower bounds are found in [1] when instead of the condition $p(x) \leqq H$, we have $p(x)$ either monotone, $p(x)$ convex, or $p(x)$ concave. The precise definitions of convex and concave are given below.

In this paper, we find sharp upper bounds for $\lambda_n(p)$ $(n = 1, 2, 3, \cdots)$ whenever $p(x)$ belongs to any one of the following sets of functions:

(a) $E_1(M, H, a)$, the set of monotone increasing functions where

$$\int_0^a p(x)dx = M \text{ and } 0 \leqq p(x) \leqq H,\ x \in [0, a]\ .$$

(b) $E_2(M, H, a)$, the set of continuous convex functions, i.e., continuous functions $p(x)$ such that

$$p(x) \leqq \frac{x_2 - x}{x_2 - x_1}p(x_1) + \frac{x - x_1}{x_2 - x_1}p(x_2),\ 0 \leqq x_1 \leqq x_2 \leqq a\ ,$$

with $\int_0^a p(x)dx = M$ and $0 \leqq p(x) \leqq H$, $x \in [0, a]$.

Received October 11, 1960.

(c) $E_3(M, a)$, the set of continuous concave functions, i.e., $-p(x)$ convex, such that $\int_0^a p(x)dx = M$, $x \in [0, a]$.

In general, the values of the maxima appear as the roots of a transcendental system of equations and are not obtained explicitly. However, explicit bounds are given in some special cases.

The methods used generalize to give bounds for the eigenvalues of a vibrating rod. Upper bounds are also found for the lowest eigenvalue of a vibrating membrane over a circular domain when the density is bounded and convex and also when the density is concave.

We make use of the following lemmas.

LEMMA 1. *Let $p(x)$ and $q(x)$ be nonnegative integrable functions defined for $x \in [a, b]$ and let $f(x)$ be nonnegative, continuous and monotone increasing in $[a, b]$. If $c \in (a, b)$ is such that $p(x) \geqq q(x)$ for $x \in (a, c)$ and $p(x) \leqq q(x)$ for $x \in (c, b)$, then*

$$\int_a^b p(x)dx = \int_a^b q(x)dx$$

implies that

$$\int_a^b p(x)f(x)dx \leqq \int_a^b q(x)f(x)dx .$$

If $f(x)$ is monotone decreasing, the inequality sign is reversed.

A proof of this lemma is given in [1].

LEMMA 2. *Let E_k be one of the classes of functions defined above. There exists a function $\rho(x) \in E_k$ such that*

$$\lambda_n(\rho) = \sup_{p(x) \in E_k} \lambda_n(p) .$$

Let $p(x) \in E_k$ for some $k = 1, 2$, or 3. By the definition of E_k, there is a number H such that $0 \leqq p(x) \leqq H$, $x \in [0, a]$. (When $k = 3$, that is when $p(x)$ is concave, we take $H = \frac{2M}{a}$.) It follows that

$$\lambda_n(p) \leqq \frac{n^2\pi^2}{Ha^2} .$$

Hence, there is a number μ such that

$$\mu = \sup_{p(x) \in E_k} \lambda_n(p) .$$

Let $E(M, H, a)$ be the set of all functions $p(x)$, $x \in [0, a]$ such that

$0 \leqq p(x) \leqq H < \infty$ and $\int_0^a p(x)dx = M$. Krein [5] has shown that there exists a subset $\{p_\nu(x)\}$ of $E(M, H, a)$ and a function $\rho(x) \in E(M, H, a)$ such that

$$\lim_{\nu\to\infty}\left(\int_0^x p_\nu(x)dx\right) = \int_0^x \rho(x)dx \ .$$

The convergence is uniform for $x \in [0, a]$ and furthermore

$$\lim_{\nu\to\infty} \lambda_n(p_\nu) = \lambda_n(\rho) \ .$$

In particular if $p(x) \in E_k$, then the functions $p_\nu(x)$ also belong to E_k. We now show that in each of the cases $k = 1, 2, 3$, $\rho(x) \in E_k$ also.

We first consider $E_1(M, H, a)$, that is, the family of all monotone increasing bounded functions $p(x)$ such that $\int_0^a p(x)dx = M$. Then $p_\nu(x) \in E_1(M, H, a)$, $(\nu = 1, 2, \cdots)$. Let

$$\sigma_\nu(x) = \int_0^x p_\nu(x)dx \ .$$

Since $p_\nu(x)$ is increasing, $\sigma_\nu(x)$ must be convex. Hence, $\lim_{\nu\to\infty} \sigma_\nu(x) = \sigma_0(x) = \int_0^a \rho(x)dx$ must also be convex. For if

$$\sigma_\nu(x) \leqq \frac{x - x_1}{x_2 - x_1}\sigma_\nu(x_2) + \frac{x_2 - x}{x_2 - x_1}\sigma_\nu(x_1) \ ,$$

$(x_1 < x < x_2)$, then the same inequality must hold in the limit. It then follows that $\rho(x)$ is increasing.

For the family $E_2(M, H, a)$, that is for convex $p(x)$, we first note that the functions $p_\nu(x)$ $(\nu = 1, 2, \cdots)$ are also convex. We now consider these functions while restricting x to lie in the interval $[\delta, a-\delta]$ where $0 < \delta < a/2$. From the convexity of $p_\nu(x)$, it follows that

$$\left|\frac{p_\nu(x + h) - p_\nu(x)}{h}\right| < H/\delta \ , \qquad (x \in [\delta, a - \delta],\ \nu = 1, 2, \cdots).$$

Hence $\{p_\nu(x)\}$ is an equicontinuous family of functions in this interval. We now consider

$$| p_\nu(x) - \rho(x) | \leqq \left| p_\nu(x) - \frac{\sigma_\nu(x + h) - \sigma_\nu(x)}{h}\right|$$

$$+ \left|\frac{\sigma_\nu(x + h) - \sigma_0(x)}{h} - \frac{\sigma_0(x + h) - \sigma_0(x)}{h}\right| + \left|\frac{\sigma_0(x + h) - \sigma_0(x)}{h} + \rho(x)\right|$$

where x, $x + h \in [\delta, a - \delta]$. Since $\dfrac{\rho_\nu(x + h) - \sigma_\nu(x)}{h} = p_\nu(x + \theta h)$ for some $0 < \theta < 1$, it follows from the equicontinuity that the first term on the right may be made small by choosing h small. The last may be made small by choosing h small since $\sigma_0'(x) = \rho(x)$. Then for fixed h, the middle term may be made small by choosing ν sufficiently large. Thus $p_\nu(x) \to \rho(x)$ as $\nu \to \infty$ in any closed interval properly contained in $(0, a)$. Hence we must have point wise convergence and $\rho(x)$ must be convex, $x \in (0, a)$.

The corresponding result for the family of functions $E_3(M, a)$, that is when $p(x)$ is concave, follows directly from the convex case by considering $\{-p_\nu(x)\}$.

LEMMA 3. *The first variation of* $\lambda_n(p)$ *with the condition* $\int_0^a p(x)dx = M$ *is*

$$\delta\lambda_n(p) = -\lambda_n(p)\int_0^a \delta p(x) u_n^2(x)dx \tag{2}$$

where $u_n(x)$ *is the normalized eigenfunction corresponding to* $\lambda_n(p)$ *and* $\int_0^a (\delta p)dx = 0$.

Consider the differential system associated with a vibrating string of linear density $p(x) + \varepsilon q(x) \geqq 0$, namely

$$(u + \varepsilon v)'' + (\lambda + \varepsilon\mu)(p(x) + \varepsilon q(x))(u + \varepsilon v) = 0 \,,$$
$$u(0) + \varepsilon v(0) = u(a) + \varepsilon v(a) = 0 \,,$$

where $\int_0^a [p(x) + \varepsilon q(x)]dx = M$. We denote the nth eigenvalue of this system by $\lambda_n(p) + \varepsilon\mu_n$ and the corresponding eigenfunction by $u_n(x) + \varepsilon v_n(x)$ where $u_n(x)$ is the eigenfunction corresponding to $\lambda_n(p)$. $u_n + \varepsilon v_n(x)$ then satisfies the equation

$$u_n'' + v_n'' + (\lambda_n(p) + \varepsilon\mu_n)(p(x) + \varepsilon q(x))(u_n + \varepsilon v_n) = 0 \,.$$

Multiplying this by $u_n(x)$ and integrating the resulting expression over the interval $(0, a)$, we get

$$-\lambda_n(p) + \varepsilon\int_0^a u_n v_n'' dx + (\lambda_n(p) + \varepsilon\mu_n)[1 + \varepsilon\int_0^a (p u_n v_n + q u_n^2)dx + 0(\varepsilon^2)] = 0 \,.$$

We have used the relation $\int_0^a u_n'' u_n dx = -\lambda_n(p)$ and taken $\int_0^a p u_n^2 dx = 1$. Solving for μ_n, we find

$$\mu_n = \frac{-\lambda_n(p)\int_0^a q(x)u_n^2(x)dx - \lambda_n\int_0^a (v_n'' u_n - v_n u_n'')dx + 0(\varepsilon)}{1 + 0(\varepsilon)} \,.$$

Integrating the second integral by parts, we find that it vanishes so that letting $\varepsilon \to 0$, we get

$$\mu_n = -\lambda_n(p)\int_0^a q(x)u_n^2(x)dx .$$

Hence

$$\delta\lambda_n(p)\,\varepsilon\mu_n = -\lambda_n(p)\int_0^a \delta p(x)u_n^2(x)dx$$

where we have taken $\delta p(x) = \varepsilon q(x)$. Since $\int_0^a [p(x) + \varepsilon q(x)]dx = M$ and $\int_0^a p(x)dx = M$, it necessarily follows that $\int_0^a \delta p(x)dx = 0$.

2. Monotone density functions. We first consider the case where $p(x)$ is a monotone increasing function such that $0 \leqq p(x) \leqq H < \infty$, that is when $p(x)\varepsilon E_1(M, H, a)$.

THEOREM 1. *Let $\lambda_n(p)$ be the nth eigenvalue of a vibrating string with fixed boundary values and with a monotone increasing density function $p(x)\varepsilon E_1(M, H, a)$. Then*

$$\lambda_n(p) \leqq \lambda_n(\rho)$$

where $\rho(x)\varepsilon E_1(M, H, a)$ is a step function with at least one and at most n discontinuities in the open interval $(0, a)$.

By Lemma 2 there exists a monotone bounded function $\rho(x) \in E_1(M, H, a)$ such that $\lambda_n(\rho) = \max_{p\in E_1} \lambda_n(p)$. Hence, letting $p(x) = \rho(x)$ in the variational formula (2), we have $\delta\lambda_n(\rho) \leqq 0$. We now show that unless $p(x) \in E_1(M, H, a)$ is a step function with at most n discontinuities $\delta\lambda_n(p) > 0$ for some $\delta p = \varepsilon q$ where $p(x) + \delta p(x) \in E_1(M, H, a)$. Hence, $\rho(x)$ must be a step function with at most n discontinuities.[1]

Let $u_n(x)$ be the eigenfunction corresponding to $\lambda_n(p)$. Denote the nodal points of $u_n(x)$ by x_k $(k = 0, 1, \cdots, n)$ where $x_0 = 0$ and $x_n = a$. Since $u_n(x)$ has only one extremum point in each of the intervals $(x_{k-1}, x_k)(k = 1, 2, \cdots, n)u_n^2(x)$ has only one maximum there. Let that point in (x_k, x_{k+1}) be $\bar{x}_k$ $(k = 1, 2, \cdots, n)$. For $k = 1, 2, \cdots, n$, we let

$$r(x) = a_k = \int_{x_{k-1}}^{\bar{x}_k} p(x)dx/(\bar{x}_k - x_{k-1}),\ x \in [x_{k-1}, \bar{x}_k] .$$

Since a_k is the mean value of $p(x)$ in $(x_{k-1}, \bar{x}_k)$ and $p(x)$ is monotone increasing, it follows that $a_{k+1} \geqq p(x)$ if $x \in [\bar{x}_k, x_k]$ $(k = 1, 2, \cdots, n-1)$ and that $a_k \leqq p(x)$ if $x \in [\bar{x}_k, x_k]$ $(k = 1, 2, \cdots, n)$. Hence, it is possible

[1] The author is indebted Z. Nehari for suggesting the variational approach used in this paper.

to find a point $\xi_k \in (\bar{x}_k, x_k)$ such that

$$r(x) = \begin{cases} a_k & \text{if } x \in [\bar{x}_k, \xi_k) \\ a_{k+1} & \text{if } x \in [\xi_k, x_k] \end{cases}, \qquad (k = 1, 2, \cdots, n)$$

satisfies the relation

$$\int_{\bar{x}_k}^{x_k} r(x)dx = \int_{\bar{x}_k}^{x_k} p(x)dx$$

$(k = 1, 2, \cdots, n)$. We have taken $a_{n+1} = H$, the upper bound of $p(x)$. In each of the intervals $(x_{k-1}, \bar{x}_k)$ and $(\bar{x}_k, x_k)$ $(k = 1, 2, \cdots, n)$, $r(x)$ and $p(x)$ satisfy the hypothesis of Lemma 1.1 relative to $u_n^2(x)$. Hence, we have

$$\int_{x_{k-1}}^{\bar{x}_k} p(x)u_n^2(x)dx \geqq \int_{x_{k-1}}^{\bar{x}_k} r(x)u_n^2(x)dx$$

and

$$\int_{\bar{x}_k}^{x_k} p(x)u_n^2(x)dx \geqq \int_{\bar{x}_k}^{x_k} r(x)u_n^2(x)dx$$

$(k = 1, 2, \cdots, n)$. Summing on k, we find that

$$\int_0^a [p(x) - r(x)]u_n^2(x)dx \geqq 0 .$$

The equality sign will hold if and only if $p(x) = r(x)$, i.e., $p(x)$ is constant or is a step function with precisely one jump in each of the intervals (x_{k-1}, x_k) $(k = 1, 2, \cdots, n)$. If we let $q(x) = r(x) - p(x)$, then for small $\varepsilon > 0$ Lemma 3 gives the result

$$\begin{aligned} \delta\lambda_n(p) &= -\lambda_n(p) \int_0^a \in q(x)u_n^2(x)dx \\ &= -\lambda_n(p)\int_0^a \delta p(x)u_n^2(x)dx > 0 \end{aligned}$$

unless $p(x) = r(x)$. Hence, $\rho(x) = r(x)$ if $\lambda_n(\rho)$ is a maximum. But $r(x)$ is a step function with at most n jumps in $(0, a)$.

Finally, we show that the maximizing density cannot be a constant so that there must be at least one jump. We first consider the lowest eigenvalue. We show that $\delta\lambda_1(p) > 0$ when $p(x) = M/a$ for a particular $\delta p = \varepsilon q$.

The eigenfunction corresponding to $\lambda_1(M/a)$ is

$$u_1(x) = \sqrt{2/a} \sin \frac{\pi x}{a} .$$

If we let

$$\delta p(x) = \varepsilon q(x) = \begin{cases} -\varepsilon & \text{if } x \in (0, a/2 + \eta) \, , \\ \varepsilon \dfrac{a/2 + \eta}{a/2 - \eta} & \text{if } x \in (a/2 + \eta, a) \, , \end{cases}$$

where $0 < \eta < a/2$ then $\int_0^a \delta p(x) dx = 0$ and

$$\delta\lambda_1(M/a) = -\lambda_1(M/a) \int_0^a \delta p(x) u_1^2(x) dx \, .$$

From the symmetry of $u_1(x)$ about the point $x = a/2$ and Lemma 1 it is easily seen that

$$\int_0^a \delta p(x) u_1^2(x) dx > 0 \, .$$

Hence, $\delta\lambda_1(M/a) > 0$ so that $\lambda_1(M/a)$ cannot be a maximum value of $\lambda_1(p)$.

The corresponding result for the higher eigenvalues can be obtained by choosing

$$\delta p(x) = \varepsilon q(x) = \begin{cases} -\varepsilon & \text{if } x \in \left(0, \dfrac{a}{2n} + \eta\right) \, , \\ \dfrac{\varepsilon(a/2n + \eta)}{\dfrac{(2n-1)a}{2n} - \eta} & \text{if } x \in (a/2n + \eta, a) \, , \end{cases}$$

where $0 < \eta < a/2n$. It then follows from the periodicity of

$$u_n(x) = \sqrt{2/a} \sin \frac{n \pi x}{a}$$

and the argument used for $\lambda_1(M/a)$ that $\lambda_n(M/a)$ cannot be a maximum value of $\lambda_n(p)$, $p \in E_1(M, H, a)$.

The upper bound of $\lambda_1(p)$, $p \in E_1(M, H, a)$ is thus given as the maximum of the lowest eigenvalue of the system.

$$u'' + \lambda p_\theta(x) u = 0, \; u(0) = u(a) = 0 \tag{3}$$

where

$$p_\theta(x) = \begin{cases} \theta H & \text{if } x \in [0, \xi a) \, , \\ H & \text{if } x \in [\xi a, a] \, ; \end{cases}$$

$0 < \theta < 1$ and $\xi = \dfrac{1 - M/Ha}{1 - \theta}$. That $\theta = 0$ may be excluded from consideration follows easily from the derivation of the form of $\rho(x)$ and the fact that the maximum of $u_1(x)$ in this case must occur in the open interval $(\xi a, a)$. For we would have $a_1 = \int_0^{\bar{x}_1} p_\theta(x) dx \neq 0$.

The eigenfunctions of (3), are [2]

$$u_n(x) = \begin{cases} \sin \sqrt{\lambda_n H}\,(1-\xi)a \cdot \sin \sqrt{\lambda_n \theta H}\, x\,, & x \in [0, \xi a)\,, \\ \sin \sqrt{\lambda_n \theta H}\, \xi a \cdot \sin \sqrt{\lambda_n H}\,(a - x)\,, & x \in [\xi a, a] \end{cases}$$

where $\lambda_n(p_\theta)$ is the nth positive root of

$$\tan(\xi a \sqrt{\lambda \theta H}) + \sqrt{\theta} \tan a(1-\xi)\sqrt{\lambda H} = 0\,.$$

We could now compute $\dfrac{d\lambda_1(p_\theta)}{d\theta}$ and determine the value which maximizes $\lambda_1(p_\theta)$.

The determination of the bounds for the higher eigenvalues is also seen to be a problem in ordinary calculus since the jumps of the step function which give the maximum must occur in the open interval $(0, a)$.

3. Convex density functions. Let $p(x)$, $x \in [0, a]$ be a continuous convex function such that $\int_0^a p(x)dx = M$ and $0 \leqq p(x) \leqq H$, that is, let $p(x) \in E_2(M, H, a)$.

THEOREM 2. *Let $\lambda_1(p)$ be the lowest eigenvalue of a string with fixed end points and with density $p(x) \in E_2(M, H, a)$. Then*

$$aM\lambda_1(p) \leqq \mu\left(\frac{aH}{M}\right)$$

where $\mu(h) = [6(h-1)t_1]^2/h^3$ and t_1 is the least positive root of

$$J_{1/3}(t)J_{2/3}\left(\frac{(2-h)^3 t}{h^3}\right) - J_{-1/3}(t)J_{-2/3}\left(\frac{(2-h)^3 t}{h^3}\right) = 0$$

if $1 < h < 2$ and $\mu(h) = h(3t_1/2)^2$ and t_1 is the least positive root of $J_{-2/3}(t) = 0$ if $h \geqq 2$. The minimum is uniquely attained for the function

$$\rho(x) = \begin{cases} \dfrac{4}{a^2}(M - aH)x + H\,, & x \in (0, a/2)\,, \\ \rho(a - x)\,, & x \in (a/2, a)\,, \end{cases} \tag{5}$$

if $1 < h = \dfrac{aH}{M} < 2$ and

$$\rho(x) = \begin{cases} [H/M](M - Hx)\,, & x \in (0, M/H)\,, \\ 0\,, & x \in (M/H, a/2)\,, \\ \rho(a - x)\,, & x \in (a/2, a)\,, \end{cases} \tag{6}$$

if $h = \dfrac{aH}{M} \geqq 2$.

It is well known that $\lambda_1(p)$ is the minimum of

$$J(u) = \frac{\int_0^a u'^2(x)dx}{\int_0^a p(x)u^2(x)dx}$$

where the minimum is taken over all functions $u \in C'$ which vanish at $x = 0$ and $x = a$. If we let

$$\bar{p}(x) = \frac{1}{2}[p(x) + p(a - x)]$$

then

$$\begin{aligned}
\lambda_1^{-1}(\bar{p}) &= \max_{u \in C'} \frac{\int_0^a \bar{p}(x)u^2(x)dx}{\int_0^a u'^2(x)dx} \\
&\leqq \max_{u \in C'} \frac{\int_0^a p(x)u^2(x)dx}{2\int_0^a u'^2(x)dx} + \max_{u \in C'} \frac{\int_0^a p(a - x)u^2(x)dx}{2\int_0^a u'^2(x)dx} \\
&= \lambda_1^{-1}(p)
\end{aligned}$$

since the eigenvalues of a string with density $p(a - x)$ are the same as those of a string with density $p(x)$. Hence any upper bound of $\lambda_1(\bar{p})$ is also an upper bound of $\lambda_1(p)$.

The differential system (1) with $p(x)$ replaced by $\bar{p}(x)$ has the same lowest eigenvalue as the system

$$u'' + \lambda\bar{p}(x)u = 0,\ u(0) = u'(a/2) = 0,\ x \in [0, a/2]. \tag{7}$$

Furthermore, since $p(x)$ is convex, so is $\bar{p}(x)$, $x \in [0, a]$, and the bound H is also a bound of $\bar{p}(x)$.

We now compare the lowest eigenvalue of the system (7) with that of the same system when $\bar{p}(x)$ is replaced by

$$\rho_1(x) = [4/a^2](M - aH)x + H,\ x \in [0, a/2]\ ,$$

if $1 < \dfrac{aH}{M} < 2$ and

$$\rho_1(x) = \begin{cases} \dfrac{H}{M}(M - Hx)\ , & x \in [0, M/H]\ , \\ 0\ , & x \in [M/H, a/2]\ , \end{cases}$$

if $\frac{aH}{M} \geqq 2$. In either case, since $\rho_1(0) = H \geqq \bar{p}(0)$ and $\int_0^{a/2} \rho_1(x)dx = \int_0^{a/2} \bar{p}(x)dx$, it follows from the convexity of $\bar{p}(x)$ that there is a point $\xi \in (0, a/2)$ such that $\rho_1(x) \geqq \bar{p}(x)$ if $x \in (0, \xi)$ and $\rho_1(x) \leqq \bar{p}(x)$ if $x \in (\xi, a/2)$. There will be strict inequality in each of these open intervals unless $\bar{p}(x) = \rho_1(x)$, $x \in [0, a/2]$. If $u(x)$ is monotone increasing in $[0, a/2]$ with $u(0) = u'(a/2) = 0$, we have by Lemma 1

$$\int_0^{a/2} \rho_1(x)u^2(x)dx \leqq \int_0^{a/2} \bar{p}(x)u^2(x)dx \; . \tag{8}$$

Since the first eigenfunction of the system (7) is a monotone increasing function, it follows from the comparison theorem [2] that

$$\lambda_1(\bar{p}) \leqq \lambda_1(\rho_1) \; .$$

There will be equality if and only if $\bar{p}(x) = \rho_1(x)$, for if $u(x)$ is the eigenfunction corresponding to the lowest eigenvalue of (7) with $\bar{p}(x)$ replaced by $\rho_1(x) \neq \bar{p}(x)$ then (8) will be a strict inequality and hence

$$\lambda_1(\rho_1) = \frac{\int_0^{a/2} u'^2(x)dx}{\int_0^{a/2} \rho_1(x)u^2(x)dx} > \frac{\int_0^{a/2} u'^2(x)dx}{\int_0^{a/2} \bar{p}(x)u^2(x)dx} \geqq \lambda_1(p) \; .$$

But $\lambda_1(\rho_1)$ is also the lowest eigenvalue of the system (1) with $p(x)$ replaced by

$$\rho(x) = \begin{cases} \rho_1(x) \; , & x \in [0, a/2] \; , \\ \rho_1(a - x) \; , & x \in [a/2, a] \; . \end{cases}$$

This is just the function (5) if $1 < \frac{aH}{M} < 2$ and the function (6) if $\frac{aH}{M} \geqq 2$. Hence we see that $\lambda_1(\rho) \geqq \lambda_1(p)$ for any bounded convex $p(x)$.

When $\rho(x)$ is defined by (5) we find that

$$\lambda_1(\rho) = \frac{\mu\left(\frac{aH}{M}\right)}{a\int_0^a p(x)dx}$$

where $\mu(h) = [6(h-1)t_1]^2/h^3$ and t_1 is the least positive root of

$$J_{1/3}(t)J_{2/3}(kt) - J_{-1/3}(t)J_{-2/3}(kt) = 0 \; ,$$

$k = \frac{(2-h)^3}{h^3}$ [4]. When $\rho(x)$ is defined by (6) we have

$$\lambda_1(\rho) = \frac{\mu\left(\frac{aH}{M}\right)}{a\int_0^a p(x)dx}$$

where $\mu(h) = h(3t_1/2)^2$ and t_1 is the least positive root of $J_{-2/3}(t) = 0$ [4].

A better bound is obtained if, instead of the bound H, we use $\bar{H} = \frac{1}{2}[p(0) + p(a)]$ for the bound of $\bar{p}(x)$. This results in a smaller value of $\mu(aH/M)$ whenever $p(0) \neq p(a)$.

For the larger eigenvalues we prove the following.

THEOREM 3. *Let $\lambda_n(p)$ be the nth eigenvalue of a vibrating string with fixed boundary values and with a convex density $p(x) \in E_2(M, H, a)$. Then*

$$\lambda_n(p) \leqq \lambda_n(\rho)$$

where $\rho(x) \in E_2(M, H, a)$ is a piecewise linear convex function with at most $(n + 2)$ pieces.

The existence of a bounded convex function $\rho(x)$ such that $\max_{p \in E_2} \lambda_n(p) = \lambda_n(\rho)$ follows from Lemma 2. It then follows by Lemma 3 that

$$\delta\lambda_n(\rho) = -\lambda_n(\rho)\int_0^a \delta\rho(x)u_n^2(x)dx \leqq 0 .$$

We now show that either $p(x)$ is a convex piecewise linear function with at most $(n + 2)$ pieces or there exists a function $q(x)$ such that $\delta\lambda_n(p) > 0$ when $\delta p = \varepsilon q$ where $p(x) + \delta p(x) \in E_2(M, H, a)$. Let $u_n(x)$ be the eigenfunction corresponding to $\lambda_n(p)$. We first find a convex function $r(x)$ such that

$$\int_0^a r(x)u_n^2(x)dx \leqq \int_0^a p(x)u_n^2(x)dx .$$

Instead of trying to find $r(x)$ directly, we carry out a preliminary construction. As in Theorem 1, we denote the minimum points of $u_n^2(x)$ by x_k $(k = 0, 1, \cdots, n)$ and the maximum points by $\bar{x}_k$ $(k = 1, 2, \cdots, n)$. We first consider each of the intervals $(\bar{x}_k, \bar{x}_{k+1})$ $(k = 1, 2, \cdots, n - 1)$ separately.

Let $L(x)$ be any linear function such that $L(x) \leqq p(x)$, $x \in (x_k, \bar{x}_{k+1})$ for some fixed integer k $(1 \leqq k \leqq n - 1)$. Then $m(x) = \max\ \{L(x), 0\}$ satisfies the inequality $0 \leqq m(x) \leqq p(x)$. Now let c_k be any number such that $c_k \geqq p(x_k)$. Then there is a number a_k such that

$$\int_{x_k}^{\bar{x}_{k+1}} [a_k(x - x_k) + c_k]dx = \int_{x_k}^{\bar{x}_{k+1}} p(x)dx . \tag{9}$$

If $a_k(x - x_k) + c_k \geqq m(x)$, $x \in (x_k, \bar{x}_{k+1})$, then we let

$$g_k(x, c_k) = a_k(x - x_k) + c_k\ , \quad x \in (x_k, \bar{x}_{k+1})\ .$$

If $a_k(x - x_k) + c_k < m(x)$ for some $x \in (x_k, \bar{x}_{k+1})$, then we redefine a_k by the condition

$$\int_{x_k}^{\xi_k}[a_k(x - x_k) + c_k]dx + \int_{\xi_k}^{\bar{x}_{k+1}} m(x)dx = \int_{x_k}^{\bar{x}_{k+1}} p(x)dx \tag{10}$$

where ξ_k satisfies the equation $a_k(\xi_k - x_k) + c_k = m(\xi_k)$. In this case, we define $g_k(x, c_k)$ by

$$g_k(x, c_k) = \begin{cases} a_k(x - x_k) + c_k\ , & x \in (x_k, \xi_k)\ , \\ m(x)\ , & x \in [\xi_k, \bar{x}_{k+1})\ . \end{cases}$$

Now consider the interval $(\bar{x}_k, x_k)$. Let $m(x) = \max\{L(x), 0\}$ where $L(x)$ is any linear function such that $L(x) \leqq p(x)$ if $x \in (\bar{x}_k, x_k)$. There is a number b_k such that

$$\int_{\bar{x}_k}^{x_k}[b_k(x - x_k) + c_k]dx = \int_{\bar{x}_k}^{x_k} p(x)dx\ . \tag{11}$$

If $b_k(x - x_k) + c_k \geqq m(x)$ for $x \in (\bar{x}_k, x_k)$, we let

$$h_k(x, c_k) = b_k(x - x_k) + c_k\ , \quad x \in (\bar{x}_k, x_k)\ .$$

If $b_k(x - x_k) + c_k < m(x)$ for some $x \in (\bar{x}_k, x_k)$, we redefine b_k by the condition

$$\int_{\bar{x}_k}^{\eta_k} m(x)dx + \int_{\eta_k}^{x_k}[b_k(x - x_k) + c_k]dx = \int_{\bar{x}_k}^{x_k} p(x)dx \tag{12}$$

where η_k satisfies the equation $b_k(\eta_k - x_k) + c_k = m(\eta_k)$. We then define $h_k(x, c_k)$ by

$$h_k(x, c_k) = \begin{cases} m(x)\ , & x \in (\bar{x}_k, \eta_k)\ , \\ b_k(x - x_k) + c_k\ , & x \in (\eta_k, x_k)\ . \end{cases}$$

We may consider a_k and b_k to be functions of c_k. They are continuous functions as is easily seen from the defining relations of a_k and b_k. It follows that there is a number $\gamma_k \geqq p(x_k)$ such that $a_k = b_k$ if $c_k = \gamma_k$. For if $c_k = p(x_k)$, the convexity of $p(x)$ implies that $a_k - b_k \geqq 0$. On the other hand, if c_k is sufficiently large, $a_k - b_k < 0$. Hence, by the continuity, the value γ_k exists such that $c_k = \gamma_k$ implies $a_k = b_k$.

In the interval $[x_0, \bar{x}_1]$, we define $g_0(x, c_0)$, in the same way that $g_k(x, c_k)$ was defined except that we specify $c_0 = p(0) = \gamma_0$. Similarly in $[\bar{x}_n, a]$ we define $h_n(x, c_n)$ as above except that we take $c_n = p(a) = \gamma_n$.

We now let

$$r_1(x) = g_0(x, \gamma_0)\ , \quad x \in [0, \bar{x}_1]\ ,$$

$$r_k(x) = \begin{cases} h_k(x, \gamma_k) , & x \in [\bar{x}_k, x_k] , \\ g_k(x, \gamma_k) , & x \in [x_k, \bar{x}_{k+1}] , \end{cases}$$
$$r_n(x) = h_n(x, \gamma_n) , \quad x \in [\bar{x}_n, x_n] ,$$

$(k = 1, 2, \cdots, n-1)$. From (9) or (10), which ever applies, we have

$$\int_{x_{k-1}}^{\bar{x}_k} r_k(x)dx = \int_{x_{k-1}}^{\bar{x}_k} p(x)dx .$$

The convexity of $p(x)$ and the definition of $r_k(x)$ imply by Lemma 1 that

$$\int_{x_{k-1}}^{\bar{x}_k} r_k(x)u_n^2(x)dx \leqq \int_{x_{k-1}}^{\bar{x}_k} p(x)u_n^2(x)dx . \tag{13}$$

Similarly from (11) or (12) we have

$$\int_{\bar{x}_k}^{x_k} r_k(x)u_n^2(x)dx \leqq \int_{\bar{x}_k}^{x_k} p(x)u_n^2(x)dx . \tag{14}$$

Furthermore, we have strict inequality unless $r_k(x) = p(x)$ in each case.

We are now able to define the function $r(x)$ by induction. We carry out the process only for $n = 3$ to avoid unnecessary detail. In $(x_0, \bar{x}_1)$, we let $m(x) = 0$ and define $r_1(x)$ as above. In $(\bar{x}_1, \bar{x}_2)$, we also define $r_2(x)$ with $m(x) = 0$. Then, comparing $r_1(\bar{x}_1)$ and $r_2(\bar{x}_1)$ we have the following alternatives:

(i) If $r_1(\bar{x}_1) > r_2(\bar{x}_1)$, we define a new function $r_2(x)$ with $m(x) = \max\{r_1(x), 0\}$, $x \in [\bar{x}_1, \bar{x}_2]$ where we define $r_1(x)$ in this interval by extrapolation.

(ii) If $r_1(\bar{x}_1) < r_2(\bar{x}_1)$, we define a new function $r_1(x)$ with $m(x) = \max\{r_2(x), 0\}$, $x \in [x_0, \bar{x}_1]$, where $r_2(x)$ is defined in this interval by extrapolation.

(iii) If $r_1(\bar{x}_1) = r_2(\bar{x}_1)$ we leave $r_1(x)$ and $r_2(x)$ as they are.

Using whichever alternative applies, we define

$$r^{(1)}(x) = \begin{cases} r_1(x) , & x \in [x_0, \bar{x}_1] , \\ r_2(x) , & x \in [\bar{x}_1, \bar{x}_2] . \end{cases}$$

Now, define $r_3(x)$, $x \in [\bar{x}_2, \bar{x}_3]$ with $m(x) = 0$ and compare $r^{(1)}(\bar{x}_2)$ and $r_3(\bar{x}_2)$. We use the same alternatives as above, the only difference being that if $r^{(1)}(\bar{x}_2) < r_3(\bar{x}_2)$ we must redefine $r^{(1)}(x)$ with $m(x) = \max\{r_3(x), 0\}$, $x \in [x_0, \bar{x}_2]$ where as above we define $r_3(x)$ by extrapolation.

It is clear that the above process can be completed for any integer n. The function which we obtain by this method we call $r(x)$. It will be a convex function since any two adjacent segments of the graph of $r(x)$ can only have a point of intersection which lies on or below the graph of $p(x)$. Since there is possibly a subinterval of $[0, a]$ where $r(x)$ may be zero, $r(x)$ may have up to $n + 2$ linear pieces.

If we sum the inequalities (13) and (14) we find that

$$\int_0^a r(x)u^2(x)dx \leqq \int_0^a p(x)u_n^2(x)dx$$

with strict inequality unless $r(x) = p(x)$ in $(0, a)$. Choosing $\delta p = \varepsilon q(x) = \varepsilon[r(x) - p(x)]$, we have that $p(x) + \delta p(x)$ is convex if $\varepsilon > 0$ is small and hence

$$\int_0^a \delta p u_n^2(x)dx < 0$$

or $\delta\lambda_n(p) > 0$ unless $p(x) \equiv r(x)$, $x \in [0, a]$. Since we must have $\delta\lambda_n(\rho) \leqq 0$, it follows that $\rho(x)$ is the same type of function as $r(x)$. From the method of determining $r(x)$, we see the $\rho(x)$ is a convex piecewise linear function with at most $n + 2$ linear segments.

We note from Theorem 2 that this is precisely the case when $n = 1$.

4. Concave density functions.

We consider the case when $p(x)$, $x \in [0, a]$ is a continuous concave function such that $\int_0^a p(x)dx = M$, that is, when $p(x) \in E_3(M, a)$.

THEOREM 4. *Let $\lambda_n(p)$ be the nth eigenvalue of a string with fixed end points and with a concave density function $p(x) \in E_3(M, a)$. Then*

$$\lambda_n(p) \leqq \lambda_n(\rho)$$

where $\rho(x) \in E_3(M, a)$ and is a piecewise linear concave function with at most n pieces.

The existence of a concave function $\rho(x)$ such that

$$\max_{p \in E_3} \lambda_n(p) = \lambda_n(\rho)$$

follows from Lemma 2. As in the previous cases, we must have $\delta\lambda_n(\rho) \leqq 0$. We show that it is always possible to find a function $q(x)$ such that

$$\delta\lambda_n(p) = -\lambda_n(p)\int_0^a \delta p(x)u_n^2(x)dx > 0$$

when $p(x) = \varepsilon q(x)$ where $p(x) + \delta p(x) \in E_3(M, a)$, unless $p(x) \in E_3(M, a)$ is a piecewise linear concave function with at most n pieces. Hence, it follows that $\rho(x)$ must be such a function.

We find the function $q(x)$ by the method used in the proof of Theorem 3. Thus, we seek a function $r(x)$ such that

$$\int_0^a r(x)u_n^2(x)dx \leqq \int_0^a p(x)u_n^2(x)dx .$$

Where $u_n(x)$ is the eigenfunction corresponding to $\lambda_n(p)$. To apply the method of Theorem 3, we consider

$$\int_0^a p(x)u_n^2(x)dx = \int_0^a [-p(x)][-u_n^2(x)]dx .$$

Then $-p(x)$ is convex and the zeros x_k $(k = 0, 1, 2, \cdots, n)$ of $u_n(x)$ are the maximum points of $-u_n^2(x)$. The maximum points $\bar{x}_k$ $(k = 1, 2, \cdots, n)$ of $u_n^2(x)$ are the minimum points of $-u_n^2(x)$.

Over each of the intervals (x_k, x_{k+1}) $(k = 0, 1, \cdots, n-1)$ we define $-r_k(x, c_k)$ where $-p(\bar{x}_k) \leqq c_k \leqq 0$. As in the convex case, there is a number γ_k such that $r_k(x, \gamma_k)$ is linear at $x = \bar{x}_k$. Using the inductive argument as before, we let $m(x) \equiv L(x)$ since $L(x)$ will be negative and form new functions $-r_k(x, \gamma_k)$. Finally we obtain $-r(x)$ which is convex and satisfies the inequality

$$\int_0^a p(x)u_n^2(x)dx \geqq \int_0^a r(x)u_n^2(x)dx .$$

Hence, choosing $q(x) = r(x) - p(x)$, we have

$$\int_0^a \delta p u_n^2(x)dx = \int_0^a \epsilon\, q(x)u_n^2(x)dx \leqq 0 ,$$

where for ε sufficiently small $p(x) + \delta p(x) \in E_3(M, a)$. Furthermore, there is strict inequality unless $p(x)$ is a concave piecewise linear function with at most n pieces. This proves the theorem.

It follows immediately from Theorem 4 that

$$\lambda_1(p) \leqq \frac{\pi^2}{aM}$$

when $p(x)$ is concave.[1] For in this case, $\rho(x)$ is a linear function. But, as was shown in the proof of Theorem 3, $\lambda_1(\bar{\rho}) \geqq \lambda_1(\rho)$ where $\bar{\rho}(x) = \frac{1}{2}[\rho(x) + \rho(a - x)]$. In this case, $\bar{\rho}(x) = M/a$ and $\lambda_1(M/a) = \pi^2/aM$.

5. The vibrating rod. The eigenvalue problem associated with a vibrating rod with clamped ends and density $p(x) \geqq 0$, $x \in [0, a]$ is

$$u^{iv} - \lambda p(x)u = 0 , \quad u(0) = u'(0) = u(a) = u'(a) = 0 . \tag{15}$$

As in the case of the string, we denote the ordered eigenvalues by

$$0 < \lambda_1(p) < \lambda_2(p) < \cdots .$$

That there should be strict inequalities in this expression has been

[1] This result has already been obtained by Z. Nehari. His proof is the one dimensional analog of that given in [7] where he shows that the lowest eigenvalue of a circular membrane with a superharmonic density $p(x, y)$ is bounded above by that of a homogeneous membrane of the same total mass.

shown in [6].

In this section, we consider the problem of finding upper bounds for these eigenvalues when $p(x)$ is restricted to be either monotone, convex or concave. In the first two cases, we require in addition that $p(x) \leqq H < \infty$. As in the case of the string, we denote the set of all functions $p(x) \geqq 0$, $x \in [0, a]$ with $\int_0^a p(x)dx = M$ where $p(x)$ is monotone increasing, convex and concave by $E_1(M, H, a)$, $E_2(M, H, a)$ and $E_3(M, a)$ respectively. The H in $E_1(M, H, a)$ and $E_2(M, H, a)$ indicates that in these cases $p(x) \leqq H$.

LEMMA 4. *Let E_k be one of the sets of functions defined above. There exists a function $\rho(x) \in E_k$ such that*

$$\lambda_n(\rho) = \sup_{p \in E_k} \lambda_n(p) .$$

This follows in exactly the same manner as the result of Lemma 2. We need only note that the result of Krein quoted in Lemma 2 may be generalized to this case. The generalization is trivial for the Green's function of the system (15) and its first partial derivatives are bounded. Krein's proof then applies word for word to this case and hence the proof of Lemma 4 follows as in the proof of Lemma 2.

LEMMA 5. *The first variation of $\lambda_n(p)$ with $\int_0^a p(x)dx = M$ is*

$$\delta\lambda_n(p) = -\lambda_n(p)\int_0^a \delta p(x)u_n^2(x)dx$$

where $u_n(x)$ is the normalized eigenfunction corresponding to $\lambda_n(p)$.

In particular we may choose $\delta p(x) = \varepsilon q(x)$ such that $\int_0^a \delta p(x) = 0$. The result is easily derived in the same way as the result of Lemma 3.

The results of Theorems 1, 3 and 4 will now generalize to the case of a vibrating rod with clamped ends. The only question which arises concerns the properties of the eigenfunction $u_n(x)$ corresponding to $\lambda_n(p)$. It must be true that $u_n(x)$ has the same general character as the nth eigenfunction of a vibrating string. In particular, it has been shown in [6] that $u_n(x)$ has exactly $n-1$ zeros in the open interval $(0, a)$. Furthermore $u_n^2(x)$ has exactly one maximum between any consecutive pairs of zeros. For suppose there are two or more maximum points between some consecutive pair of zeros. Then $u_n'(x)$ must have at least $n+4$ zeros in $[0, a]$. Hence $u_n''(x)$, $u_n'''(x)$ and $u_n^{(iv)}(x)$ must have at least $n+3$, $n+2$, and $n+1$ zeros respectively in the open interval $(0, a)$. This leads to a contradiction if $p(x) > 0$ since $u_n^{(iv)} = \lambda_n p(x)u_n(x)$ may have only $n-1$ zeros in $(0, a)$. If $p(x) \geqq 0$, we may apply the same argument with $p(x)$ replaced by $p(x)+\varepsilon$, $\varepsilon>0$. Thus, if $u_{n\varepsilon}(x)$ is the nth eigenfunction, $u^2_{n\varepsilon}(x)$ has n maximum points in $(0, a)$. Letting $\varepsilon \to 0$, we see that the same must be true of the nth eigenfunction when the rod density is $p(x) \geqq 0$.

From these observations, Lemmas 4 and 5, and the arguments used in Theorems 1, 3 and 4, we have the following result.

THEOREM 5. *Let $\lambda_n(p)$ be the* nth *eigenvalue of a rod with clamped ends and density $p(x)$, $x \in [0, a]$, such that $\int_0^a p(x)dx = M$.*

(a) *If $p(x)$ is monotone increasing and bounded*

$$\lambda_n(p) \leqq \lambda_n(\rho)$$

where $\rho(x)$, $x \in [0, a]$ is an increasing step function with at least one and at most n discontinuities in the open interval $(0, a)$ and $\int_0^a \rho(x)dx = M$.

(*b*) *If $p(x)$ is convex and bounded*

$$\lambda_n(p) \leqq \lambda_n(\rho)$$

where $\rho(x)$, $x \in [0, a]$ is a bounded piecewise linear convex function with at most $n+2$ linear pieces and $\int_0^a \rho(x)dx = M$.

(c) *If $p(x)$ is concave*

$$\lambda_n(p) \leqq \lambda_n(\rho)$$

where $\rho(x)$, $x \in [0, a]$ is a piecewise linear concave function with at most n linear pieces and $\int_0^a \rho(x)dx = M$.

In the case of the lowest eigenvalue, the density which gives the upper bound may be obtained precisely when $p(x)$ is convex or concave. It follows from the Rayleigh quotient as in Theorem 2 that for $\bar{p}(x) = \frac{1}{2}[p(x) + p(a-x)]$

$$\lambda_1(\bar{p}) \geqq \lambda_1(p) .$$

This and the above theorem thus show that when $p(x)$ is convex, $\rho(x)$ is symmetric and piece wise linear with at most three linear pieces and that when $p(x)$ is concave, $\rho(x)$ is a constant. This result may also be obtained by the method used in the proof of Theorem 2.

6. The membrane. We consider a vibrating membrane stretched with uniform unit tension over a disk $D = \{(x, y) \mid x^2 + y^2 < R^2\}$. We assume the areal density of the membrane is given by the measurable function $p(x, y)$ where

$$\iint_D p(x, y)dxdy = M .$$

For such a membrane with a fixed boundary, the eigenvalues and eigenfunctions are determined by the integral equation [8]

$$u(x, y) = \iint_D G(x, y, \xi, \eta)p(\xi, \eta)u(\xi, \eta)d\xi d\eta \tag{16}$$

where $G(x, y, \xi, \eta)$ is the Green's function of D. We denote the first eigenvalue by $\lambda_1(p)$ and the corresponding eigenfunction by $u_1(x, y)$.

We find upper bounds for $\lambda_1(p)$ by use of the following result.

LEMMA 6. *The lowest eigenvalue of a circular membrane with fixed boundary and integrable density $p(x, y)$ is always less than that of a circular membrane with fixed boundary and density.*

$$\bar{p}(x, y) = p(r) = \frac{1}{2\pi}\int_0^{2\pi} p(r\cos\theta, r\sin\theta)d\theta .$$

Proof. We use the fact the first eigenvalue is given by the infimum of the Rayleigh quotient

$$R(u) = \frac{\iint_D (u_x^2 + u_y^2)dxdy}{\iint_D p(x, y)u^2(x, y)dxdy}$$

where the infimum is taken over all functions $u(x, y) \in C'$ such that $u(x, y)$ vanishes on the boundary D. In particular, the lowest eigenvalue of a circular membrane with density $p(r)$ is given by

$$\lambda_1(p(r)) = \inf_{u \in C'} \frac{\iint_D (u_x^2 + u_y^2)dxdy}{\iint_D p(r)u^2(x, y)dxdy} .$$

We note that

$$p(r) = \frac{1}{2\pi}\int_0^{2\pi} p(r, \phi)d\phi = \frac{1}{2\pi}\int_0^{2\pi} p(r, \phi + \theta)d\phi = \bar{p}(r, \theta) .$$

Hence, it follows that

$$
\begin{aligned}
\frac{1}{\lambda_1(p(r))} &= \sup_{u\in C'} \frac{\iint_D \bar{p}(r,\phi)u^2(r,\phi)r d\phi\, dr}{\iint_D (u_x^2 + u_y^2)dxdy} \\
&= \sup_{u\in C'} \frac{\iint_D \left(\frac{1}{2\pi}\int_0^{2\pi} p(r,\phi+\theta)d\theta\right)u^2(r,\phi)rd\phi dr}{\iint_D (u_x^2 + u_y^2)dxdy} \\
&\leqq \frac{1}{2\pi}\int_0^{2\pi} d\theta \cdot \sup_{u\in C'} \frac{\iint_D p(r,\phi+\theta)u^2(r,\phi)rd\phi dr}{\iint_D (u_x^2 + u_y^2)dxdy} \\
&= \frac{1}{2\pi}\int_0^{2\pi} \frac{1}{\lambda_1(p)} d\theta = \frac{1}{\lambda_1(p)}
\end{aligned}
$$

i.e., $\lambda_1(p) \leqq \lambda_1(\bar{p})$ since $\lambda_1(p)$ does not depend on θ. We may now prove the following.

THEOREM 6. *The lowest eigenvalue of a circular membrane with fixed boundary and a bounded convex density $p(x, y)$ is less than the lowest eigenvalue of a circular membrane with density*

$$
q(r) = \begin{cases} 0\,, & 0 < r \leqq R - H/\alpha\,, \\ \alpha(r - R) + H\,, & R - H/\alpha < r \leqq R\,, \end{cases}
$$

if $R > H/\alpha$ and

$$
q(r) = \alpha(r - R) + H
$$

if $R < H/\alpha$ where α is such that $2\pi\int_0^R q(r)rdr = M$.

We first note that since $p(x, y)$ is convex, so is

$$
p(r) = \frac{1}{2\pi}\int_0^{2\pi} p(r\cos\phi, r\sin\phi)d\phi\ .
$$

For suppose r_1 and r_2 are such that $-R \leqq r_1 < r_2 \leqq R$. By the convexity of $p(x, y)$ we have

$$
p\left(\frac{r_1 + r_2}{2}\cos\phi, \frac{r_1 + r_2}{2}\sin\phi\right) \leqq \tfrac{1}{2}[p(r_1\cos\phi, r_1\sin\phi) + p(r_2\cos\phi, r_2\sin\phi)]\ .
$$

Integrating this with respect to ϕ, we have

$$
p\left(\frac{r_1 + r_2}{2}\right) \leqq \tfrac{1}{2}[p(r_1) + p(r_2)]\ .
$$

We now consider

$$\frac{1}{\lambda_1(p(r))} = \frac{\iint_D \bar{p}(x, y)u^2(x, y)dxdy}{\iint_D (u_x^2 + u_y^2)dxdy}$$

where $u(x, y)$ is the eigenfunction corresponding to $\lambda_1(p(r))$. For any function $u_1(x, y) \in C'$, we then have

$$\frac{1}{\lambda_1(p(r))} \geqq \frac{\iint_D \bar{p}(x, y)u_1^2(x, y)dxdy}{\iint_D (u_{1x}^2 + u_{1y}^2)dxdy}$$

In particular, if $u_1(x, y) = u_1(r)$ is the eigenfunction corresponding to the first eigenvalue $\lambda_1(q)$ of a membrane with density $q(r)$, it is a decreasing function of r. This is easily seen by considering the differential equation which is equivalent to the integral equation (16) [3]. By Lemma 1, we thus have

$$2\pi\int_0^R \bar{p}(r)u_1^2(r)rdr \geqq 2\pi\int_0^R q(r)u_1^2(r)rdr \;. \tag{17}$$

Hence,

$$\frac{1}{\lambda_1(p(r))} \geqq \frac{\iint_D q(r)u_1^2(r)rdrd}{\iint_D \text{grad}\, u_1^2\, rdrd} = \frac{1}{\lambda_1(q)} \,.$$

This same method yields a result if $p(x, y)$ is a concave function. For $p(r)$ is also concave and the inequality (17) holds if we choose $q(r) = \iint_D p(x, y)dxdy = M$. Hence we find that

$$\lambda_1(p) \leqq \frac{\pi j_0^2}{R^2} M$$

where j_0 is the least positive zero of $J_0(x) = 0$. As pointed out in [1], this result is a corollary to a theorem of Nehari [7] which says that if $p(x, y)$ is superharmonic in D, then $\lambda_1(p) \leqq \pi \frac{j_0^2}{R^2} M$. Since a concave function is superharmonic, this implies the above result.

REFERENCES

1. D. Banks, *Bounds for the eigenvalues of some vibrating systems*, Pacific J. Math., **10** (1960), 439-474.

2. L. Collatz, *Eigenwertprobleme*, Chelsea Publishing Co., 1948

3. R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Interscience Publishers, 1953.

4. E. Kamke, *Differentialgleichungen, Lösungsmethoden und Lösungen*, Chelsea Publishing Co., 1944.

5. M. G. Krein, *On certain problems on the maximum and minimum of characteristic values and on Lyapunov zones of stability*, Amer. Math. Soc. Trans., Series 2, **1** (1955), 163-187.

6. W. Leighton and Z. Nehari, *On the oscillation of solutions of self-adjoint linear differential equations of the fourth order*, Trans. Amer. Math. Soc., **89**, (1958), 325-377.

7. Z. Nehari, *On the principal frequency of a membrane*, Pacific J. Math., **8**, (1958), 285-293.

8. F. G. Tricomi, *Integral Equations*, Interscience Publishers, 1957.

UNIVERSITY OF CALIFORNIA, DAVIS

ON THE FIELD OF RATIONAL FUNCTIONS OF ALGEBRAIC GROUPS

A. BIALYNICKI-BIRULA

0. Introduction. Let K be an algebraically closed field of characteristic 0, let k be a subfield of K and suppose that G is a (k, K) algebraic group, i.e., an algebraic group defined over k and composed of K-rational points. Let $k(G)$ denote the fields of k-rational functions on G. G_k denotes the subgroup of G composed of all k-rational points of G. If $g \in G_k$ then the regular mapping $L_g(R_g)$ of G onto G defined by $L_g x = gx$ $(R_g x = xg)$ induces an automorphism of $k(G)$ denoted by $g_l(g_r)$. Let D_k denote the Lie algebra of all k-derivations of $k(G)$ (i.e., of all derivations of $k(G)$ that are trivial on k) which commute with g_r, for every $g \in G_k$.

For any subset A of $k(G)$ let $G(A)$ denote the subgroup of G composed of all elements g such that $g_r(f) = f$, for every $f \in A$. In the sequel we shall always assume that G_k is dense in G.

The main result of this paper is the following theorem:

THEOREM 1. *Let F be a subfield of $k(G)$ containing k. Then the following three conditions are equivalent*:

(1) *F is $(G_k)_l$ — stable*

(2) *F is D_k — stable*

(3) *$F = k(G/G(F))$ and so F coincides with the field of all elements of $k(G)$ that are fixed under $G(F)_r$.*

By means of the theorem, we establish a Galois correspondence between a family of subgroups of G and the family of $(G_k)_l$-stable subalgebras of the algebra of representative functions of G.

The author wishes to express his thanks to Professor G.P. Hochschild and Professor M. Rosenlicht for a number of instructive conversations on the subject of this note.

1. Let K be an algebraically closed field of characteristic 0, let k be a subfield of K and suppose that V, W are (k, K) — algebraic varieties. Let $k(V)$, $k(W)$ denote the fields of k-rational functions on V and W, respectively. If A is a subset of $k(V)$ then $k(A)$ denotes the fields generated by k and A.

The following result is known[1]:

(1) Let F be a rational mapping of V onto a dense subset of W and let φ be the cohomomorphism corresponding to F. Then there exists

Received September 28, 1960, in revised form November 14, 1960.

[1] See e.g. [2],

an open subset $W_1 \subset W$ such that $F^{-1}(x)$ contains exactly $[k(V): \varphi(k(W))]$ elements, for every $x \in W_1$.

LEMMA 1. *Let A be a subset of $k(V)$ and suppose that there exists a dense set $V_1 \subset V$ and an open subset $V_2 \subset V$ such that for any two distinct points x_1, x_2, where $x_1 \in V_1, x_2 \in V_2$, there exists a function $f \in A$ which is defined at x_1, x_2 and $f(x_1) \neq f(x_2)$. Then $k(A) = k(V)$.*

Proof. Let B be a finite subset of A, say $B = \{f_1, \cdots, f_n\}$. Then F_B denotes the rational mapping $F_B\colon V \to K^n$ defined by $F_B(x) = (f_1(x), \cdots, f_n(x))$ and $W_B = (F_B(V)^- \subset K^n$. Let $\Delta(W_B)$ be the diagonal of $W_B \times W_B$ and $V_B = ((F_B \times F_B)^{-1}\Delta(W_B))^- \subset V \times V$. Then there exists a finite subset $B_0 \subset A$ such that $V_{B_0} \subset V_B$, for every finite subset $B \subset A$ (since $V \times V$ satisfies the minimal condition for closed sets). Let V_0 be an open subset of V such that F_{B_0} is regular on V_0. We may assume that $V_0 = V_2 = V$, since we may replace V by $V_0 \cap V_2$. If $x_1 \in V_1$, $x_2 \in V$ and $x_1 \neq x_2$ then there exists $f \in A$ such that f is defined at x_1, x_2 and $f(x_1) \neq f(x_2)$. Hence $(x_1, x_2) \notin V_{\{f\}}$ and so $(x_1, x_2) \notin V_{B_0}$. Thus $F_{B_0}(x_1) \neq F_{B_0}(x_2)$. Therefore, for every $x \in F_{B_0}(V_1)$, $F_{B_0}^{-1}(x)$ contains exactly one element. But $F_{B_0}(V_1)$ is dense in W_{B_0}. Hence it follows from (i) that $[k(V): k(B_0)] = 1$, i.e., $k(V) = k(B_0)$. Thus $k(V) = k(A)$.

Let G be a (k, K) – algebraic group. Suppose that G_k is dense in G. Let D be the Lie algebra of all derivations of $K(G)$ commuting with g_r, for every $g \in G$, and let D_k denote the Lie algebra consisting of all derivations from D that map $k(G)$ into $k(G)$. Let $k[D]$ ($K[D]$) denote the k-algebra (K – algebra) of transformations generated by the identity map and $D_k(D)$.

If $d \in D_k$ then d restricted to $k(G)$ is a k-derivation commuting with g_r, for every $g \in G_k$. On the other hand if d_1 is a k-derivation of $k(G)$ commuting with g_r, for every $g \in G_k$, then there exists a unique extension d of d_1 to a K-derivation of $K(G)$, and the extension belongs to D_k. Hence we may identify D_k and the Lie algebra of all k-derivations of $k(G)$ that commute with g_r, for every $g \in G_k$.

(ii)[2] If $f \in K(G)$ and f is defined at a point $g \in G$ then df is defined at g, for any $d \in K[D]$.

LEMMA 2. *Let $f \in K(G)$ and suppose that f is defined at $g \in G_k$. If $f \neq 0$ then there exists $d \in k[D]$ such that $(df)(g) \neq 0$.*

Proof. Suppose that $f \neq 0$. If $f(g) \neq 0$ then the identity element of $k[D]$ satisfies the desired condition. Hence we may assume that $f(g) = 0$, Let $\mathcal{O}_k(\mathcal{O}_K)$ denote the local ring of g in $k(G)$ $(K(G))$ and let $m_k(m_K)$ be the maximal ideal of $\mathcal{O}_k(\mathcal{O}_K)$. Then $f \in m_K$. Let

[2] See [4] p.51,

$x_1, \cdots, x_m$ be elements of m_k such that $x_1 + m_k^2, \cdots, x_m + m_k^2$ is a k-basis of m_k/m_k^2. The $x_1 + m_K^2, \cdots, x_m + m_K^2$ is a K-basis of m_K/m_K^2. Hence every mapping $(x_1, \cdots, x_m) \to k$ can be extended to a derivation ∂: $\mathcal{O}_K \to K$. On the other hand $f \neq 0$ and so there exists an integer t such the $f \in m_K^t - m_K^{t+1}$. Hence $f = \sum_{i_1+\cdots+i_m=t} a_{i_1,\ldots,i_m} x_1^{i_1}, \cdots, x_m^{i_m} + f_1$, where $f_1 \in m_K^{t+1}$, $a_{i_1,\ldots,i_m} \in K$ and at least one $a_{i_1,\ldots,i_m}$ is different from zero. Let ∂_i be the derivation of $\mathcal{O}_K$ into K such than $\partial_i\, x_j = \delta_{ij}$, where $\delta_{ij} = \begin{cases} 0 \text{ if } i \neq j \\ 1 \text{ if } i = j. \end{cases}$ It is known[3], that there exist $d_i \in D_k$ such that $(d_i f)(g) = \partial_i f$ for every $f \in \mathcal{O}_K$. Then $(d_1^{i_1} \cdots d_m^{i_m}) f(g) = i_1! \cdots i_m! a_{i_1,\ldots,i_m} \neq 0$ if $a_{i_1,\ldots,i_m} \neq 0$. Hence the lemma is proved.

If A is a subset of $k(G)$ then $G(A)$ denotes the subgroup of G composed of all elements g such that g_r leaves the elements of A fixed. For any $A \subset k(G)$, $G(A)$ is a k-closed subgroup of G.

(iii)[4] Let G_1 be a k-closed subgroup of G. Then G/G_1 is defined over k. Let φ be the cohomomorphism of $k(G/G_1)$ into $k(G)$ corresponding to the canonical mapping $G \to G/G_1$. Then $\varphi(k(G/G_1))$ coincides with the subfield of all elements of $k(G)$ which are fixed under g_r, for every $g \in G_1$. In the sequel we shall identify $k(G/G_1)$ and $\varphi(k(G/G_1))$.

Proof of the theorem.

Implications $(3) \Rightarrow (1)$ and $(3) \Rightarrow (2)$ are obvious.

$(1) \Rightarrow (3)$[5]. Let $g_1 \in G_k$, $g_2 \in G$ and $G(F)g_1 \neq G(F)g_2$. Then $g_2 g_1^{-1} \notin G(F)$. Hence there exists $f_0 \in F$ such that $(g_2 g_1^{-1})_r f_0 \neq f_0$. Therefore there exists an element $g \in G_k$ such that $(g_2 g_1^{-1})_r f_0$ and f_0 are defined at g and $(g_2 g_1^{-1})_r f_0(g) \neq f_0(g)$, i.e., $f_0(g_2 g_1^{-1} g) \neq f_0(g)$, $(g_1^{-1} g)_l f_0(g_2) \neq (g_1^{-1} g)_l f_0(g_1)$. Let $f = (g_1^{-1} g)_l f_0$. Then $f \in F$ since $g_1^{-1} g \in G_k$; f is defined at g_1 and g_2, and $f(g_1) \neq f(g_2)$. Thus it follows from Lemma 1 that $F = k(G/G(F))$, because $G(F) \cdot G_k / G(F)$ is dense in $G/G(F)$.

$(2) \Rightarrow (3)$. Let $f_1, \cdots, f_n$ be a set of generators of F over k, and let V_1 be an open subset of G such that $f_1, \cdots, f_n$ are regular on V_1. We may assume that $V_1 = G(F)V_1$. Let $g_1 \in V_1 \cap G_k$, $g_2 \in V_1$, $G(F)g_1 \neq G(F)g_2$. Then $g_2 g_1^{-1} \notin G(F)$ and so there exists f_i such that $(g_2 g_1^{-1})_r f_i \neq f_i$. We know that $(g_2 g_1^{-1})_r f_i$ and f_i are defined at g_1. Hence it follows from Lemma 2 that there exists an element $d \in k[D]$ such that

$$d((g_2 g_1^{-1})_r f_i)(g) \neq (df_i)(g), \text{ i.e., } (df_i)(g_1) \neq (df_i)(g_2)\,.$$

Therefore, for any pair of distinct elements $G(F)g_1$, $G(F)g_2$ such that

$$G(F)g_1 \in G(F) \cdot G_k \cap V_1/G(F) \text{ and } G(F)g_2 \in V_1/G(F)\,,$$

[3] See [4] p. 51,

[4] See Proposition 2, p. 495 in [5].

[5] This part of the proof is modeled after the proof of Lemma 5.3 p. 515 in [3].

there exists an element $f \in F$ which is defined at $G(F)g_1$, $G(F)g_2$ and such that $f(G(F)g_1) \neq f(G(F)g_2)$. But $V_1/G(F)$ is an open subset of $G/G(F)$, and $G(F)G_k \cap V_1/G(F)$ is dense in $G/G(F)$. Hence it follows from Lemma 1 that $F = k(G/G(F))$.

This completes the proof of the theorem.

2. Applications. As a consequence of Lemma 2 one can get the following corollary:

COROLLARY. *If α is an automorphism of $k(G)$ commuting with D_k and leaving the elements of k fixed then there exists $h \in G_k$ such that $\alpha = h_r$.*

Proof. α induces a rational map $F_\alpha: G \to G$. Let $g \in G_k$ be a point such that F_α is defined at g and let $F_\alpha(g) = h^{-1}g$ Then $h \in G_k$ and $f(g) = (\alpha f)(h^{-1}g)$, for every $f \in k(G)$ that is defined at g. Hence $(df)g = (\alpha(df))(h^{-1}g)$, for every $d \in k[D]$. But $(\alpha(df))(h^{-1}g) = (h_r^{-1}(\alpha(df)))(g)$ and d commutes with α and h_r^{-1}. Therefore $(df)(g) = (d(h_r^{-1}(\alpha f)(g)))$. Hence it follows from Lemma 2 that $f = h_r^{-1}(\alpha f)$. Thus $h_r f = \alpha f$, for every f that is defined at g. Therefore $h_r f = \alpha f$, for every $f \in k(G)$.

It follows from the corollary that if F is a D_k — stable subfield of $k(G)$ containing k then every D_k — automorphism of $k(G)$ leaving the elements of F fixed belongs to $G(F)_r$, i.e., the D_k — Galois group of $k(G)$ over F coincides with $G(F)_r$. Combining this result and the above theorem we obtain that there exists the usual one to one Galois correspondence between D_k — stable subfields of $k(G)$ containing k and k-closed subgroups of G.

Let $k[G]$ denote the ring of regular (i.e., representative) functions on G. Let $\mathscr{R}$ be the family of all $(G_k)_l$ — stable (or, equivalently, D_k — stable) subrings R of $k[G]$ containing k and satisfying the following condition if $f \in R, g \in R$ and $f/g \in k[G]$ then $f/g \in R$. Let $\mathscr{G}$ denote the family of all k-closed subgroups H of G such that G/H is isomorphic to an open subset of an affine variety.

THEOREM 2. *The mappings $H \to k[G] \cap k(G/H)$ and $R \to G(R)$ establish a Golois correspondence between $\mathscr{G}$ and $\mathscr{R}$*[6].

Proof. $H \in \mathscr{G}$ then $k[G] \cap k(G/H) \in \mathscr{R}$ and $G(k[G] \cap k(G/H)) = H$, since $k(G/H)$ is generated by $k[G] \cap k(G/H)$.

Now, if $R \in \mathscr{R}$ then $G(R) \in \mathscr{G}$. In fact, if $R \in \mathscr{R}$, then $k(R)$ is $(G_k)_l$ — stable and so $k(R) = k(G/G(R))$. For every $f \in R$, $(G_k)_l f$ generates a finite dimensional k-vector space, Hence there exists a finitely generated over k $(G_k)_l$ — stable subring R_0 of R such that $k(R_0) = k(R)$. Let W denote

[6] C.f. [1] p. 324.

the affine variety that has R_0 as its coordinate ring. One can define a structure of a G-homogeneaus space on W, since $K[R_0]$ is G_ι — stable. Let η be the canonical mapping of $G/G(R)$ into W. Then η commutes with the action of G and is birational. Hence η is an isomorphism of $G/G(R)$ onto an open subset $\eta(G/G(R))$ of W.

Moreover, $R = k[G] \cap k(G/G(R))$, since $R \in \mathscr{R}$ and $k(R) = k(G/G(R))$. This completes the proof of the theorem.

Added in Proof. The equivalence (1) $\Longleftrightarrow$ (2) of Theorem 1 in the case where k is algebraically closed has been proved by E. Abe and T. Kanno (Tohoku Math. Jour. 2nd series **11** (1959), 376–384).

REFERENCES

1. P. Cartier, *Dualité de Tannaka des graoups et des algèbres de Lie*, Comptes Rendus de l'Académie des Sciences, Paris, **242** (1956), 322-325
2. C. Chevalley, *Fondements de la Géometrie Algebrique*, Paris, L'Ecole Normale Superieure, 1958.
3. G. Hochschild and G. D. Mostow, *Representations and representative functions of Lie groups*, Annals of Math., **66** (1957), 495-542.
4. M. Rosenlicht, *A note on derivations and differentials on algebraic varieties*, Portugaliae Mathematica, **16** (1957), 43-55.
5. A. Weil, *On algebraic groups and homogeneous spaces*, Amer. J. Math., **77** (1955), 493-512.

UNIVERSITY OF CALIFORNIA AND POLISH ACADEMY OF SCIENCES

SIMPLE PATHS ON CONVEX POLYHEDRA

THOMAS A BROWN

1. Introduction. In problems of linear programming, one sometimes wants to find all vertices of a given convex polyhedron. An algorithm for finding all such vertices will often define a path which passes from vertex to vertex along the edges of the polyhedron in question [1], and thus it is natural to ask, as Balinski does in [2], whether or not one can always find a path along the edges of a convex polyhedron which visits each vertex once and only once. This question has been answered in the negative independently by Grünbaum and Motzkin [5] and the author [3]. The purpose of the present paper is to present a modification of the results of [3], and answer certain questions which were asked by Grünbaum and Motzkin.

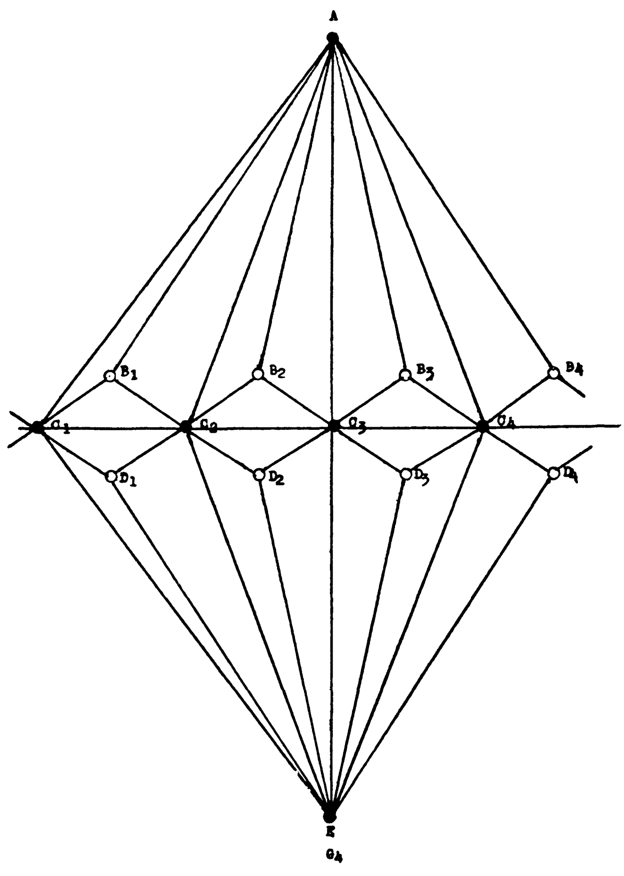


Figure 1.

Received September 5, 1960 in revised form October 20, 1960.

2. Path numbers and path lengths. For any graph G with $n(G)$ nodes we let $m(G)$ denote the number of disjoint simple paths required to cover all vertices of G, and let $p(G)$ denote the maximum number of nodes contained in a simple path on G. We call $m(G)$ the "path number" of G and $p(G)$ the "path length" of G. If G can be represented as the edges and vertices of a convex polyhedron in three-dimensional space, we say that G is "3-polyhedral". Now let

$$p(n) = \min\{p(G)\colon\ G \text{ is 3-polyhedral and } n(G) = n\}$$
$$m(n) = \max\{m(G)\colon G \text{ is 3-polyhedral and } n(G) = n\}\ .$$

We will show, by means of a class of examples, that $m(n) \geqq (n - 10)/3$ and $p(n) \leqq (2n + 13)/3$ for all n.

3. The graphs G_k. Let the graph $G_k (k \geqq 3)$ have $3k + 2$ vertices, which we will denote by a, b_i, c_i, d_i, and $e(i$ ranging from 1 to $k)$. Let the edges of G_k be (a, b_i), (a, c_i), (e, d_i), (e, c_i), (c_i, c_{i+1}), (c_i, b_i), (c_i, d_i), (d_i, c_{i+1}), and (b_i, c_{i+1}). Thus a and e are of valence $2k$, the c_i are of valence 8, and the b_i and d_i are of valence 3. See Figure 1 for a drawing of G_4. G_k can be represented as a triangulation of the plane, and it is easy to show by induction [4] that if $n(G) \geqq 4$ and G can be

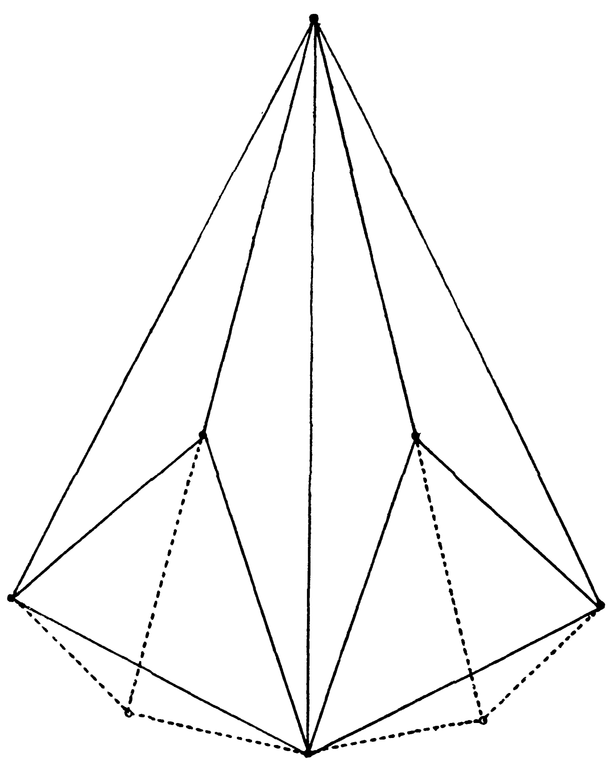


Figure 2.

represented as a triangulation of the plane, then G can be represented as the edges and vertices of a convex polyhedron in 3-space. Alternatively, one could apply the "Fundamentalsatz der Konvexen Typen" of E. Steinitz [6]. But in the case of G_k it is really unnecessary to use any such general results, for G_k is clearly the graph of the polyhedron obtained by appropriately truncating a bipyramid whose base is a regular $2k$-gon (Figure 2 illustrates how the top half of a bipyramid should be truncated in obtaining G_4).

If we color a, c_i, and e black and let b_i and d_i be white (where i ranges from 1 to k), then G_k consists of $n+2$ black nodes and $2n$ white ones. Since each white node has only black neighbors, each simple path in G_k must contain at most one more white node than black. Thus at least $2k-(k+2)=k-2$ disjoint simple paths are required to visit every node of G_k. The following set of paths shows that the path-number of G_k is, in fact, exactly $k-2$:

$$b_1 \to c_1 \to d_1 \to e \to d_2 \to c_2 \to b_2 \to a \to b_3 \to c_3 \to d_3$$
$$b_i \to c_i \to d_i \quad (i = 4, \cdots, k)\,.$$

Similarly, since no simple path can contain more than $k+2$ black vertices, it follows that no simple path can contain more than

$$(k+2)+(k+3)=2k+5$$

vertices. It is easy to construct simple paths containing exactly this many vertices, and thus the path-length of G_k is $2k+5$. Since $n(G_k)=3k+2$, it follows that if $n \equiv 2 (\bmod 3)$,

$$p(n) \leqq \frac{2n+11}{3}$$
$$m(n) \geqq \frac{n-8}{3}\,.$$

To get bounds for $n \equiv 1 \pmod 3$, consider the graph G_k^- obtained by omitting one white vertex from G_k. For $n \equiv 0 \pmod 3$, consider the graph G_k^+ obtained by adjoining to G_k a vertex connected to c_1, d_1, and e. It follows that

$$\left.\begin{aligned} p(n) &\leqq \frac{2n+13}{3} \\ m(n) &\geqq \frac{n-10}{3} \end{aligned}\right\} n \equiv 1 \pmod 3 \qquad \left.\begin{aligned} p(n) &\leqq \frac{2n+13}{3} \\ m(n) &\geqq \frac{n-9}{3} \end{aligned}\right\} n \equiv 0 \pmod 3\,.$$

Grünbaum and Motzkin asked if $n(G)=p(G)$ provided all of the faces of the polyhedron representing G were triangles, and our examples show that this is not the case. They further conjectured that

$$\max_{n(G)=n} m(G)\cdot p(G) \geqq n^{1+\gamma} \qquad \text{for some } \gamma > 0 .$$

Our examples show that

$$\max_{n(G)=n} m(G)\cdot p(G) \geqq \frac{2n^2 - 7n\,130}{9} .$$

Thus for any $\gamma < 1$ we can find an N_γ such that

$$\max_{n(G)=n} m(G)\cdot p(G) > n^{1+\gamma} \qquad \text{for all } n \geqq N_\gamma .$$

Furthermore, this result is the best possible in a sense; for since $m(G) < n$ and $p(G) \leqq n$, it follows that

$$\max_{n(G)=n} m(G)\cdot p(G) < n^2 \qquad \text{for all } n .$$

I want to thank Dr. Michel Balinski for drawing this subject to my attention, and the referee for making me aware of the paper by Grünbaum and Motzkin.

Bibliography

1. Michel L. Balinski, *An Algorithm for Finding All Vertices of Convex Polyhedral Sets,* Doctoral Dissertation, Princeton University, June 1959.
2. ———, *On the graph structure of convex polyhedra in n-space.* Pacific J. Math., (to appear).
3. T. A. Brown, *Hamiltonian Paths on Convex Polyhedra,* unpublished note, the RAND Corporation, August, 1960.
4. ———, *The Representation of Planar Graphs by Convex Polyhedra,* unpublished note, the RAND Corporation, August, 1960.
5. B. Grünbaum, and T. S. Motzkin, *Longest Simple Paths in Polyhedral Graphs,* (to appear).
6. E. Steinitz, and H. Rademacher, *Vorlesungen über die Theorie des Polyeder,* Springer, Berlin. 1934.

THE RAND CORPORATION AND HARVARD UNIVERSITY

SOME CONGRUENCES FOR THE BELL POLYNOMIMALS

L. CARLITZ

1. Let $\alpha_1, \alpha_2, \alpha_3, \cdots$ denote indeterminates. The Bell polynomial $\phi_n(\alpha_1, \alpha_2, \alpha_3, \cdots)$ may be defined by $\phi_0 = 1$ and

$$\phi_n = \phi_n(\alpha_1, \alpha_2, \alpha_3, \cdots) = \sum \frac{n!}{k_1!(1!)^{k_1}k_2!(2!)^{k_2}\cdots}\alpha_1^{k_1}\alpha_2^{k_2}\cdots, \tag{1.1}$$

where the summation is over all nonnegative integers k_j such that

$$k_1 + 2k_2 + 3k_3 + \cdots = n .$$

For references see Bell [2] and Riordan [5, p. 36]. The general coefficient

$$A_n(k_1, k_2, k_3, \cdots) = \frac{n!}{k_1!(1!)^{k_1}k_2!(2!)^{k_2}\cdots} \tag{1.2}$$

is integral; this is evident from the representation

$$A_n(k_1, k_2, k_3, \cdots) = \frac{n!}{k_1!(2k_2)!(3k_3)!\cdots}\cdot\frac{(2k_2)!}{k_2!(2!)^{k_2}}\frac{(3k_3)!}{k_3!(3!)^{k_3}}\cdots$$

and the fact that the quotient

$$\frac{(rk)!}{k!(r!)^k}$$

is integral [1, p. 57].

The coefficient $A_n(k_1, k_2, k_3, \cdots)$ resembles the multinomial coefficient

$$M(k_1, k_2, k_3 \cdots) = \frac{(k_1 + k_2 + k_3 + \cdots)!}{k_1!k_2!k_3\cdots} .$$

If p is a fixed prime it is known [3] that $M(k_1, k_2, k_3, \cdots)$ is prime to p if and only if

$$k_i = \sum_j a_{ij}p^j \qquad (0 \leqq a_{ij} < p) ,$$

$$k_1 + k_2 + k_3 + \cdots = \sum_j a_jp^j \qquad (0 \leqq a_j < p)$$

and

$$\sum_i a_{ij} = a_j \qquad (j = 0, 1, 2, \cdots) .$$

It does not seem easy to find an analogous result for $A_n(k_1, k_2, k_3, \cdots)$. For some special results see § 3 below.

Received October 31, 1960

Bell [2] showed that

$$\phi_p \equiv \alpha_1^p + \alpha_p \pmod{p} \tag{1.3}$$

and also determined the residues (mod p) of $\phi_{p+1}, \phi_{p+2}, \phi_{p+3}$. He also obtained an expression for the residue of ϕ_{p+r} as a determinant of order $r+1$. Generalizing (1.3) we shall show first that

$$\phi_{p^r} \equiv \alpha_1^{p^r} + \alpha_p^{p^{r-1}} + \cdots + \alpha_{p^r} \pmod{p} \tag{1.4}$$

and that

$$\phi_{pn}(\alpha_1, \alpha_2, \alpha_3, \cdots) \equiv \phi_n(\phi_p, \alpha_{2p}, \alpha_{3p}, \cdots) \pmod{p} \tag{1.5}$$

for all $n \geqq 1$. Note that on the right the first argument in ϕ_n is ϕ_p and not α_p.

2. From (1.1) we get the generating function

$$\sum_{n=0}^{\infty} \phi_n \frac{t^n}{n!} = \exp\left(\alpha_1 t + \alpha_2 \frac{t^2}{2!} + \alpha_3 \frac{t^3}{3!} + \cdots\right). \tag{2.1}$$

Indeed this may be taken as the definition of ϕ_n. Differentiating with respect to t we get

$$\sum_{n=0}^{\infty} \phi_{n+1} \frac{t^n}{n!} = \sum_{n=0}^{\infty} \phi_n \frac{t^n}{n!} \sum_{r=0}^{\infty} \alpha_{r+1} \frac{t^r}{r!} ,$$

so that

$$\phi_{n+1} = \sum_{r=0}^{n} \binom{n}{r} \phi_{n-r} \alpha_{r+1} . \tag{2.2}$$

Since the binomial coefficient

$$\binom{pn}{r} \equiv 0 \pmod{p}$$

unless $p|r$ and

$$\binom{pn}{pr} \equiv \binom{n}{r} \pmod{p}$$

it follows from (2.2) that

$$\phi_{pn+1} \equiv \sum_{r=0}^{n} \binom{n}{r} \phi_{p(n-r)} \alpha_{pr+1} \pmod{p} . \tag{2.3}$$

If for brevity we put

$$A(t) = \sum_{r=1}^{\infty} \alpha_r t^r / r! ,$$

so that

$$\sum_{n=0}^{\infty} \phi_n \frac{t^n}{n!} = \exp A(t) ,$$

it is easily seen by repeated differentiation and by (1.3) that

$$\sum_{n=0}^{\infty} \phi_{n+p} \frac{t^n}{n!} \equiv \{(A'(t))^p + A^{(p)}(t)\} e^{A(t)} \pmod{p} . \tag{2.4}$$

(By the statement

$$\sum_{n=0}^{\infty} A_n \frac{t^n}{n!} \equiv \sum_{n=0}^{\infty} B_n \frac{t^n}{n!} \pmod{m} ,$$

where A_n, B_n are polynomials with integral coefficients, is meant the system of congruences

$$A_n \equiv B_n \pmod{m} \qquad (n = 0, 1, 2, \cdots)) .$$

Hurwitz [4, p. 345] has proved the lemma that if $a_1, a_2, a_3, \cdots$ are arbitrary integers then

$$\left(\sum_{n=1}^{\infty} a_n \frac{x^n}{n!}\right)^k \equiv 0 \pmod{k!} .$$

The proof holds without change when the a_n are indeterminates. Since

$$A'(t) = \sum_{n=0}^{\infty} \alpha_{n+1} \frac{t^n}{n!} ,$$

it follows easily from Hurwitz's lemma that

$$(A'(t)^p = \left(\alpha_1 + \sum_{n=1}^{\infty} \alpha_{n+1} \frac{t^n}{n!}\right)^p \equiv \alpha_1^p \pmod{p} .$$

Thus (2.4) becomes

$$\sum_{n=0}^{\infty} \phi_{n+p} \frac{t^n}{n!} \equiv \left(\alpha_1^p + \sum_{r=0}^{\infty} \alpha_{r+p} \frac{t^r}{r!}\right) \sum_{n=0}^{\infty} \phi_n \frac{t^n}{n!} ,$$

which yields

$$\phi_{n+p} \equiv (\alpha_1^p + \alpha_p)\phi_n + \sum_{r=1}^{n} \binom{n}{r} \alpha_{r+p} \phi_{n-r} \pmod{p} . \tag{2.5}$$

In particular, for $n = 0$, (2.5) reduces to Bell's congruence (1.3). Similarly

$$\phi_{p+1} \equiv (\alpha_1^p + \alpha_p)\alpha_1 + \alpha_{p+1} \equiv \phi_p \alpha_1 + \alpha_{p+1} ,$$
$$\phi_{p+2} \equiv \phi_p \phi_2 + 2\alpha_{p+1}\alpha_1 + \alpha_{p+2} ,$$

and so on.

We remark that (2.5) is equivalent to Bell's congruence involving a determinant [2, p. 267, formula (6.5)]. Also for $s = \alpha_1 = \alpha_2 = \cdots$, (2.5) reduces to

$$(2.5)' \qquad \begin{aligned} a_{n+p}(s) &\equiv (s^p + s)a_n(s) + s \sum_{r=1}^{n} \binom{n}{r} a_{n-r}(s) \\ &\equiv a_{n+1}(s) + s^p a_n(s) \qquad \pmod p , \end{aligned}$$

where [5, p. 76]

$$a_n(s) = \phi_n(s, s, \cdots) = \sum_k S(n, k)s^k$$

and $S(n, k)$ denotes the Stirling number of the second kind. The congruence (2.5)' is due to Touchard [6].

If in (2.5) we replace n by pn we get

$$(2.6) \qquad \phi_{p(n+1)} \equiv \phi_p \phi_{np} + \sum_{r=1}^{n} \binom{n}{r} \alpha_{p(r+1)} \phi_{p(n-r)} \pmod p$$

for all $n = 0, 1, 2, \cdots$. Thus ϕ_{pn} is congruent to a polynomial in ϕ_p, $\alpha_{2p}, \alpha_{3p}, \cdots$ alone. Moreover, comparing (2.6) with (2.2), it is clear that

$$(2.7) \qquad \phi_{pn} \equiv \phi_n(\phi_p, \alpha_{2p}, \alpha_{3p}, \cdots) \pmod p ,$$

so that we have proved (1.5).

Replacing n by pn in (2.7) we get

$$\phi_{p^2 n} \equiv \phi_{pn}(\phi_p, \alpha_{2p}, \alpha_{3p}, \cdots) \equiv \phi_n(\phi_p^p + \alpha_{p^2}, \alpha_{2p^2}, \alpha_{3p^2}, \cdots) .$$

In particular for $n = 1$

$$\phi_{p^2} \equiv \phi_p^p + \alpha_{p^2} \equiv \alpha_1^{p^2} + \alpha_p^p + \alpha_{p^2} .$$

Again replacing n by pn we get

$$\phi_{p^3 n} \equiv \phi_n(\phi_{p^2}^p + \alpha_{p^3}, \alpha_{2p^3}, \alpha_{3p^3}, \cdots) ,$$

so that in particular

$$\phi_{p^3} \equiv \phi_{p^2}^p + \phi_{p^3} \equiv \alpha_1^{p^3} + \alpha_p^{p^2} + \alpha_{p^2}^p + \alpha_{p^3} .$$

Continuing in this way we see that

$$(2.8) \qquad \phi_{p^r n} \equiv \phi_n(\phi_{p^r}, \alpha_{2p^r}, \alpha_{3p^r}, \cdots) \pmod p$$

and

$$(2.9) \qquad \phi_{p^r} \equiv \phi_{p^{r-1}}^p + \alpha_{p^r} \equiv \alpha_1^{p^r} + \alpha_p^{p^{r-1}} + \cdots + \alpha_{p^r} \pmod p .$$

We have therefore proved (1.4) as well as the more general congruence (2.8).

Since

$$\phi_2 = \alpha_1^2 + \alpha_2 ,$$
$$\phi_3 = \alpha_1^3 + 3\alpha_1\alpha_2 + \alpha_3 ,$$
$$\phi_4 = \alpha_1^4 + 6\alpha_1^2\alpha_2 + 4\alpha_1\alpha_3 + 3\alpha_2^2 + \alpha_4 ,$$

it follows from (2.8) that

$$\begin{cases}\phi_{2p^r} \equiv \phi_{p^r}^2 + \alpha_{2p^r} , \\ \phi_{3p^r} \equiv \phi_{p^r}^3 + 3\phi_{p^r}\alpha_{2p^r} + \alpha_{3p^r} , \\ \phi_{4p^r} \equiv \phi_{p^r}^4 + 6\phi_{p^r}^2\alpha_{2p^r} + 4\phi_{p^r}\alpha_{3p^r} + 3\alpha_{2p^r}^2 + \alpha_{4p^r} ,\end{cases} \tag{2.10}$$

and so on.

We note also that (2.3) implies

$$\begin{cases}\phi_{p^r+1} \equiv \phi_{p^r}\alpha_1 + \alpha_{p^r+1} , \\ \phi_{2p^r+1} \equiv \phi_{2p^r}\alpha_1 + 2\phi_{p^r}\alpha_{p^r+1} + \alpha_{2p^r+1} , \\ \phi_{3p^r+1} \equiv \phi_{3p^r}\alpha_1 + 3\phi_{2p^r}\alpha_{p^r+1} + 3\phi_{p^r}\alpha_{2p^r+1} + \alpha_{3p^r+1} .\end{cases} \tag{2.11}$$

3. By means of (1.5) we can obtain certain congruences for the coefficient $A(k_1, k_2, k_3, \cdots)$. Indeed by (1.1) and (1.3)

$$\begin{aligned}&\phi_n(\phi_p, \alpha_{2p}, \alpha_{3p}, \cdots) \\ &\equiv \sum A_n(k_1, k_2, k_3, \cdots)(\alpha_1^p + \alpha_p)^{k_1}\alpha_{2p}^{k_2}\alpha_{3p}^{k_3} \cdots \qquad \pmod p ,\end{aligned} \tag{3.1}$$

where the summation is over nonnegative k_j such that

$$k_1 + 2k_2 + 3k_3 + \cdots = n .$$

The right member of (3.1) is equal to

$$\sum_{(k_j)} A_n(k_1, k_2, k_3, \cdots) \sum_{r=0}^{k_1} \binom{k_1}{r} \alpha_1^{p(k_1-r)}\alpha_p^r\alpha_{2p}^{k_2}\alpha_{3p}^{k_3} \cdots . \tag{3.2}$$

On the other hand

$$\phi_{pn} = \sum A_{pn}(h_1, h_2, h_3, \cdots)\alpha_1^{h_1}\alpha_2^{h_2}\alpha_3^{h_3} \cdots , \tag{3.3}$$

summed over

$$h_1 + 2h_2 + 3h_3 + \cdots = p_n . \tag{3.4}$$

It follows from (1.5) that

$$A_{pn}(h_1, h_2, h_3, \cdots) \equiv 0 \qquad \pmod p$$

except possibly when

$$h_j = 0 \qquad (j > 1,\ p \nmid j) . \tag{3.5}$$

When this condition is satisfied (3.4) becomes

$$h_1 + p(h_p + 2h_{2p} + \cdots) = pn ;$$

consequently $h_1 = pk_1$ and (3.3) becomes

$$\phi_{pn} \equiv \sum A_{pn}(pk_1, 0, \cdots, 0, h_p, \cdots)\alpha_1^{pk_1}\alpha_h^{h_p}\alpha_{2p}^{h_{2p}} \cdots .$$

We have therefore proved the following result:

THEOREM 1. *The coefficient* $A_{pn}(h_1, h_2, h_3, \cdots)$ *occurring in* (3.3) *is certainly divisible by* p *unless* (3.5) *is satisfied and* $h_1 = pk_1$. *If these conditions are satisfied then*

$$A_{pn}(h_1, h_2, h_3, \cdots) \equiv \binom{k_1}{h_p} A_n(k_1 - h_p, h_p, h_{2p}, \cdots) \pmod{p} .$$

If we make use of (1.4) we obtain the following simpler

THEOREM 2. *Let*

$$h_1 + 2h_2 + 3h_3 + \cdots = p^r .$$

Then the coefficient $A_{p^r}(h_1, h_2, h_3, \cdots)$ *is divisible by* p *except when*

$$h_i = 0 \qquad (i \neq j) , \qquad h_j = p^s ,$$

for some j, *in which case*

$$A_{p^r}(h_1, h_2, h_3, \cdots) \equiv 1 \pmod{p} .$$

Using (2.10) and (2.11) we can obtain additional results. For example take

$$h_1 + 2h_2 + 3h_3 + \cdots = 2p^r .$$

Then $A_{2p^r}(h_1, h_2, h_3, \cdots)$ is divisible by p unless (i) all $h_s = 0$ $(s \neq j)$, $h_j = 1$ or 2; (ii) all $h_s = 0$ $(s \neq i, j)$, $h_i = h_j = 1$. In case (i) $A \equiv 1$, in case (ii) $A \equiv 2 \pmod{p}$.

For $n = 3p^r$ the corresponding results are more complicated.

4. We turn now to the polynomial $C_n(\alpha_1, \alpha_2, \alpha_3, \cdots)$, the cycle indicator of the symmetric group [5, p. 68]:

$$\begin{aligned} (4.1) \qquad C_n = C_n(\alpha_1, \alpha_2, \alpha_3, \cdots) &= \phi_n(\alpha_1, \alpha_2, 2!\alpha_3, \cdots) \\ &= \sum \frac{n!}{k_1!k_2!k_3 \cdots} \left(\frac{\alpha_1}{1}\right)^{k_1}\left(\frac{\alpha_2}{2}\right)^{k_2}\left(\frac{\alpha_3}{3}\right)^{k_3} \cdots , \end{aligned}$$

where the summation is over all nonnegative k_j such that

$$k_1 + 2k_2 + 3k_3 + \cdots = n .$$

It is convenient to define $C_0 = 1$.

Put

$$c_n(k_1, k_2, k_3, \cdots) = \frac{n!}{k_1!k_2!k_3 \cdots 1^{k_1}2^{k_2}3^{k_3} \cdots}, \tag{4.2}$$

the general coefficient of C_n. Clearly $c_n(k_1, k_2, k_3, \cdots)$ is integral and indeed a multiple of $A_n(k_1, k_2, k_3, \cdots)$.

From (4.1) we get the generating function

$$G(t) = \sum_{n=0}^{\infty} G_n \frac{t^n}{n!} = \exp\left(\alpha_1 t + \frac{1}{2}\alpha_2 t^2 + \frac{1}{3}\alpha_3 t^3 + \cdots\right). \tag{4.3}$$

For brevity put

$$C(t) = \sum_{n=1}^{\infty} \frac{1}{n} \alpha_n t^n .$$

Differentiating (4.3) with respect to t we get

$$G'(t) = C'(t)G(t) ,$$

that is

$$\sum_{n=0}^{\infty} C_{n+1} \frac{t^n}{n!} = \sum_{r=0}^{\infty} \alpha_{r+1} t^r \sum_{n=0}^{\infty} C_n \frac{t^n}{n!} .$$

This implies

$$C_{n+1} = \sum_{r=0}^{n} \frac{n!}{r!} \alpha_{n-r+1} C_r , \tag{4.4}$$

so that

$$C_{n+1} \equiv \alpha_1 C_n \pmod{n} . \tag{4.5}$$

By repeated differentiation of (4.3) we get (compare (2.4))

$$\frac{d^p}{dt^p} G(t) \equiv \{(C'(t))^p + C^{(p)}(t)\} G(t) \pmod{p} . \tag{4.6}$$

Now since

$$C'(t) = \sum_{n=0}^{\infty} \alpha_{n+1} t^n , \quad C^{(p)}(t) = \sum_{n=0}^{\infty} (n + p - 1)! \alpha_{n+1} \frac{t^n}{n!} ,$$

it is clear that

$$(C'(t))^p \equiv \alpha_1^p, \quad C^{(p)}(t) \equiv -\alpha_p \pmod{p} ;$$

at the last step we have used Wilson's theorem. Thus (4.6) becomes

$$\sum_{n=0}^{\infty} C_{n+p}\frac{t^n}{n!} \equiv (\alpha_1^p - \alpha_p)\sum_{n=0}^{\infty} C_n\frac{t^n}{n!}\ ,$$

so that

$$C_{n+p} \equiv (\alpha_1^p - \alpha_p)C_n \pmod{p}\ . \tag{4.7}$$

In particular we have

$$C_p \equiv \alpha_1^p - \alpha_p \pmod{p} \tag{4.8}$$

and

$$C_{n+rp} \equiv (\alpha_1^p - \alpha_p)^r C_n \pmod{p} \tag{4.9}$$

We remark that for $p = 3, 5, 7$, (4.8) is in agreement with the explicit values of C_n given in [5, p. 69].

By (4.9) with $n = 0$ we find that the coefficient

$$c_{rp}(k_1, k_2, k_3, \cdots) \equiv 0 \pmod{p}$$

unless all k_j except k_1 and k_p vanish and k_1 is a multiple of p; in this case we have

$$c_{rp}(pk, 0, \cdots, 0, k_p, \cdots) \equiv (-1)^{k_p}\binom{r}{k} \pmod{p}\ . \tag{4.10}$$

References

1. P. Bachmann, *Niedere Zahlentheorie*, vol. 1, Leipzig, 1902.
2. E. T. Bell, *Exponential polynomials*, Annals of Math., **35** (1934), 258-277.
3. L. E. Dickson, *Theorems on the residues of multinomial coefficients with respect to a prime modulus*, Quarterly J. Math., **33** (1902), 378-384.
4. A. Hurwitz, *Über die Entwickelungscoeffizienten der lemniscatischen Functionen*, Mathematische Werke, Basel, **2**, (1933), 342-373.
5. J. Riordan, *An Introduction to combinatorial analysis*, New York, 1958.
6. J. Touchard, *Propriétés arithmétiques de certains nombres recurrents*, Ann. Soc. Sci. Bruxelles, **A53** (1933), 21-31.

Duke University

EXTENSIONS OF HOMOMORPHISMS

PAUL CIVIN

1. Introduction. A multiplication was introduced by R. Arens [1] [2] into the second conjugate space B^{**} of a Banach algebra, B, which made B^{**} into a Banach algebra. The algebra of the second conjugate space was studied by Civin and Yood [3], with particular attention given to the case where B was $L(\mathfrak{G})$, the group algebra of the locally compact abelian group $\mathfrak{G}$. Among the results they noted was that the algebra $M(\mathfrak{G})$ of finite regular Borel measures on $\mathfrak{G}$ was isomorphic as an algebra with a quotient algebra of $L^{**}(\mathfrak{G})$. With $\mathfrak{H}$ also a locally compact abelian group, P. J. Cohen showed [4, p. 220] that any homomorphism of $L(\mathfrak{G})$ into $M(\mathfrak{H})$ has an extension which was a homomorphism of $M(\mathfrak{G})$ into $M(\mathfrak{H})$.

In §3 we discuss the extensions of homomorphisms defined on a Banach algebra A into either the second conjugate algebra B^{**} of a Banach algebra B or certain of its quotient algebras. The result of Cohen quoted above is included in Theorem 3.7 when $\mathfrak{G}$ and $\mathfrak{H}$ are compact groups. In §4 we indicate, for compact $\mathfrak{H}$, a class of homomorphisms from $L(\mathfrak{G})$ into $M(\mathfrak{H})$, which are induced by homomorphisms of $L(\mathfrak{G})$ into $L^{**}(\mathfrak{H})$.

2. Notation. The notation of Civin and Yood [3] is used throughout. If A is a Banach algebra, A^*, A^{**}, $\cdots$ denote the various conjugate spaces of A. For $f \in A^*$, $x \in A$, $\langle f, x\rangle \in A^*$ is defined by $\langle f, x\rangle(y) = f(xy)$, $y \in A$. For $F \in A^{**}$, $f \in A^*$, $[F, f] \in A^*$ is defined by $[F, f](x) = F(\langle f, x\rangle)$, $x \in A$. Also for $F \in A^{**}$, $G \in A^{**}$ the multiplication FG is defined in A^{**} by $FG(f) = F([G, f])$, $f \in A^*$.

For some purposes, Arens [2] also considers a second multiplication $F\cdot G$ defined for F and G in A^{**} in a manner similar to the above, except that at the first stage, $\langle f | x\rangle \in A^*$ is defined by $\langle f | x\rangle(y) = f(yx)$, $f \in A^*$, $x, y \in A$. Arens calls the multiplication in A *regular* provided that $F\cdot G = GF$ for all $F, G \in A^{**}$. Clearly, if A is commutative, then A^{**} is commutative if and only if the multiplication in A is regular. The same notation as above, in terms of bilinear functionals, is used in the sequel with respect to a multiplication in A^{****} which comes from the first of the above multiplications in A^{**}.

If π is the natural mapping of A into A^{**}, we say that a mapping φ defined on A^{**} into a set $\mathfrak{S}$ is an extension of a mapping ρ defined on A into $\mathfrak{S}$ if $\varphi(\pi x) = \rho(x)$ for $x \in A$.

For any subset $\mathfrak{J}$ in A^*, we use the notation $\mathfrak{J}^{\perp}$ for $\{F \in A^{**} | F(f) = 0, f \in \mathfrak{J}\}$.

Received December 12, 1960. This research was supported by the National Science Foundation, grant NSF-G-14, 111.

For a commutative Banach algebra A, we let $\mathfrak{Y}(A)$ denote the closed subspace of A^* generated by the multiplicative linear functionals. If $A = L(\mathfrak{G})$, the group algebra of the locally compact group $\mathfrak{G}$, we write $\mathfrak{Y}(\mathfrak{G})$ in place of $\mathfrak{Y}(L(\mathfrak{G}))$.

3. **Extension of homomorphisms.** We first consider the possibility of extending a bounded homomorphism of the Banach algebra A into the Banach algebra B^{**} to a w^*-continuous homomorphism of A^{**} into B^{**}. Throughout this section we adopt the notation π for the natural mapping of A into A^{**} and σ for the natural mapping of B^* into B^{***}.

3.1 THEOREM. *Let A and B be Banach algebras. Let φ be a bounded homomorphism of A into the center of B^{**}. Then there is a unique w^*-continuous homomorphism ψ of A^{**} into B^{**} which is the extension of φ.*

Proof. Let $f \in B^{**}$, and $x, y \in A$. Then $\langle\varphi^*\sigma f, x\rangle(y) = \varphi^*\sigma f(xy) = \varphi(xy)(f) = \varphi(y)\varphi(x)(f) = \varphi(y)([\varphi(x), f]) = \varphi^*\sigma[\varphi(x), f](y)$. Thus $\langle\varphi^*\sigma f, x\rangle = \varphi^*\sigma[\varphi(x), f]$. For any $G \in A^{**}$, $[G, \varphi^*\sigma f](x)\ G(\langle\varphi^*\sigma f, x\rangle) = G(\varphi^*\sigma[\varphi(x), f] = \sigma^*\varphi^{**}G([\varphi(x), f]) = \sigma^*\varphi^{**}G\varphi(x)(f) = \varphi(x)\sigma^*\varphi^{**}G(f) = \varphi(x)([\sigma^*\varphi^{**}G, f]) = \varphi^*\sigma[\sigma^*\varphi^{**}G, f](x)$. Consequently, $[G, \varphi^*\sigma f] = \varphi^*\sigma[\sigma^*\varphi^{**}G, f]$. Therefore for any $F \in A^{**}$, $F([G, \varphi^*\sigma f]) = F(\varphi^*\sigma[\sigma^*\varphi^{**}G, f]) = \sigma^*\varphi^{**}F([\sigma^*\varphi^{**}G, f])$. Hence $\sigma^*\varphi^{**}(FG)(f) = FG(\varphi^*\sigma f) = F([G, \varphi^*\sigma f]) = \sigma^*\varphi^{**}F([\sigma^*\varphi^{**}G, f]) = \sigma^*\varphi^{**}F\sigma^*\varphi^{**}G(f)$. Thus $\sigma^*\varphi^{**}$ is a homomorphism of A^{**} into B^{**}.

For $x \in A$, and $f \in B^*$, $\sigma^*\varphi^{**}(\pi x)(f) = \pi x(\varphi^*\sigma f) = \varphi^*\sigma f(x) = \sigma f(\varphi(x)) = \varphi(x)(f)$. Thus $\sigma^*\varphi^{**}(\pi x) = \varphi(x)$ and $\sigma^*\varphi^{**}$ is an extension of φ.

Let $G \in A^{**}$, $G_\alpha \in A^{**}$ and suppose $G = w^*\text{-}\lim G_\alpha$. Then for any $f \in B^*$, $\lim \sigma^*\varphi^{**}G_\alpha(f) = \lim G_\alpha(\varphi^*\sigma f) = \sigma^*\varphi^{**}G(f)$, and so $\sigma^*\varphi^{**}$ is w^*-continuous.

The assertion of uniqueness follows from the following.

3.2 LEMMA. *Let A and B be Banach algebras, and let φ be any bounded linear transformation of A into B^{**}. Then $\sigma^*\varphi^{**}$ is the only w^*-continuous extension of φ to a transformation of A^{**} into B^{**}.*

Proof. That $\sigma^*\varphi^{**}$ is a w-continuous extension was given above. Suppose that ψ is a w^*-continuous extension of φ, so that $\psi(\pi x) = \varphi(x)$ for all $x \in A$. Let $G \in A^{**}$ and let $\{x_\alpha\}$ be a net in A such that w^*-lim $\pi x_\alpha = G$. Then for $f \in B^*$, $\psi(G)(f) = \lim\ \psi(\pi x_\alpha)f = \lim\ \varphi(x_\alpha)(f) = \lim\ \varphi^*\sigma f(x_\alpha) = \lim\ \pi x_\alpha(\varphi^*\sigma f) = G(\varphi^*\sigma f) = \sigma^*\varphi^{**}G(f)$. Hence $\psi(G) = \sigma^*\varphi^{**}G$.

If B is commutative with a regular multiplication, an alternative proof of Theorem 3.1 may be given on the basis of the following lemma and Theorem 6.1 of [3].

3.3 LEMMA. *If B is a commutative Banach algebra with a regular multiplication then σ^* is a homomorphism of B^{****} into B^{**}.*

Proof. Since multiplication in B is regular, B^{**} is [2] a commutative algebra. Let $U, V \in B^{****}$. For $f \in B^*$, and $F, G \in B^{**}$, $\langle\sigma f, F\rangle(G) = \sigma f(FG) = FG(f) = GF(f) = G([F,f]) = \sigma[F,f](G)$, and therefore $\langle\sigma f, F\rangle = \sigma[F,f]$. Also $[V, \sigma f](F) = V(\langle\sigma f, F\rangle) = V(\sigma[F,f]) = \sigma^* V[F,f] = (\sigma^* V)F(f) = F\sigma^* V(f) = F([\sigma^* V, f]) = \sigma[\sigma^* V, f](F)$. Thus $[V, \sigma f] = \sigma[\sigma^* V, f]$. Consequently $\sigma^*(UV)(f) = UV(\sigma f) = U([V, \sigma f]) = U(\sigma[\sigma^* V, f]) = \sigma^* U([\sigma^* V, f]) = \sigma^* U\sigma^* V(f)$ and σ^* is a homomorphism a claimed.

We note that it is impossible in general to conclude that the range of the extension of φ is in the center of B^{**} even though the range of φ is in the center. For let $A = B$ be a commutative algebra whose multiplication is not regular, and let $\varphi = \pi$. Then the w^*-continuous extension of π is the identity map and B^{**} is not commutative.

One further example is in order, to see that in general a bounded homomorphim φ from A into B^{**} does not admit a w^*-continuous extension as a homomorphism from A^{**} into B^{**}. For this purpose let A be the group algebra of the integers, $\mathfrak{G}$, and let $B = A$. Let t_γ, $\gamma \in \mathfrak{G}$ be the translation operator on A^*, defined by $t_\gamma f(\alpha) = f(\alpha + \gamma)$, $f \in A^*$, and α, $\gamma \in \mathfrak{G}$. Let $e \in A^*$ correspond to the function identically one on $\mathfrak{G}$. Let $\mathfrak{J} = \{F \in A^{**} \mid F(t_\gamma f) = F(f), \text{ for all } \gamma \in \mathfrak{G}, f \in A^*\}$. Then as noted in formula (3.2) of [3],

$$GF = G(e)F, \quad F \in \mathfrak{J}, \quad G \in A^{**}. \tag{3.1}$$

In particular any $F \in \mathfrak{J}$ with $F(e) = 1$ is an idempotent. As noted in [3], $\mathfrak{J}$ is a two sided ideal in A^{**} with only zero in common with the center of A^{**}. Since $\mathfrak{G}$ is a discrete group A has an identity and thus [3, Lemma 5.4] A^{**} has an identity E. Let F be a nonzero idempotent in $\mathfrak{J}$. Thus $E - F$ is also an idempotent. Let $\varphi(x) = \pi x(E - F)$. Since πA is in the center of A^{**}, $\varphi(x)$ is a homomorphism of A into A^{**}. If φ had a w^*-continuous extension as a homomorphism, the extension ψ would have the value $\psi(G) = G(E - F)$, $G \in A^{**}$. We now show that ψ is not a homomorphism. As noted above F is not in the center of A^{**}, so we may pick $H \in A^{**}$ such that $HF \neq FH$. Also pick $G \in A^{**}$ such that $G(e) = 1$. Then $\psi(GH) = GH(E - F) = GH - GHF = GH - (GH)(e)F$. Now e is a multiplicative linear functional on A, and so by Lemma 3.6 of [3], $(GH)(e) = G(e)H(e) = H(e)$. Thus $\psi(GH) = GH - H(e)F = GH - HF$. On the other hand $\psi(G)\psi(H) = (G - GF)(H - HF) = (G - F)(H - H(e)F) = GH - FH - H(e)GF + H(e)F = GH - FH$. Since $FH \neq HF$, $\psi(GH) \neq \psi(G)\psi(H)$ and ψ is not a homomorphism.

Before turning to other types of extensions we note one further item on the matter of w^*-continuity of homomorphisms.

3.4 LEMMA. *If A and B are Banach algebras and ψ is a bounded homomorphism of A^{**} into the center of B^{**}, then there is a w^*-continuous homomorphism ρ of A^{**} into B^{**} such that $\psi(\pi x) = \rho(\pi x)$ for $x \in A$.*

Proof. Since $\psi\pi$ is a homomorphism of A into the center of B^{**}, we may take $\rho = \sigma^*\psi^{**}\pi^{**}$ and apply Theorem 3.1.

Homomorphisms of A^{**} into B^{**} which are not w^*-continuous exist, as may be seen in the following example. Let $\mathfrak{G}$ be an infinite compact group and let $A = B$ be the group algebra of $\mathfrak{G}$. Then by Lemma 3.8 of [3], A^{**} has a right identity E which is not an identity. Define for $F \in A^{**}$, $\psi(F) = EF$. Then $\psi(FG) = EFG = EFEG = \psi(F)\psi(G)$. However ψ although bounded is not w^*-continuous. For let $G \in A^{**}$ and let $\{x_\alpha\}$ be a net such that $w^* - \lim \pi x_\alpha = G$. Then if ψ were w^*-continuous we would have $\psi(G) = \lim \psi(\pi x_\alpha) = \lim E\pi x_\alpha = \lim \pi x_\alpha = G$. However, $\psi(G) = EG$ and $EG \neq G$ for some $G \in A^{**}$.

We next turn to the question of extending homomorphisms from A into certain quotient algebras of B^{**} in the case in which both A and B are commutative. We must first characterize the w^*-closed ideals of a second conjugate algebra.

3.5 LEMMA. *Let A be a commutative Banach algebra. Let $\mathfrak{J}$ be a w^*-closed subspace of A^{**} and let $\mathfrak{J}_0 = \{f \in A^* \mid F(f) = 0, F \in \mathfrak{J}\}$. Then $\mathfrak{J}$ is an ideal of A^{**} if and only if $[G, f] \in \mathfrak{J}_0$ for all $G \in A^{**}, f \in \mathfrak{J}_0$.*

Proof. Since $\mathfrak{J}$ is w^*-closed, $\mathfrak{J} = \mathfrak{J}_0{}^\perp$. Suppose $\mathfrak{J}$ is an ideal of A^{**}. For any $F \in \mathfrak{J}, G \in A^{**}$, and $f \in \mathfrak{J}_0$, $FG \in \mathfrak{J}$ and $FG(f) = 0$. Therefore $F([G, f]) = 0$ for all $F \in \mathfrak{J}$, and so by definition $[G, f] \in \mathfrak{J}_0$. Suppose next that the stated condition holds. Let $F \in \mathfrak{J}$ and $G \in A^{**}$. For any $f \in \mathfrak{J}_0$, $[G, f] \in \mathfrak{J}_0$ and thus $FG(f) = F([G, f]) = 0$. Consequently $FG \in \mathfrak{J}_0{}^\perp = \mathfrak{J}$ and $\mathfrak{J}$ is a right ideal. For any $x \in A, \pi x$ is in the center of A^{**}, hence if $F \in \mathfrak{J}$, $\pi x F = F\pi x \in \mathfrak{J}$. Since πA is w^*-dense in A^{**} and left multiplication is w^*-continuous [2], we see that $GF \in \mathfrak{J}$ for any $G \in A^{**}$, and thus $\mathfrak{J}$ is an ideal of A^{**}.

3.6 THEOREM. *Let A and B be commutative Banach algebras. Let $\mathfrak{J}$ be a w^*-closed ideal of B^{**}. Suppose that φ is a bounded homomorphism of A into the center of $B^{**}/\mathfrak{J}$. Then there exists a w^*-closed ideal $\mathfrak{J}'$ of A^{**} and a homomorphism ψ of $A^{**}/\mathfrak{J}'$ into $B^{**}/\mathfrak{J}$ such that if π is the natural embedding of A into A^{**}, then $\psi(\pi x + \mathfrak{J}') = \varphi(x), x \in A$.*

Proof. Since $\mathfrak{I}$ is w^*-closed, $\mathfrak{I} = \mathfrak{I}_0{}^{\perp}$ where $\mathfrak{I}_0 = \{f \in B^* \mid F(f) = 0$ for all $F \in \mathfrak{I}\}$. Let β be the linear space isometric isomorphism of $\mathfrak{I}_0{}^*$ onto $B^{**}/\mathfrak{I}$ defined for $F_0 \in \mathfrak{I}_0{}^*$ by $\beta F_0 = F + \mathfrak{I}$ where $F \in B^{**}$ is an arbitrary extension of F_0. Define multiplication in $\mathfrak{I}_0{}^*$ so that β (and thus β^{-1}) is an algebra isomorphism. For $f \in \mathfrak{I}_0$, define $\varphi_* f$ by $\varphi_* f(x) = (\beta^{-1}\varphi(x))(f)$, $x \in A$. Then $\varphi_* f$ is linear and since φ is bounded $\|\varphi_* f(x)\| \leqq \|\varphi\| \, \|x\| \, \|f\|$, and $\varphi_* f \in A^*$.

Let $\mathfrak{I}_0'$ be the w^*-closure of the range of φ_*, and let $\mathfrak{I}' = \mathfrak{I}'_0{}^{\perp}$. Clearly $\mathfrak{I}'$ is w^*-closed. We next show that $\mathfrak{I}'$ is an ideal of A^{**}. Let $f \in \mathfrak{I}_0$. Then for any $x, y \in A$, $\langle \varphi_* f, x\rangle(y) = \varphi_* f(xy) = (\beta^{-1}\varphi(xy))f = (\beta^{-1}\varphi(yx))(f)$, since the range of φ is commutative. Suppose that $\varphi(y) = U + \mathfrak{I}$, and $\varphi(x) = V + \mathfrak{I}$ so that $\varphi(yx) = UV + \mathfrak{I}$. Then $(\beta^{-1}\varphi(xy))(f) = UV(f) = U([V, f])$. Since $f \in \mathfrak{I}_0$, and $\mathfrak{I} = \mathfrak{I}_0{}^{\perp}$ is an ideal, $g = [V, f] \in \mathfrak{I}_0$ by Lemma 3.5. Hence $(\beta^{-1}\varphi(yx))(f) = U(g) = (\beta^{-1}\varphi(y))(g) = \varphi_* g(y)$, for all $y \in A$. We therefore have $\langle \varphi_* f, x\rangle = \varphi_* g$ and so $\langle \varphi_* f, x\rangle \in \mathfrak{I}'_0$ for any $x \in A$ and $f \in \mathfrak{I}_0$. Suppose next that $g \in \mathfrak{I}_0'$, and $x \in A$. Say $g = w^*\text{-}\lim \varphi_* f_\alpha$ with $f_\alpha \in \mathfrak{I}_0$. Then for $y \in A$, $\langle g, x\rangle(y) = g(xy) = \lim \varphi_* f_\alpha(xy) = \lim \langle \varphi_* f_\alpha, x\rangle(y)$, and hence $\langle g, x\rangle = w^* - \lim \langle \varphi_* f_\alpha, x\rangle$. However, by the above, $\langle \varphi_* f_\alpha, x\rangle \in \mathfrak{I}_0'$, and $\mathfrak{I}_0'$ is w^*-closed so $\langle g, x\rangle \in \mathfrak{I}_0'$ for any $g \in \mathfrak{I}_0'$ and $x \in A$.

Let $G \in A^{**}$ and let $f \in \mathfrak{I}_0'$. Let $\{x_\alpha\}$ be a net in A such that $w^*\text{-}\lim \pi x_\alpha = G$. Then $[G, f](x) = G(\langle f, x\rangle) = \lim \pi x_\alpha(\langle f, x\rangle) = \lim \langle f, x\rangle(x_\alpha) = \lim f(x x_\alpha) = \lim f\langle x_\alpha x) = \lim \langle f, x_\alpha\rangle(x)$ for $x \in A$. Consequently $[G, f] = w^*\text{-}\lim \langle f, x_\alpha\rangle$, and is thus in $\mathfrak{I}_0'$ as $\mathfrak{I}_0'$ is w^*-closed. Hence, by Lemma 3.5, $\mathfrak{I}' = \mathfrak{I}_0'{}^{\perp}$ is a w^*closed ideal of A^{**}.

For $F \in A^{**}$, define $\gamma F(f) = F(\varphi_* f)$ for $f \in \mathfrak{I}_0$. Clearly γF is a bounded linear functional on $\mathfrak{I}_0$, and so has an extension of the same norm which is an element of B^{**}. We again denote the extension by γF. Thus γ is a bounded linear map from A^{**} into B^{**}. Note that if $F_1 - F_2 \in \mathfrak{I}'$ and $f \in \mathfrak{I}_0$, then $\gamma(F_1 - F_2)(f) = (F_1 - F_2)(\varphi_* f)\ 0$, and thus $\gamma F_1 - \gamma F_2 \in \mathfrak{I}$. Thus for any $F \in F_0 + \mathfrak{I}$, $\|\gamma F_0 + \mathfrak{I}'\| = \|\gamma F + \mathfrak{I}\| \leqq \|\gamma F\| \leqq \|F\| \, \|\varphi_*\|$ and hence $\|\gamma F_0 + \mathfrak{I}\| \leqq \|F_0 + \mathfrak{I}'\| \, \|\varphi_*\|$.

Define ψ on $A^{**}/\mathfrak{I}'$ by $\psi(F + \mathfrak{I}') = \gamma F + \mathfrak{I}$. By the above, we see that ψ is a bounded linear mapping of $A^{**}/\mathfrak{I}'$ into $B^{**}/\mathfrak{I}$. Also for $x \in A$, $\psi(\pi x + \mathfrak{I}') = \gamma \pi x + \mathfrak{I}$. Since $\gamma \pi x(f) = \pi x(\varphi_* f) = \varphi_* f(x) = (\beta^{-1}\varphi(x))(f)$ for $f \in \mathfrak{I}_0$, $\gamma \pi x - \beta^{-1}\varphi(x) \in \mathfrak{I}$, and $\psi(\pi x + \mathfrak{I})' = \varphi(x)$.

Thus all that remains is to see that ψ satisfies the required multiplicative property of a homomorphism. Let $F, G \in A^{**}$. To see that $\psi(FG) = \psi(F)\psi(G)$, we must show that for $f \in \mathfrak{I}_0$, $\{\gamma(F)\gamma(G) - \gamma(FG)\}(f) = 0$. Since $\{\gamma(F)\gamma(G) - \gamma(FG)\}(f) = \gamma(F)([\gamma(G), f]) - FG(\varphi_* f) = F(\varphi_*[\gamma(G), f] - [G, \varphi_* f])$, it suffices if we show that $\varphi_*[\gamma(G), f] - [G, \varphi_* f] = 0$. Let $x, y \in A$ and suppose that $\varphi(x) = U + \mathfrak{I}$, $\varphi(y) = V + \mathfrak{I}$, and thus $\varphi(xy) = \varphi(yx) = VU + \mathfrak{I}$. It follows that $\langle \varphi_* f, x\rangle(y) = \varphi_* f(xy) = VU(f) = V([U, f])$. Now, since $f \in \mathfrak{I}_0$, $[U, f] \in \mathfrak{I}_0$ by Lemma 3.5. We therefore have $\langle \varphi_* f, x\rangle(y) =$

$\varphi_*[U,f](y)$ for all $y \in A$, and consequently $\langle \varphi_* f, x\rangle = \varphi_*[U,f]$. Thus $[G, \varphi_* f](x) = G(\langle \varphi_* f, x\rangle) = G(\varphi_*[U,f]) = \gamma G([U,f] = (\gamma G)U(f)$. On the other hand, $\varphi_*[\gamma G, f](x) = U([\gamma G, f]) = U\gamma G(f)$. Since under our hypothesis $\varphi(x) = U + \mathfrak{J}$ is in the center of $B^{**}/\mathfrak{J}$, $U\gamma G(f) = (\gamma G)U(f)$ for $f \in \mathfrak{J}_0$ and we have the desired result.

It should be noted that the ideal $\mathfrak{J}'$ in general is dependent on the homomorphism φ. Two instances should be noted where this is not the case. The first, when $\mathfrak{J}' = 0$, has already been treated in the discussion of w^*-continuous extensions of homomorphisms of A into the center of B^{**}. The other is the following.

3.7 THEOREM. *Let A and B be commutative Banach algebras. Let φ be a homomorphism of A into $B^{**}/\mathfrak{Y}^{\perp}(B)$. Then there is a homomorphism ψ of $A^{**}/\mathfrak{Y}^{\perp}(B)$ such that $\psi(\pi x + \mathfrak{Y}^{\perp}) = \varphi(x)$.*

Proof. If in the proof of Theorem 3.6, $\mathfrak{J}_0 = \mathfrak{Y}(B)$, it follows from Lemma 3.6 of [3] that for any $f \in \mathfrak{J}_0$ which is a multiplicative linear functional on B, that $\varphi_* f$ is a multiplicative linear functional on A. Hence, the norm closure of the range of φ_* is contained in $\mathfrak{Y}(A)$. In view of Lemma 3.6 of [3], the subspace $\mathfrak{Y}^{\perp}(A)$ is a w^*-closed ideal of A^{**}, and if used in the role of $\mathfrak{J}'$ affords the same conclusion. Note that the homomorphism φ is not postulated to be bounded or with range in the center of $B^{**}/\mathfrak{Y}^{\perp}(B)$. This is legitimate since in view of Theorem 3.7 of [3], $B^{**}/\mathfrak{Y}^{\perp}$ is automatically commutative and semi-simple, and thus φ is automatically bounded.

If A and B are the group algebras of the compact groups $\mathfrak{G}$ and $\mathfrak{H}$, then $A^{**}/\mathfrak{Y}^{\perp}(A)$ and $B^{**}/\mathfrak{Y}^{\perp}(B)$ may be identified with the measure algebras $M(\mathfrak{G})$ and $M(\mathfrak{H})$ respectively by Theorem 3.18 of [3]. Thus Theorem 3.7 includes in the case of compact groups, the result of P. J. Cohen [4] quoted in the introduction.

4. Group algebras. Let $\mathfrak{G}$ be a locally compact abelian group. As in §3, we denote the group algebra of $\mathfrak{G}$ by $L(\mathfrak{G})$ and the algebra of finite regular Borel measures on $\mathfrak{G}$ by $M(\mathfrak{G})$. For notational purposes, it is also convenient to identify the character group $\hat{\mathfrak{G}}$ of $\mathfrak{G}$ with the subset of $L^*(\mathfrak{G})$ consisting of the nonzero multiplicttive linear functional on $L(\mathfrak{G})$. The topology of $\hat{\mathfrak{G}}$ is then in agreement with the w^*-topology of $\hat{\mathfrak{G}}$ as a subset of $L^*(\mathfrak{G})$.

Suppose that $\mathfrak{H}$ is a locally compact abelian group. A continuous homomorphism ν of $\mathfrak{G}$ into $\mathfrak{H}$ is called *nonsingular* if for every Borel set E is $\mathfrak{H}$ with zero Haar measure, $\nu^{-1}(\mathfrak{E})$ is of zero Haar measure in $\mathfrak{G}$.

A complete characterization of all homomorphisms φ of $L(\mathfrak{G})$ into $M(\mathfrak{H})$ was given by P. J. Cohen [4]. He utilized the function φ_* from $\hat{\mathfrak{H}}$ into $\{\hat{\mathfrak{G}}, 0\}$ defined by $\varphi_* f(x) = \varphi(x)(f)$, $x \in L(\mathfrak{G})$, $f \in \hat{\mathfrak{H}}$.

4.1 THEOREM. (P. J. Cohen) *Let $\mathfrak{G}$ and $\mathfrak{H}$ be locally compact abelian groups, φ a homomorphism of $L(\mathfrak{G})$ into $M(\mathfrak{H})$, φ_* the induced map of $\hat{\mathfrak{H}}$ into, $\{\hat{\mathfrak{G}}, 0\}$. Then there are a finite number of sets $\mathfrak{K}_i$, which are cosets of open subgroups of $\hat{\mathfrak{H}}$, and continuous maps ψ_i: $\hat{\mathfrak{K}}_i \to \mathfrak{G}$, such that*

$$\psi_i(x + y - z) = \psi_i(x) + \psi_i(y) - \psi_i(z) \tag{4.1}$$

for all x, y and z in $\mathfrak{K}_i$, with the following property: There is a decomposition of $\hat{\mathfrak{H}}$ into the disjoint union of sets $\mathfrak{S}_j$, each lying in the Boolean ring generated by the sets $\mathfrak{K}_i$, such that on each $\mathfrak{S}_j$, φ_* is either identically zero or agrees with some ψ_i, where $\mathfrak{S}_j \subset \mathfrak{K}_i$.

Conversely, for any map of $\hat{\mathfrak{H}}$ into $\{\hat{\mathfrak{G}}, 0\}$, there is a homomorphism of $L(\mathfrak{G})$ into $M(\mathfrak{H})$ which induces it. The map φ carries $L(\mathfrak{G})$ into $L(\mathfrak{H})$ if and only if φ_*^{-1} of every compact subset of $\hat{\mathfrak{G}}$ is compact.

Suppse that the sets $\mathfrak{K}_i$ are cosets of the subgroups $\mathfrak{U}_i$ of $\hat{\mathfrak{H}}$. There is a closed subgroup $\mathfrak{H}_i$ of $\mathfrak{H}$, $\mathfrak{H}_i = \{h \in \mathfrak{H} \mid (h, \hat{h}) = 1, \hat{h} \in \mathfrak{U}_i\}$, such that $\mathfrak{U}_i$ may be viewed [6, p. 130] as the character group of $\mathfrak{H}/\mathfrak{H}_i$. Let $a_i \in \mathfrak{K}_i$, and define ψ_i': $\mathfrak{U}_i \to \mathfrak{G}$ by

$$\psi_i'(x) = \psi_i(a_i + x) - \psi_i(a_i), \qquad x \in \mathfrak{U}_i. \tag{4.2}$$

The condition (4.1) on ψ_i is then equivalent to the assertion that ψ_i' is a homomorphism of $\mathfrak{U}_i$ into $\hat{\mathfrak{G}}$, and ψ_i' is continuous along with ψ_i. We may also consider the dual homomorhism ρ_i: $\mathfrak{G} \to \hat{\mathfrak{U}}_i = \mathfrak{H}/\mathfrak{H}_i$, defined dy

$$(\psi_i'(x), g) = (x, \beta_i(g)), \qquad x \in \mathfrak{U}_i = (\mathfrak{H}/\mathfrak{H}_i)\hat{}, g \in \mathfrak{G}. \tag{4.3}$$

In view of the Cohen theorem, the homomorphism ψ is determined by the sets $\mathfrak{K}_i, \mathfrak{S}_j$ and the functions β_i. The notation introduced above will be used in the sequel without further comment. We also use the notation ρ_* as the mapping of $L^*(\mathfrak{H})$ into $L^*(\mathfrak{G})$ which is defined by $\rho_* f(x) = \rho(x)(f)$, $x \in L(\mathfrak{G})$, $f \in L^*(\mathfrak{H})$, whenever ρ is a bounded linear map of $L(\mathfrak{G})$ into $L^{**}(\mathfrak{H})$.

4.2 LEMMA. *Let λ be a nonsingular homomorphism of $\mathfrak{G}$ into a locally compact abelian group $\mathfrak{K}$. Then λ induces a homomorphism ρ of $L(\mathfrak{G})$ into $L^{**}(\mathfrak{K})$ such that for $f \in \hat{\mathfrak{K}}$, $\rho_*(f) = f \circ \lambda$.*

Proof. For $k \in L^*(\mathfrak{K})$, define $\lambda_*(k)$ by

$$\lambda_*(k)(\alpha) = k \circ \lambda(\alpha), \qquad \alpha \in G.$$

We first must show that λ_* is a well-defined bounded linear mapping of $L^*(\mathfrak{K})$ into $L^*(\mathfrak{G})$. Suppose that K_1 andK_2 are two bounded Borel measurable functions on $\mathfrak{K}$ such that $k_1(\beta) = k_2(\beta)$ for almost all β in $\mathfrak{K}$. Let $\mathfrak{E} = \{\alpha \in \mathfrak{G} \mid k_1(\lambda(\alpha)) \neq k_2(\lambda(\alpha))\}$. Then $\mathfrak{E} = \lambda^{-1}(\lambda(\mathfrak{E}))$ and by the hypothesis of non-singularity $\mathfrak{E}$ has measure zero in $\mathfrak{G}$. Since it is now immediate that $|\lambda_*(k)(\alpha)| \leqq \|k\|$ for almost all α in $\mathfrak{G}$, it follows that λ_* is a bounded linear map of $L^*(\mathfrak{K})$ into $L^*(\mathfrak{G})$.

For $x \in L(\mathfrak{G})$, define $\rho(x)$ on $L^*(\mathfrak{K})$ by

$$\rho(x)(f) = \lambda_* f(x), \qquad f \in L^*(\mathfrak{K}).$$

Clearly $\rho(x) \in L^{**}(\mathfrak{K})$, and ρ is a bounded linear mapping from $L(\mathfrak{G})$ into $L^{**}(\mathfrak{K})$, and $\rho_* f = f \circ \lambda$.

We next show that ρ satisfies the multiplicative condition for a homomorphism. Let $x, y \in L(\mathfrak{G})$ and $f \in L^*(\mathfrak{K})$. Then

$$\begin{aligned} \rho(xy)(f) = \lambda_* f(xy) &= \int_{\mathfrak{G}} \lambda_* f(\alpha) \int_{\mathfrak{G}} x(\beta) y(\alpha - \beta)\, d\beta\, d\alpha \\ &= \int_{\mathfrak{G}} \int_{\mathfrak{G}} f(\lambda)(\alpha)) x(\beta) y(\alpha - \beta)\, d\beta\, d\alpha \\ &= \int_{\mathfrak{G}} \int_{\mathfrak{G}} f(\lambda(\alpha + \beta)) x(\beta) y(\alpha)\, d\beta\, d\alpha. \end{aligned}$$

For any $z \in L(\mathfrak{K})$, and $\delta \in \mathfrak{K}$, it is easily seen [3] that $\langle f, z\rangle(\delta) = \int_{\mathfrak{K}} f(z + \delta) z(\gamma)\, d\gamma$. Therefore,

$$\begin{aligned} [\rho(y), f](z) = \rho(y)(\langle f, z\rangle) = \lambda_* \langle f, z\rangle(y) &= \int_{\mathfrak{G}} \lambda_* \langle f, z\rangle(\alpha) y(\alpha)\, d\alpha \\ = \int_{\mathfrak{G}} \langle f, z\rangle(\lambda(\alpha)) y(\alpha) d\alpha &= \int_{\mathfrak{G}} \int_{\mathfrak{K}} f(\gamma + \lambda(\alpha)) z(\gamma) y(\alpha)\, d\gamma\, d\alpha. \end{aligned}$$

Since the order of integration may be reversed, we see that for $\gamma \in \mathfrak{K}$, $[\rho(y), f](\gamma) = \int_{\mathfrak{G}} f(\gamma + \lambda(\beta)) y(\beta)\, d\beta$. Hence,

$$\begin{aligned} \rho(x)\rho(y)(f) = \rho(x)([\rho(y), f]) = \lambda_*[\rho(y), f](x) &= \int_{\mathfrak{G}} \lambda_*[\rho(y), f](\alpha) x(\alpha)\, d\alpha \\ = \int_{\mathfrak{G}} [\rho(y), f](\lambda(\alpha)) x(\alpha)\, d\alpha &= \int_{\mathfrak{G}} \int_{\mathfrak{G}} f(\lambda(\alpha) + \lambda(\beta)) y(\beta) x(\alpha)\, d\beta\, d\alpha . \end{aligned}$$

Since we thus have $\rho(xy)(f) = \rho(x)\rho(y)(f)$, for all $f \in L^*(K)$, ρ is a homomorphism.

4.3 THEOREM. *Let $\mathfrak{G}$ and $\mathfrak{H}$ be locally compact abelian groups, with $\mathfrak{H}$ compact. Let φ be a homomorphism of $L(\mathfrak{G})$ into $M(\mathfrak{H})$. Let $M(\mathfrak{H})$ be regarded as $L^{**}(\mathfrak{H})/\mathfrak{Y}^{\perp}(\mathfrak{H})$, and let θ be the natural mapping of $L^{**}(\mathfrak{H})$ onto $L^{**}(\mathfrak{H})/\mathfrak{Y}^{\perp}(\mathfrak{H})$. Then if each homomorphism β_i, determined by φ, is nonsingular, there is a homomorphism ρ of $L(\mathfrak{G})$ into $L^{**}(\mathfrak{H})$ such that $\varphi = \theta \circ \rho$.*

Proof. The justification for considering $M(\mathfrak{H})$ as $L^{**}(\mathfrak{H})/\mathfrak{Y}^{\perp}(\mathfrak{H})$ is

Theorem 3,18 of [3].

If $\varphi_*(f) = 0$ for all $f \in \mathfrak{S}_j$, define ρ_j: $L(\mathfrak{G}) \to L^{**}(\mathfrak{H})$ by $\rho_j(x) = 0$, $x \in L(\mathfrak{G})$.

Suppose that $\mathfrak{S}_j \subset \mathfrak{K}_i \subset \hat{\mathfrak{H}}$, and $\varphi_*(f) = \psi_i(f)$ for $f \in \mathfrak{S}_j$. In view of (4.1), the homomorphism ψ_i' of U_i into $\hat{G}$ may be defined by $\psi_i'(k) = \psi_i(k + k_i) - \psi_i(k_i)$ for an arbitrary $k_i \in \mathfrak{S}_j$. The dual homomorphism β_i of $\mathfrak{G}$ into $\mathfrak{H}/\mathfrak{H}_i$ is by hypothesis nonsingular. Thus by Lemma 4.2, there is a homomorphism ρ_j' of $L(G)$ into $L^{**}(\mathfrak{H}/\mathfrak{H}_i)$ such that $\rho_{j*}'(k) = k \circ \beta_i$, for $k \in (\mathfrak{H}/\mathfrak{H}_i)\hat{} = \mathfrak{U}_i$.

For $f \in L(\mathfrak{H}/\mathfrak{H}_i)$ define $\theta_i(f)$ on $\mathfrak{H}$ by $\theta_i(f)(\beta) = f(\beta + \mathfrak{H}_i)$. Suppose that the Haar measure on $\mathfrak{H}_i$ is normalized so that the measure of $\mathfrak{H}_i$ is one. The formula relating integration on a group with that on a quotient group shows that θ_i is an isometric isomorphism of $L(\mathfrak{H}/\mathfrak{H}_i)$ into $L(\mathfrak{H})$. Thus by Theorem 6.1 of [3], θ_i^{**} is a homomorphism of $L^{**}(\mathfrak{H}/\mathfrak{H}_i)$ into $L^{**}(\mathfrak{H})$. Also for any $u \in L(\mathfrak{H}/\mathfrak{H}_i)$, and $f \in L^*(\mathfrak{H})$,

$$\begin{aligned}\theta_i^* f(u) = f(\theta_i u) &= \int_{\mathfrak{H}} f(\beta)\theta_i(u)(\beta)\, d\beta \\ &= \int_{\mathfrak{H}/\mathfrak{H}_i}\int_{\mathfrak{H}_i} f(\beta + \gamma)\theta_i(u)(\beta + \gamma)\, d\gamma\, d\dot{\beta},\end{aligned}$$

where $d\dot{\beta}$ is the Haar measure on $\mathfrak{H}/\mathfrak{H}_i$. Thus

$$\theta_i^* f(u) = \int_{\mathfrak{H}/\mathfrak{H}_i} u(\dot{\beta}) \int_{\mathfrak{H}_i} f(\beta + \gamma)\, d\gamma\, d\dot{\beta}\ ,$$

and we conclude that $\theta_i^* f(\dot{\beta}) = \int_{\mathfrak{H}_i} f(\beta + \gamma)\, d\gamma$.

It is well known that in a group algebra the pointwise multiplication by a character is an automorphism of the algebra. We next show that the same situation prevails in the second conjugate algebra of a group algebra. Let $\mathfrak{T}$ be a locally compact abelian group and define, for $\eta \in \hat{\mathfrak{T}}$, $\eta \circ g$ and $\eta \circ g$ by pointwise multiplication on $\mathfrak{T}$ if $x \in L(\mathfrak{T})$ and $g \in L^*(\mathfrak{T})$. Define $\eta \circ G(g) = G(\eta \circ g)$ for $G \in L^{**}(\mathfrak{T})$. Clearly the map $G \to \eta \circ G$ is a one-to-one bounded linear map of $L^{**}(\mathfrak{T})$ onto itself. Let $F, G \in L^{**}(\mathfrak{T})$ and $g \in L^*(\mathfrak{T})$. It remains for us to show that $(\eta \circ F)(\eta \circ G)(g) = \eta \circ (FG)(g)$. Since $(\eta \circ F)(\eta \circ G)(g) = \eta \circ F([\eta \circ G, g]) = F(\eta \circ [\eta \circ G, g])$, while $\eta \circ (FG)(g) = FG(\eta \circ g) = F([G, \eta \circ g])$, it suffices if we show that for all $x \in L(\mathfrak{T})$, $\eta \circ [\eta \circ G, g](x) = [G, \eta \circ g](x)$. Now $\eta \circ [\eta \circ G, g](x) = [\eta \circ G, g](\eta \circ x) = \eta \circ G(\langle g, \eta \circ x \rangle) = G(\eta \circ \langle g, \eta \circ x \rangle)$, while $[G, \eta \circ g](x) = G(\langle \eta \circ g, x \rangle)$, so it suffices if we show that for all $y \in L(\mathfrak{T})$, $\eta \circ \langle g, \eta \circ x \rangle(y) = \langle \eta \circ g, x \rangle(y)$. Since $\eta \circ \langle g, \eta \circ x \rangle(y) = g((\eta \circ x)(\eta \circ y)) = g(\eta \circ xy) = \eta \circ g(xy) = \langle \eta \circ g, x \rangle(y)$, the original assertion follows.

Define the mapping ρ_j by

$$\rho_j(x) = k_i^{-1} \circ \theta_i^{**} \rho_j'(\psi_i(k_i) \circ x)\ , \qquad x \in L(\mathfrak{G})\ , \tag{4.4}$$

where the dot at each occurrence indicates multiplication of the appropriate functions. Since $k_i \in \hat{\mathfrak{H}}$, and $\psi_i(k_i) \in \hat{\mathfrak{G}}$, ρ_j is a composite of four homomorphisms and is thus a homomorphism of $L(\mathfrak{G})$ and $L^{**}(\mathfrak{H})$.

Suppose that $f \in \mathfrak{S}_j \subset \mathfrak{K}_i$, so that $\varphi_* f = \psi_i f$. Since $\mathfrak{K}_i$ is a coset of $\mathfrak{U}_i$, there is a $k \in \mathfrak{U}_i$ such that $f = k_i + k$. We use the same notation for k when it is viewed as a member of $(\mathfrak{H}/\mathfrak{H}_i)\hat{}$. For any $x \in L(\mathfrak{G})$, $\rho_{j*}f(x) = \rho_j(x)(f) = k_i^{-1}\circ\theta_i^{**}\rho_j'(\psi_i(k_i)\circ x)(f) = \theta_i^{**}\rho_j'(\psi_i(k_i)\circ x)(k) = \rho_j'(\psi_i(k_i)\circ x)\theta_i^*(k)$. From the formula obtained earlier for θ_i^*, it is immediate that θ_i^* simply transfers k from being viewed as a member of $\mathfrak{U}_i \subset \hat{\mathfrak{H}}$, to being viewed as a member of $(\mathfrak{H}/\mathfrak{H}_i)\hat{} \subset L^*(\mathfrak{H}/\mathfrak{H}_i)$. Thus

$$\begin{aligned}\rho_{j*}f(x) = \rho_j'(\psi_i(k_i)\circ x)(k) &= \int_{\mathfrak{G}} \rho_j'(k)(\alpha)\psi_i(k_i)(\alpha)x(\alpha)\,d\alpha \\ &= \int_{\mathfrak{G}} (k, \beta_i(\alpha))\psi_i(k_i)(\alpha)x(\alpha)\,d\alpha = \int_{\mathfrak{G}} (\psi_i'(k), \alpha)\psi_i(k_i)(\alpha)x(\alpha)\,d\alpha\ ,\end{aligned}$$

by use of (4.3). Thus by use of the definition of ψ_i' in terms of k_i, we have

$$\begin{aligned}\rho_{j*}f(x) &= \int_{\mathfrak{G}} (\psi_i(k + k_i) - \psi_i(k_i), \alpha)(\psi_i(k_i), \alpha)x(\alpha)\,d\alpha \\ &= \int_{\mathfrak{G}} (\psi_i(f), \alpha)x(\alpha)\,d\alpha = \int_{\mathfrak{G}} \varphi_* f(\alpha)x(\alpha)\,d\alpha\ .\end{aligned}$$

We therefore conclude that $\rho_{j*}f(x) = \varphi_* f(x)$ for all $x \in L(\mathfrak{G})$ or that $\rho_{i*}f = \varphi_* f$ for $f \in \mathfrak{S}_i$.

Now, by the Cohen theorem, $\hat{\mathfrak{H}}$ is the disjoint union of the sets $\mathfrak{S}_j$. The characteristic function of $\mathfrak{S}_j$ is then the Fourier transform of an idempotent measure in $M(\mathfrak{H}) = L^{**}(\mathfrak{H})/\mathfrak{Y}^\perp(\mathfrak{H})$. Let F_j be any member of $L^{**}(\mathfrak{H})$ such that θF_j is the Fourier transform of the characteristic function of $\mathfrak{S}_j$. Then $F_j^2 - F_j \in \mathfrak{Y}^\perp(\mathfrak{H})$. Now, Theorem 3.15 of [3] states that $\mathfrak{Y}^\perp(\mathfrak{H})$ is the radical of $L^{**}(\mathfrak{H})$, and therefore Theorem 2.3.9 of [5] yields $E_j \in L^{**}(\mathfrak{H})$ such that $E_j^2 = E_j$ and $\theta E_j = \theta F_j$.

We next show that if $i \neq j$, then $E_i F E_j = 0$ for any $F \in L^{**}(\mathfrak{H})$. Suppose that $f \in \hat{\mathfrak{H}}$, then Lemma 3.6 of [3] yields

$$E_i F E_j(f) = E_i(f)F(f)E_j(f)\ .$$

For $f \in \hat{\mathfrak{H}}$, $E_k(f) = F_k(f) = \chi(\mathfrak{S}_k)(f)$, where $\chi(\mathfrak{S}_k)$ is the characteristic function of $\mathfrak{S}_k$. Thus since S_i and S_j are disjoint $E_i F E_j(f) = 0$. Hence $E_i F E_j \in \mathfrak{Y}^\perp$, the radical of $L^{**}(\mathfrak{H})$. For a compact group $\mathfrak{H}$, the radical is also the right annihilator of $L^{**}(\mathfrak{H})$ by Theorem 3.5 of [3]. Thus since $E_i = E_i^2$, $E_i F E_j = E_i(E_i F E_j) = 0$.

Let ρ be defined on $L(\mathfrak{G})$ by

$$\rho(x) = E_1\rho_1(x)E_1 + \cdots + E_r\rho_r(x)E_r, \qquad x \in L(\mathfrak{G}),$$

where $\hat{\mathfrak{H}} = \mathfrak{S}_1 \cup \cdots \cup \mathfrak{S}_r$. Clearly ρ is a bounded linear transformation of $L(\mathfrak{G})$ into $L^{**}(\mathfrak{H})$, and to see that ρ is a homomorphism it suffices if we show that $E_i\rho_i(xy)E_i = E_i\rho_i(x)E_i\rho_i(y)E_i$. The latter equality is ertablished by an identical argument to that used above to show $E_iFE_j = 0$ for $i \neq j$. Thus ρ is a homomorphism of $L(\mathfrak{G})$ into $L^{**}(\mathfrak{H})$.

To see that $\theta\circ\rho = \varphi$, it suffices if we show that $\varphi_*(f) = (\theta\circ\rho)_*(f)$ for $f \in \hat{\mathfrak{H}}$. Suppose that $f \in \mathfrak{S}_k$. Then for $x \in L(\mathfrak{G})$, $(\theta\circ\rho)_*(f)(x) = \theta\circ\rho(x)(f) = E_k\rho_k(x)E_k(f)$, since $E_i(f) = 0$ if $i \neq k$. Thus $(\theta\circ\rho)_*(f)(x) = \rho_k(x)(f) = \varphi_* f$ as was shown earlier.

References

1. Richard Arens, *Operations in function classes*, Monat. für Math. **55** (1951), 1-19.
2. ———, *The adjoint of a bilinear operation*, Proc. Amer. Math. Soc., **2** (1954), 839-848.
3. P. Civin and B. Yood, *The second conjugate space of a Banach algebra as an algebra*, Pacific J. Math., **11** (1961), 820-847.
4. P. J. Cohen, *On homomorphisms of group algebras*, Amer. J. Math., **82** (1960), 213-226.
5. C. E. Rickart, *General Theory of Banach Algebras*, Van Nostrand, New York, 1960.
6. André Weil, *L' intégration dans les groupes topologique et ses appicationes*, Herman et Cie, Paris, 1953.

UNIVERSITY OF FLORIDA
UNIVERSITY OF OREGON

ASYMPTOTIC DECAY OF SOLUTIONS OF DIFFERENTIAL INEQUALITIES

PAUL J. COHEN AND MILTON LEES†

1. Introduction. Let A be an operator in a Hilbert space H, and let $u(t)$, $0 \leqq t < \infty$ be a strongly continuously differentiable function of t with values in H such that $Au(t)$ is continuous. We say that $u(t)$ has property S if, as $t \to \infty$, it cannot vanish faster than *every* exponential, unless identically zero. A sufficient condition for all solutions of the abstract differential inequality

$$\left\| \frac{du}{dt} - Au \right\| \leqq \phi(t) \, \| u \| \, , \qquad 0 \leqq t < \infty \, , \tag{1.1}$$

to have property S was determined by P. D. Lax [1]. The required condition is that there exists an infinite sequence of lines parallel to the imaginary axis whose abcissae λ_n tend to $-\infty$ and on which the resolvent operator $(A - \lambda)^{-1}$ is uniformly bounded by some constant d^{-1}, and that $\sup \phi(t) < d$.

In this paper we give another sufficient condition for all of the solutions of (1.1) to have property S. We require that the operator A be symmetric, i.e., $(Au, v) = (u, Av)$, for all u and v in the domain of A, and that the function $\phi(t)$ be continuous and in $L^p(0, \infty)$, for some p in $1 \leqq p \leqq 2$. Actually, under these conditions, we prove a slightly stronger result; namely, that there exist constants $K > 0$ and μ such that the non-trivial solutions of (1.1) satisfy $\| u(t) \| \geqq Ke^{\mu t}$.

The restriction in Lax's result on the size of $\phi(t)$ cannot be lessened in general. For in the contrary case he constructed a solution of (1.1) that, as $t \to \infty$, behaves like $\exp(-bt^2)$, b being a positive linear function of $\sup \phi(t)$. It is therefore natural to ask whether there exist solutions of (1.1) which, as $t \to \infty$, tend to zero faster than $\exp(-\lambda t^2)$, for every $\lambda > 0$. We shall show that, at least for symmetric operators, this is only possible for the trivial solution. More generally, we obtain results that relate the rate of decay at infinity of the solutions of (1.1) to the asymptotic behavior of the function $\phi(t)$.

In the final portion of this paper we derive similar results for solutions of concrete parabolic differential inequalities. Results concerning the asymptotic behavior of solutions of parabolic partial differential ine-

Received October 1960. This research was supported by the United States Air Force through the Air Force Office of Scientific Research of the Air Research and Development Command, under contract No. AF49(638)-253. Reproduction in whole or in part is permitted for any purpose of the United States Government.

† Now at New York University, the Institute for Mathematical Sciences.

qualities have been obtained recently by M. H. Protter [2].

2. The estimate from below. Throughout this paper A will denote a symmetric operator in a Hilbert space H, and $u(t)$ will denote a strongly continuously differentiable function defined for $0 \leqq t < \infty$ with values in H such that $Au(t)$ to continuous. We shall also assume that $\phi(t)$ is a positive continuous function belonging to $L^p(0, \infty)$, for some p in the interval $1 \leqq p \leqq 2$.

THEOREM 1. *If $u(t)$ is a solution of the abstract differential inequality*

$$\left\| \frac{du}{dt} - Au \right\| \leqq \phi(t) \| u \| , \qquad 0 \leqq t < \infty , \tag{2.1}$$

and $u(0) \neq 0$, then there exists $K > 0$ and μ such that

$$\| u(t) \| \geqq K e^{\mu t} , \qquad 0 \leqq t < \infty . \tag{2.2}$$

The proof of Theorem 1 requires several lemmas concerning operators in finite-dimensional Hilbert spaces. Let D be a symmetric operator in a finite-dimensional Hilbert space F. Since F is finite-dimensional and D is symmetric, there is no loss of generality in assuming that D is in diagonal form.

For any real number λ and any vector v in F, denote by $P_\lambda v$ the projection of v onto the subspace of F spanned by those eigenvectors of D whose eigenvalues are not less than λ. Since D is in diagonal form, we have

$$(DP_\lambda v, P_\lambda v) \geqq \lambda \| P_\lambda v \|^2 . \tag{2.3}$$

Similarly, if we define $R_\lambda v = v - P_\lambda v$, then

$$(DR_\lambda v, R_\lambda v) \leqq \lambda \| R_\lambda v \|^2 . \tag{2.4}$$

Let ρ be an arbitrary positive number, and define a sequence $\{t_n\}$ as follows: $t_0 = 0$, and t_n, for positive integers n, is determined from the relation

$$\int_{t_n}^{t_{n+1}} \phi(\eta) d\eta = \rho , \tag{2.5}$$

where $t_{n+1} = \infty$ if

$$\int_{t_n}^{\infty} \phi(\eta) d\eta < \rho .$$

LEMMA 1. *Let $v(t)$, $0 \leqq t < \infty$, be a differentiable function of t with values in F such that*

$$\left\| \frac{dv}{dt} - Dv \right\| \leqq \phi(t) \| v \| , \qquad 0 \leqq t < \infty . \tag{2.6}$$

Assume that $v(0) \neq 0$ *and that*

$$\| P_\lambda v(t_n) \| \geqq \| R_\lambda v(t_n) \| . \tag{2.7}$$

Then there exists $\rho_0 > 0$ *such that for all* $\rho \leqq \rho_0$

$$2 \| P_\lambda v \| \geqq \| R_\lambda v \| , \qquad t_n \leqq t \leqq t_{n+1} . \tag{2.8}$$

Proof. The set $T = \{t\colon 2 \| R_\lambda v \| \leqq \| P_\lambda v \|\}$ is closed, and the inequality (2.8) obviously holds for each t in T. Thus it is sufficient to prove (2.8) for t in CT, the complement of T. Since CT is an open set of reals, it can be represented as a denumerable union of disjoint open intervals. Therefore it suffices to prove (2.8) for a generic open interval, $a < t < b$ say, forming this union.

We have

$$\| P_\lambda v(t) \| < 2 \| R_\lambda v(t) \| , \qquad a < t < b , \tag{2.9}$$

and

$$\| P_\lambda v(a) \| \geqq \| R_\lambda v(a) \| . \tag{2.10}$$

Since the space F is finite dimensional, D is a bounded operator (the bound for D may depend on the dimension of F), and this implies that the inequality (2.6) can have only one solution with prescribed initial value $v(0)$. Thus $v(t)$ can never vanish since $v(0) \neq 0$. It follows now from (2.9) that $R_\lambda v(t)$ is nonzero in $a < t < b$, so that we can form the function

$$f(t) = \frac{\| P_\lambda v(t) \|^2}{\| R_\lambda v(t) \|^2} . \tag{2.11}$$

Differentiating $f(t)$, we find that

$$\begin{aligned} \| R_\lambda v \|^4 \frac{df}{dt} &= 4 \| R_\lambda v \|^2 Re\Big(P_\lambda v, P_\lambda \frac{dv}{dt}\Big) \\ &\quad - 4 \| P_\lambda v \|^2 Re\Big(R_\lambda v, R_\lambda \frac{dv}{dt}\Big) . \end{aligned} \tag{2.12}$$

Since $v(t)$ satisfies the inequality (2.6) and P_λ and R_λ are projections, we can write

$$P_\lambda \frac{dv}{dt} = D(P_\lambda v) + Q_1 \tag{2.13}$$

and

$$P_\lambda \frac{dv}{dt} = D(R_\lambda v) + Q_2 \,, \tag{2.14}$$

where

$$\| Q_i \| \leqq \phi(t) \| v \| \,, \qquad (i = 1, 2) \,. \tag{2.15}$$

It follows from (2.13) and (2.15) that

$$Re\left(P_\lambda v, P_\lambda \frac{dv}{dt}\right) \geqq (P_\lambda v, DP_\lambda v) - \phi(t) \| v \|^2 \,.$$

Applying (2.3) to the first term on the right, we obtain

$$Re\left(P_\lambda v, P_\lambda \frac{dv}{dt}\right) \geqq \lambda \| P_\lambda v \|^2 - \phi(t) \| v \|^2 \,. \tag{2.16}$$

Similarly, we have

$$Re\left(R_\lambda v, R_\lambda \frac{dv}{dt}\right) \leqq \lambda \| R_\lambda v \|^2 + \phi(t) \| v \|^2 \,. \tag{2.17}$$

Entering the estimates (2.16) and (2.17) into the right side of (2.12), we find that

$$\| R_\lambda v \|^4 \frac{df}{dt} \geqq -8\phi(t) \| v \|^4 \,. \tag{2.18}$$

Here we have made use of the inequalities $\| P_\lambda v \| \leqq \| v \|$ and $\| R_\lambda v \| \leqq \| v \|$. It follows from (2.9) that

$$\| v \|^2 = \| P_\lambda v \|^2 + \| R_\lambda v \|^2 \leqq 5 \| R_\lambda v \| \ .$$

This inequality and (2.18) imply that

$$\frac{df}{dt} \geqq -200\phi(t) \,,$$

and therefore

$$f(t) \geqq f(a) - 200\int_a^t \phi(\eta) d(\eta) \,. \tag{2.19}$$

Now, according to (2.10) and (2.11), $f(a) \geqq 1$. Therefore if we make use of (2.5), we conclude from (2.19) that

$$\frac{\| P_\lambda v(t) \|^2}{\| R_\lambda v(t) \|^2} \geqq 1 - 200\rho \geqq \tfrac{1}{4} \,,$$

provided that $800\rho_0 = 3$. This completes the proof of the lemma.

LEMMA 2. *Let $v(t)$ satisfy the conditions of Lemma 1. If*

$$\pi = \lambda - 200\rho(t_{n+1} - t_n)^{-1} , \tag{2.20}$$

then, for all $\rho \leqq \rho_0$,

$$\| P_\pi v(t_{n+1}) \| \geqq \| R_\pi v(t_{n+1}) \| . \tag{2.21}$$

Proof. First, assume that

$$\| P_\lambda v(t) \| \leqq 2 \| R_\pi v(t) \| , \qquad t_n < t < t_{n+1} . \tag{2.22}$$

Setting

$$g(t) = \frac{\| P_\lambda v(t) \|^2}{\| R_\pi v(t) \|^2} , \tag{2.23}$$

we obtain

$$\begin{aligned} \| R_\pi v \|^4 \frac{dg}{dt} &= 4 \| R_\pi v \|^2 Re\left(P_\lambda v,\, P_\lambda \frac{dv}{dt}\right) \\ &\quad - 4 \| P_\pi v \|^2 Re\left(R_\pi v,\, R_\pi \frac{dv}{dt}\right) . \end{aligned} \tag{2.24}$$

As in the proof of Lemma 1, we have

$$Re\left(R_\pi v,\, R_\pi \frac{dv}{dt}\right) \leqq \pi \| R_\pi v \|^2 + \phi(t) \| v \|^2 . \tag{2.25}$$

Inserting the estimates (2.16) and (2.25) into the right side of (2.24), we conclude that

$$\| R_\pi v \|^4 \frac{dg}{dt} \geqq 4(\lambda - \pi) \| R_\pi v \|^2 \| P_\lambda v \|^2 - 8\phi(t) \| v \|^2 . \tag{2.26}$$

Since $\| P_\pi v \| \geqq \| P_\lambda v \|$, (2.22) implies that $\| v \| \leqq 5 \| R_\pi v \|^2$, which, when inserted into (2.21), yields

$$\frac{dg}{dt} \geqq 200\rho_0(t_{n+1} - t_n)^{-1} \frac{\| P_\lambda v \|^2}{\| R_\lambda v \|^2} - 200\phi(t) . \tag{2.27}$$

Here we have employed the inequality $\| R_\pi v \| \leqq \| R_\lambda v \|$. By Lemma 1, $4 \| P_\lambda v \|^2 \geqq \| R_\lambda v \|^2$, so that we obtain from (2.27)

$$\frac{dg}{dt} \geqq 200\rho_0(t_{n+1} - t_n)^{-1} - 200\phi(t) . \tag{2.28}$$

Finally, when we integrate (2.28) between t_n and t_{n+1} and apply (2.7) and (2.23) to the result, we obtain the desired inequality (2.21).

Now assume that there is a value of $t < t_{n+1}$ such that $\| P_\lambda v \| > 2 \| R_\pi v \|$. Let $\bar{t}$ be the last such value of t. If $\bar{t} = t_{n+1}$ there is nothing to prove, so we assume that $\bar{t} < t_{n+1}$. In this situation we have that

$$\| P_\lambda v(t) \| \leqq 2 \| R_\pi v(t) \| , \qquad \bar{t} < t < t_{n+1} ,$$

and

$$\| P_\pi v(\bar{t}) \| \geqq 2 \| R_\pi v(\bar{t}) \| .$$

The reasoning emplyed to prove Lemma 1 can now be used to establish the inequality

$$\| P_\pi v(t) \| \geqq \frac{\sqrt{11}}{2} \| R_\pi v(t) \| , \qquad \bar{t} \leqq t \leqq t_{n+1} ,$$

which certainly implies (2.21).

From the sequence $\{t_n\}$ we form the series

$$\sigma = \sum_{n=0}^{\infty} (t_{n+1} - t_n)^{-1} .$$

Our assumption that $\phi(t)$ belongs to $L^p(0, \infty)$, for some p in the interval $1 \leqq p \leqq 2$, implies that σ converges. This is clear when $p = 1$ since in this case the series has only a finite number of nonzero terms. Assume that $1 < p \leqq 2$. Applying Hölder's inequality to (2.5), we obtain the nequality

$$\rho \leqq \left(\int_{t_n}^{t_{n+1}} \phi^p(\eta) d\eta \right)^{1/p} (t_{n+1} - t_n)^{1/q} ,$$

where $p^{-1} + q^{-1} = 1$. Therefore

$$(t_{n+1} - t_n)^{-1} \leqq \rho^{-1} \left(\int_{t_n}^{t_{n+1}} \phi^p(\eta) d\eta \right)^{q/p} ,$$

which, since $q \geqq p$, implies that σ converges.

Also, we note here that our assumption that $\phi(t)$ belongs to $L^p(0, \infty)$, for some p in the interval $1 \leqq p \leqq 2$, implies that there exist constants C_1 and C_2 such that

$$\int_0^t \phi(\eta) d\eta \leqq C_1 t + C_2 . \tag{2.29}$$

From now on we shall assume that ρ has the fixed value ρ_0.

LEMMA 3. *Let $v(t)$ satisfy the conditions of Lemma 1. If*

$$\| P_\lambda v(0) \| \geqq \| R_\lambda v(0) \| , \tag{2.30}$$

then

$$\| v(t) \| \geqq \tfrac{1}{2} e^{-\sigma_2} \| v(0) \| e^{\mu t} , \qquad 0 \leqq t < \infty , \tag{2.31}$$

where $\mu = \lambda - 200\rho_0\sigma - 3C_1$.

Proof. Set $\lambda_0 = \lambda - 200\rho_0\sigma$. We assert that

$$2 \| P_{\lambda_0} v(t) \| \geqq \| R_{\lambda_0} v(t) \| , \qquad 0 \leqq t < \infty . \tag{2.32}$$

Let t be arbitrary. Then for some n, $t_n \leqq t \leqq t_{n+1}$. It follows from (2.30), Lemma 1 and Lemma 2 that

$$2 \| P_\pi v(t) \| \geqq \| R_\pi v(t) \| , \qquad t_n \leqq t \leqq t_{n+1} . \tag{2.33}$$

Hence the inequality

$$\| P_{\lambda_0} v \| \geqq \| P_\pi v \| \geqq \tfrac{1}{2} \| R_\pi v \| \geqq \tfrac{1}{2} \| R_{\lambda_0} v \|$$

implies (2.32) for this particular value of t.

It follows from (2.32) that

$$\| v(t) \| \leqq 3 \| P_{\lambda_0} v(t) \| , \qquad 0 \leqq t < \infty . \tag{2.34}$$

Set $z(t) = P_{\lambda_0} v(t)$. Then by (2.34) $z(t)$ is a solution of the differential inequality

$$\left\| \frac{dz}{dt} - Dz \right\| \leqq 3\phi(t) \| z \| , \qquad 0 \leqq t < \infty . \tag{2.35}$$

Differentiating $\| z \|^2$, and taking (2.35) into account, we get

$$\frac{d}{dt} \| z \|^2 = 2Re\left(z, \frac{dz}{dt}\right) \geqq 2Re(z, Dz) - 6\phi(t) \| z \|^2 . \tag{2.36}$$

Since $z(t) = P_{\lambda_0} v(t)$, it follows from (2.3) and (2.36) that

$$\frac{d}{dt} \| z \|^2 \geqq (2\lambda_0 - 6\phi(t)) \| z \|^2 . \tag{2.37}$$

Consequently, if we integrate (2.37), we obtain

$$\| v(t) \|^2 \geqq \| z(t) \|^2 \geqq \| z(0) \|^2 \exp\left[2\int_0^t (\lambda_0 - 3\phi(\eta) d\eta)\right] \geqq \tfrac{1}{4} e^{-C_2} \| v(0) \|^2 e^{2\mu t} ,$$

which is equivalent to (2.31).

To pass from the finite to the infinite dimentional case, we have to show that the cut-off parameter λ can be selected independently of the dimension of the space F.

LEMMA 4. *Let $v(t)$ satisfy the conditions of Lemma 1. Then there exists a λ, depending only on $\| v(0) \|$, $\| v(1) \|$ and $\phi(t)$, such that*

$$\| P_\lambda v(1) \| \geqq \| R_\lambda v(1) \| . \tag{2.38}$$

Proof. Define $w(t) = v(1 - t)$. Then $w(t)$ is a solution of the differential inequality

$$\left\| \frac{dw}{dt} + Dw \right\| \leqq \phi(t) \| w \| , \qquad 0 \leqq t < \infty . \tag{2.39}$$

If for some λ

$$\| P_\lambda v(1) \| < \| R_\lambda v(1) \| ,$$

then

$$\| P_{-\lambda} w(0) \| > \| R_{-\lambda} w(0) \| . \tag{2.40}$$

Applying Lemma 3 to (2.39) and (2.40), we find that

$$\| v(0) \| = \| w(1) \| \geqq \tfrac{1}{2} e^{-\sigma_2} \| w(0) \| e^m ,$$

where

$$m = -\lambda - 200\rho_0\sigma - 3C_1 .$$

Hence

$$\lambda \geqq \log\left[\frac{2 \| v(0) \|}{\| v(1) \|} e^{\sigma_2} \right] - 200\rho_0 - 3C_1 . \tag{2.41}$$

Thus if λ is chosen smaller than the right side of (2.41), then the desired inequality (2.38) holds.

3. **Proof of Theorem 1.** Let k be an arbitrary positive integer. Using the continuity of the derivative of $v(t)$, one can show that for any $\varepsilon > 0$ there exists a $\delta = \delta(\varepsilon, k) > 0$ such that

$$\left\| \frac{d}{dt} u(t) - \frac{u(t+h) - u(t)}{h} \right\| < \varepsilon , \tag{3.1}$$

for $|h| < \delta$ and $|t| \leqq k$.

We subdivide the interval $0 \leqq t \leqq k$ into equal subintervals of length Δ, where $\Delta < \delta$, and

$$\| Au(t+h) - Au(t) \| < \varepsilon , \tag{3.2}$$

for $|h| < \Delta$. We assume that the point $t = 1$ is included in the subdivision.

Let G be the subspace of H generated by $u(0), u(\Delta), u(2\Delta), \cdots, u(k)$. Let $A_0 = EA$, where E is the projection of H onto the subspace G. Clearly, the operator A_0 restricted to the subspace G is symmetric.

For any subdivision point $j\Delta$, we have

$$\left\| \frac{u((j+1)\Delta) - u(j\Delta)}{\Delta} - A_0 u(j\Delta) \right\| \leqq \left(1 + \frac{\varepsilon}{M}\right) \phi(j\Delta) \| u(j\Delta) \| , \tag{3.3}$$

where M is the infimum of $\phi(t) \| u(t) \|$ for $0 \leqq t \leqq k$. Let $u_0(t)$ be equal to $u(t)$ at the subdivision points and be linear in between. Note that $u_0(t)$ has its values in the finite-dimensional subspace G of H.

It follows from (3.2) and (3.3) that

$$\| D_+u_0(t) - A_0u_0(t) \| \leqq \left(1 + \frac{2\varepsilon}{M}\right)\phi(j\varDelta) \| u(j\varDelta) \| , \tag{3.4}$$

where D_+ denotes right differentiation, and $j\varDelta \leqq t \leqq (j+1)\varDelta$. By taking $\varDelta$ sufficiently small and taking into account the continuity of $\phi(t)$, we obtain

$$\| D_+u_0 - A_0u_0 \| \leqq 2\phi(t) \| u_0 \| , \qquad 0 \leqq t \leqq k . \tag{3.5}$$

By Lemma 4, there is a $\bar{\lambda} = \bar{\lambda}(\| u(0) \|, \| u(1) \|, 2\phi(t))$ such that

$$\| P_{\bar{\lambda}}u_0(1) \| \geqq \| R_{\bar{\lambda}}u_0(1) \| . \tag{3.6}$$

Now we observe that the lemmas of the preceding section remain valid when $v(t)$ has a right derivative everywhere and is continuously differentiable, except at a finite number of points. Once this observation is made, we can conclude from (3.5), (3.6) and Lemma 3 that

$$\| u_0(t) \| \geqq \tfrac{1}{2} \| u(1) \| \exp\left[\int_1^t \psi(\eta)d\eta\right] , \qquad 1 \leqq t \leqq k ,$$

where

$$\psi(\eta) = \bar{\lambda} - 400\rho_0\sigma - 6\phi(\eta) .$$

Hence

$$\| u_0(t) \| \geqq \bar{K}e^{\bar{\mu}t} , \qquad 1 \leqq t \leqq k .$$

Letting $\varDelta \to 0$, we conclude that

$$\| u(t) \| \geqq \bar{K}e^{\bar{\mu}t} , \qquad 1 \leqq t \leqq k ,$$

which is easily seen to imply the inequality (2.2) of Theorem 1.

In the proof of Theorem 1 we tacitly assumed that $u(t)$ never vanishes. The proof of this fact is easy. For let t_0 denote the first value of t for which $u(t)$ is zero. Since $u(0) \neq 0$, $t_0 > 0$. According to Theorem 1, we have $\| u(t) \| \geqq Ke^{\mu t}$, for $0 \leqq t < t_0$, which shows that $u(t)$ cannot possibly vanish at t_0.

4. An A priori inequality. In this section we derive an a priori inequality for a class of functions with a prescribed rate of decay at infinity.

LEMMA 5. *Let $\psi(t)$ belong to $L^2(0, a)$, for every $a > 0$, and define*

$$\beta(t) = \lambda\int_0^t (t - \eta)\psi^2(\eta)d\eta . \tag{4.1}$$

Let $U(t)$ be a strongly continuously differentiable mapping from $0 \leqq t < \infty$ with values in H such that $AU(t)$ is continuous. If the support of $U(t)$ is contained in $0 < \varepsilon \leqq t < \infty$ and

$$\lim_{t\to\infty} \| U(t) \| \exp \beta(t) = 0 \ , \tag{4.2}$$

for every $\lambda > 0$, then

$$\lambda\int_0^\infty e2^{\beta(t)}\psi^2(t) \| U(t) \|^2 dt \leqq \int_0^\infty e^{2\beta(t)} \left\| \frac{dU}{dt} - AU \right\|^2 dt \ , \tag{4.3}$$

provided that the left side is finite.

Proof. We may assume that $U(t)$ vanishes for all sufficiently large values of t. For in the general case we can approximate U by the sequence $U_n(t) = \zeta_n(t)U(t)$, $\zeta_n(t)$ being a C^∞ function equal to one for $t \leqq n$, zero for $t \geqq n + 1$ and $0 \leqq \zeta_n \leqq 1$ in between. As $n \to \infty$, the inequality (4.3) for U_n goes over into (4.3) for U.

Now consider the integral

$$I = \int_0^\infty e^{2\beta(t)} \left\| \frac{dU}{dt} - AU \right\|^2 dt \ .$$

If we make the transformation $U(t) = e^{-\beta(t)}V(t)$, then

$$I = \int_0^\infty \left\| \frac{dV}{dt} - AV - \frac{d\beta}{dt} V \right\|^2 dt \ . \tag{4.4}$$

It follows from the elementary inequality

$$(a - b)^2 \geqq -2ab$$

and (4.4) that

$$I \geqq -2\int_0^\infty Re\left(\frac{dV}{dt}, \frac{d\beta}{dt} V + AV\right) dt \ . \tag{4.5}$$

We have

$$\begin{aligned} \int_0^\infty \left(\frac{dV}{dt}, \frac{d\beta}{dt} V\right) dt &= \int_0^\infty \frac{d}{dt}\left(V, \frac{d\beta}{dt} V\right) dt \\ &\quad - \int_0^\infty \left(V, \frac{d^2\beta}{dt^2} V + \frac{d\beta}{dt}\frac{dV}{dt}\right) dt \ . \end{aligned} \tag{4.6}$$

The first integral on the right vanishes since $V(t)$ has compact support. Hence

$$-2\int_0^\infty Re\left(\frac{dV}{dt}, \frac{d\beta}{dt} V\right) dt = \lambda\int_0^\infty e^{2\beta(t)}\psi^2(t) \| U(t) \|^2 dt \ . \tag{4.7}$$

In view of (4.5) and (4.7) it is sufficient to prove that

$$\int_0^\infty Re\Big(\frac{dV}{dt}, AV\Big)dt = 0 . \tag{4.8}$$

Taking into account the symmetry of A, we have that

$$\Big(\frac{dV}{dt}, AV\Big) = \frac{d}{dt}(V, AV) - \overline{\Big(\frac{dV}{dt}, AV\Big)} , \tag{4.9}$$

the bar denoting complex conjugation. Therefore

$$2\,Re\Big(\frac{dV}{dt}, AV\Big) = \frac{d}{dt}(V, AV) ,$$

from which (4.8) follows directly by integration. This completes the proof of the lemma.

5. A special instance of Theorem 1. As a first application of Lemma 5, we give a direct proof of a slightly weaker version of Theorem 1 in the case that $\phi(t)$ belongs to $L^2(0, \infty)$.

THEOREM 2. *Let $u(t)$ be a solution of the abstract differential inequality*

$$\Big\|\frac{du}{dt} - Au\Big\| \leqq \phi(t)\,\|u\| , \qquad 0 \leqq t < \infty , \tag{5.1}$$

where $\phi(t)$ belongs to $L^2(0, \infty)$. If

$$\lim_{t\to\infty} \|u(t)\|\, e^{\lambda t} = 0 \tag{5.2}$$

for every $\lambda > 0$, then u has property S, i.e., it vanishes identically for $0 \leqq t < \infty$.

Proof. Since $\phi \in L^2(0, \infty)$, it follows from (4.1) that (we take $\psi = \phi$)

$$\beta(t) \leqq \lambda t \int_0^\infty \phi^2(\eta)d\eta .$$

Therefore (5.2) implies that

$$\lim_{t\to\infty} \|u(t)\| \exp \beta(t) = 0 \tag{5.3}$$

for every $\lambda > 0$. Let $\zeta(t)$ be a C^∞ function equal to one for $0 < 2\varepsilon \leqq t$, equal to zero for $0 \leqq t \leqq \varepsilon$ and $0 \leqq \zeta \leqq 1$ in between. Set $U(t) = \zeta(t)u(t)$. Because of (5.3) and the fact that

$$\int_0^\infty \phi^2(t)e^{2\beta(t)}\, \|U(t)\|^2 dt < \infty ,$$

all of the conditions of Lemma 5 are met, and therefore

$$\lambda\int_{2\varepsilon}^{\infty}\phi^2(t)e^{2\beta(t)}\,\|u(t)\|^2 dt \leqq \int_{\varepsilon}^{2\varepsilon} e^{2\beta(t)}\left\|\frac{dU}{dt} - AU\right\|^2 dt + \int_{2\varepsilon}^{\infty}\phi^2(t)\,\|u(t)\|^2 dt\ .$$

If $\lambda \geqq 2$ then

$$\int_{3\varepsilon}^{\infty}\phi^2(t)e^{2\beta(t)}\,\|u(t)\|^2 dt \leqq \int_{\varepsilon}^{2\varepsilon} e^{2\beta(t)}\left\|\frac{dU}{dt} - AU\right\|^2 dt\ . \tag{5.4}$$

Using the monotonic character of $\beta(t)$, we get from (5.5) that

$$\int_{3\varepsilon}^{\infty}\phi^2(t)\,\|u(t)\|^2 dt \leqq \exp\left[\beta(2\varepsilon) - \beta(3\varepsilon)\right]\int_{\varepsilon}^{2\varepsilon}\left\|\frac{dU}{dt} - AU\right\|^2 dt\ . \tag{5.5}$$

Since $\beta(2\varepsilon) - \beta(3\varepsilon) \to -\infty$ as $\lambda \to \infty$, it follows from (5.5) that

$$\int_{3\varepsilon}^{\infty}\phi^2(t)\,\|u(t)\|^2 dt = 0\ .$$

Therefore $u(\mathrm{t}) = 0$ for $t \geqq 3\varepsilon$. Since ε is arbitrary, $u(t)$ vanishes identically for $0 \leqq t < \infty$.

In much the same way we can prove the following result for bounded ϕ.

THEOREM 3. *Let $u(t)$ be a solution of the abstract differential inequality* (5.1), *where $\phi(t) \leqq$ const. If*

$$\lim_{t\to\infty}\|u(t)\|\exp(\lambda t^2) = 0$$

for every $\lambda > 0$, then $u(t)$ vanishes identically.

More generally, we have the

THEOREM 4. *Let $u(t)$ be a solution of the abstract differential inequality* (5.1). *Assume that $\phi(t)$ belong to $L^2(0, a)$ for every $a > 0$, and*

$$\phi^2(t) \leqq \exp\left[\lambda\int_0^t (t-\eta)\phi^2(\eta)d\eta\right],$$

for all sufficiently large t and λ. If

$$\lim_{t\to\infty}\|u(t)\|\exp\beta(t) = 0\ ,$$

for every $\lambda > 0$, then $u(t)$ vanishes identically.

6. Parabolic differential inequalities. Let G be a bounded domain in the real Euclidean n-space R^n. For two real functions $u(x)$ and $v(x)$

belonging to $L^2(G)$ we denote by

$$(u, v) = \int_G u(x)v(x)dx$$

their real scalar product and by $\| u \|_0 = (u, u)^{1/2}$ the associated norm. Let $H_1^0(G)$ denote the closure of $C_0^\infty(G)$, the C^∞ functions on G with compact support in G, relative to the norm

$$\| u \|^2 = \int_G \left(| u(x) |^2 + \sum_{i=1}^n \left| \frac{\partial u}{\partial x_i}(x) \right|^2 \right) dx .$$

Consider the differential operator

$$L = \sum_{i,j=1} \frac{\partial}{\partial x_i}\left(a^{ij}(x) \frac{\partial}{\partial x_j} \right) , \tag{6.1}$$

where $a^{ij}(x) = a^{ji}(x)$. We assume that there exist positive constants m and M such that, for all x in $\bar{G}$ and all real vectors $\zeta = (\zeta_1, \zeta_2, \cdots, \zeta_n)$,

$$m \sum_{i=1}^n \zeta_i^2 \leqq \sum_{i,j=1}^n a^{ij}(x)\zeta_i\zeta_j \leqq M \sum_{i=1}^n \zeta_i^2 . \tag{6.2}$$

Thus L is a real elliptic differential operator.

If $u \in H_1^0(G)$ we say that $Lu \in L^2(G)$ when $\{a^{ij}(x)\}(\partial u/\partial x_j)$ is differentiable with respect to x_i (in the sense of distributions) and

$$\frac{\partial}{\partial x_i}\left(a^{ij}(x) \frac{\partial u}{\partial x_j} \right) \in L^2(G) .$$

It is not difficult to show that

$$(Lu, v) = (u, Lv) , \tag{6.3}$$

for u and v in $H_1^0(G)$ and Lu and Lv in $L^2(G)$, the common value of (6.3) is

$$(Lu, v) = - \int_G \left(\sum_{i,j=1}^n a^{ij}(x) \frac{\partial u}{\partial x_i} \frac{\partial u}{\partial x_j} \right) dx . \tag{6.4}$$

Thus the operator L is formally self-adjoint.

Let $\Gamma(t) = \exp\left[\gamma \int_0^t \psi^2(\eta) d\eta \right]$, and introduce the function

$$\sigma(t) = \lambda \int_0^t \Gamma(\eta) \int_0^\eta \Gamma^{-1}(\zeta)\phi^2(\zeta) d\zeta d\eta . \tag{6.5}$$

The function $\sigma(t)$ is non-decreasing provided that γ and λ are nonnegative. We also note that

$$\Gamma \frac{d}{dt}\left[\Gamma^{-1} \frac{d\sigma}{dt} \right] = \lambda \phi^2 . \tag{6.6}$$

If the functions $\phi(t)$ and $\psi(t) \in L^2(0, \infty)$, there exists a constant Γ_0, depending on γ, such that

$$\sigma(t) \leqq \lambda \Gamma_0 t \ . \tag{6.7}$$

We introduce the norm

$$\| u \|_1^2 = \int_G \sum_{i=1}^n \left| \frac{\partial u}{\partial x_i}(x) \right|^2 dx \ ,$$

which is equivalent to the norm defined above for $H_1^0(G)$.

LEMMA 6. *Let $\phi(t)$ and $\psi(t)$ belong to $L^2(0, \infty)$. Let $Z(t)$ be a strongly continuously differentiable mapping from $0 \leqq t < \infty$ with values in $H_1^0(G)$ such that $LZ(t) \in L^2(G)$ is continuous. If the support of $Z(t)$ is contained in $0 < t_0 \leqq t < \infty$ and*

$$\lim_{t \to \infty} \| Z(t) \|_1 e^{\lambda t} = 0 \ , \tag{6.8}$$

for every $\lambda > 0$, then

$$\begin{aligned} \int_0^\infty \Gamma^{-1}(t) e^{2\sigma(t)} \left\| LZ - \frac{\partial Z}{\partial t} \right\|_0^2 dt &\geqq \lambda \int_0^\infty \Gamma^{-1}(t) e^{2\sigma(t)} \phi^2(t) \| Z \|_0^2 dt \\ &\quad + m\gamma \int_0^\infty \Gamma^{-1}(t) e^{2\sigma(t)} \psi^2(t) \| Z \|_1^2 dt \ . \end{aligned} \tag{6.9}$$

Proof. The integrals on the right side of (6.9) are finite because of (6.7) and (6.8). As in Lemma 5, we may assume that $Z(t)$ is identically zero for all sufficiently large values of t. Set $Z(t) - e^{-\sigma(t)} V(t)$. Then if J denotes the integral on the left side of (6.9), we have

$$J \geqq -2 \int_0^\infty \Gamma^{-1}(t) \left(\frac{dV}{dt}, LV + V \frac{d\sigma}{dt} \right) dt \ . \tag{6.10}$$

Integrating by parts and using the fact that $V(t)$ has compact support, we find that

$$-2 \int_0^\infty \Gamma^{-1}(t) \left(\frac{dV}{dt}, V \frac{d\sigma}{dt} \right) dt = \lambda \int_0^\infty \Gamma^{-1}(t) e^{2\sigma(t)} \phi^2(t) \| Z \|_0^2 dt \ . \tag{6.11}$$

In proving (6.11) we have made use of (6.6).

Since L is real and symmetric, we have

$$-2 \int_0^\infty \Gamma^{-1}(t) \left(\frac{dV}{dt}, LV \right) dt = - \int_0^\infty \Gamma^{-1}(t) \frac{d}{dt} (V, LV) dt \ . \tag{6.12}$$

Another integration by parts yields

$$-2 \int_0^\infty \Gamma^{-1}(t) \left(\frac{dV}{dt}, LV \right) dt = -\gamma \int_0^\infty \Gamma^{-1}(t) \psi^2(t) (V, LV) dt \ . \tag{6.13}$$

In view of (6.2) and (6.4) we have $(V, LV) \leqq -m \| V \|_1^2$, so that (6.13) implies that

$$(6.14) \qquad -2\int_0^\infty \Gamma^{-1}(t)\left(\frac{dV}{dt}, LV\right)dt \geqq \gamma m \int_0^\infty \Gamma^{-1}(t) e^{2\sigma(t)} \psi^2(t) \| Z \|_1^2 dt \, .$$

Combining (6.10), (6.11) and (6.14), we get (6.9).

THEOREM 5. *Let $\phi(t)$ and $\psi(t)$ belong to $L^2(0, \infty)$. Let $u(t)$ be a strongly continuously differentiable function from $0 \leqq t < \infty$ with values in $H_1^0(G)$ such that $Lu(t) \in L^2(G)$ is continuous. If $u(t)$ satisfies the differential inequality*

$$(6.5) \qquad \left\| Lu - \frac{\partial u}{\partial t} \right\|_0^2 \leqq \phi^2(t) \| u \|_0^2 + \psi^2(t) \| u \|_1^2 \, , \qquad 0 \leqq t < \infty \, ,$$

and

$$\lim_{t \to \infty} \| u(t) \|_1 e^{\lambda t} = 0 \, ,$$

for every $\lambda > 0$, then u vanishes identically.

Theorem 5 follows from Lemma 6 in much the same way that Theorem 2 follows from Lemma 5, and for this reason the proof will be omitted.

If in Theorem 5 we only assume that $\phi(t)$ is bounded, then we can deduce from Lemma 6 that only the trivial solution of (6.15) can vanish faster than $\exp(-\lambda t^2)$, for every $\lambda > 0$.

BIBLIOGRAPHY

1. P. D. Lax, *A stability theorem for solutions of abstract differential equations, and its application to the study of the local behavior of solutions of elliptic equations*, Comm. Pure and Appl. Math., **9** (1956), 747-766.

2. M. H. Protter, *Properties of Solutions of Parabolic Equations and Inequalities*, Univ. of California Tech. Report, Prepared under Contract AF 49(638)-398 (1960).

STANFORD UNIVERSITY
CALIFORNIA INSTITUTE OF TECHNOLOGY

SELF-INTERSECTION OF A SPHERE ON A COMPLEX QUADRIC

I. FÁRY

1. The real part S^n of a quadric V in complex, affine $(n+1)$-space is a sphere. The self-intersection of S^n *in* V is the same as the self-intersection of a "vanishing cycle," introduced by Lefschetz, and plays a certain role in [4], [5]. We will compute here this self-intersection number, using elementary tools.

Let us introduce some notations. P_{n+1} denotes the complex projective space of algebraic dimension $n+1$, hence of topological dimension

$$\dim P_{n+1} = 2n + 2 .$$

To each projective sub-space P_k of P_{n+1} a *positive* orientation can be given, thus it can be considered as a *cycle* p_{2k}. Then we agree that

$$\text{(1)} \qquad \textit{if } k + l = n + 1, \text{ then } (p_{2k}, p_{2l}) = 1 \text{ in } P_{n+1}$$

be true for the *intersection numbers* of cycles. This is the usual convention, the one in [1], for example; in [7] another convention is adopted.

Let $x_1, \cdots, x_{n+2}$ be a fixed system of projective coordinates in P_{n+1}. Then

$$\text{(2)} \qquad Q_n : x_1^2 + \cdots + x_{n+2}^2 = 0$$

is a *non-singular quadric*; $\dim Q_n = 2n$. The points of P_{n+1} whose last coordinate is non-zero form a complex affine space C_{n+1}, and

$$V = Q_n \cap C_{n+1} = [x : x \in Q_n, x_{n+2} \neq 0]$$

is a *non-singular affine quadric.* If $z \in C_{n+1}$, we denote by $z_1, \cdots, z_{n+2}$ those coordinates for which $z_{n+2} = i$ where $i^2 = -1$; thus $z_1, \cdots, z_{n+1}$ are affine coordinates in C_{n+1}. Then

$$V\colon z_1^2 + \cdots + z_{n+1}^2 = 1 \qquad (z \in C_{n+1})$$
$$S^n\colon z_1^2 + \cdots + z_{n+1}^2 = 1, z_1 \cdots, z_{n+1} \textit{ reals}$$

are the equations of an affine quadric and its real part respectively; *this real part* S^n is, of course, a sphere. We consider S^n with an arbitrarily chosen and fixed orientation as a cycle s. It is well known (see, for example, [2], p. 35, (g)) that

Received September 12, 1960.

(3) *the homology class s, of the cycle whose carrier is S^n, generates $H_n(V; Z)$,*

where Z denotes the ring of integers.

As dim $V = 2$ dim S^n, the self intersection number

$$(s, s) = (S^n S^n)\ , \qquad (\text{in } V)\ , \tag{4}$$

of s in V, is well defined; we may write (S^n, S^n) for this self intersection number, because (s, s) does not depend on the orientation of S^n, used in (3).

2. M. F. Atiyah communicated to me his computation of the intersection number (4) for $n = 2$, showing that the sign in [2], p. 35 (10) is not the right one.[1] The determination of the sign of (4) given below is a generalization to n dimensions of the construction of Atiyah. In [2] we used only the fact that (4) is not zero, if n is even, hence other results of that paper are not invalidated by the false sign in (10), p. 35. The mistaken sign is "classical." Wrong sign appears in [4], p. 93, Théorème sur les Γ_{d-1} de C_u, I, [5] on top of p. 16, [8], p. 102, (3), and [7], p. 104, Theorem 45 (although in [7] not the convention (1) is used, the alternation of the sign in question is independent of any convention). After the completion of the present paper [6] appeared, where the classical mistake in sign is corrected (see (11.3) on p. 161). The results of [1] are in agreement with the sign (5) below.

3. Using the notations and conventions introduced above, we will prove the following theorem.

THEOREM. *Let s be the homology class of the oriented sphere S^n in $H_n(Q_n; Z)$ where $n = 2h$ is even. Let us denote by (s, s) the self-intersection number of s computed with the convention* (1). *Then*

$$(s, s) = \begin{cases} -2\ , & \textit{if } h = \dfrac{n}{2} \textit{ is odd}\ ; \\[2ex] +2\ , & \textit{if } h = \dfrac{n}{2} \textit{ is even}\ ; \end{cases} \tag{5}$$

holds true.

[1] I take the opportunity to correct another mistake in [2], also noticed by Atiyah. In Proposition 2, p. 27, we have to suppose that the singularity in question is *conical*. In [2], Proposition 2 is stated without proof; Atiyah gave an example showing that the statement does not hold true, if the singularity is *not conical*, and gave a proof with the correct hypothesis. Proposition 2 is used in [2] only in connection with conical singularities; thus other results of [2] are not affected by the incomplete formulation of that Proposition.

4. We prepare the proof of this theorem; for the first part of the proof, see [1]. (See also [3], pp. 230–232.) In order to describe easily linear sub-spaces of Q_n, we introduce new projective coordinates in P_{n+1}:

$$\begin{matrix} u_j = x_{2j-1} + ix_{2j} \\ v_j = x_{2j-1} - ix_{2j} \end{matrix} \quad j = 1, \cdots, h+1 \qquad (i^2 = -1) .$$

Let us notice that

(6) $$u_j = v_j = 0 \textit{ if and only if } x_{2j-1} = x_{2j} = 0 .$$

The equation of Q_n is

$$u_1 v_1 + \cdots + u_{h+1} v_{h+1} = 0 ,$$

in the new coordinates.

We consider the following linear sub-spaces of Q_n:

(7) $$A: \quad u_j = 0 , \quad j = 1, \cdots, h, h+1 ;$$

(8) $$B: \quad u_j = 0 , \quad j = 1, \cdots, h ; \quad v_{h+1} = 0 ;$$

(9) $$C: \quad v_j = 0 , \quad j = 1, \cdots, h+1 .$$

Let us remark that,

(10) $$A \cap C = \phi , \quad B \cap C \textit{ is just one point,}$$

by (6).

LEMMA 1. *Let X be one of the projective spaces A, B, C. If, in the system of equations defining X, we replace an even number of equations $u_j = 0$ by the corresponding $v_j = 0$, or vice versa, we define a new linear sub-space of Q_n belonging to the same continuous system as X. Similarly, without leaving the continuous system containing B, we may replace $u_h = 0$, $v_{h+1} = 0$ in (8) by $v_h = 0$ and $u_{h+1} = 0$.*

Proof. Let us suppose that we want to replace $v_1 = 0$, $v_2 = 0$ in (9) by $u_1 = 0$, $u_2 = 0$. Let us consider the linear space

$$\begin{matrix} \alpha v_2 + \beta u_1 = 0 , \\ -\alpha v_1 + \beta u_2 = 0 , \end{matrix} \quad v_3 = 0, \cdots, v_{h+1} = 0 ,$$

defined for every $(\alpha, \beta) \neq (0, 0)$. This projective space is clearly contained in Q_n. For $(1, 0)$ we have C and for $(0, 1)$ the desired replacement. The last statement of the lemma is proved similarly using the system

$$\begin{aligned} \alpha u_h + \beta u_{h+1} &= 0 , \\ -\beta v_h + a v_{h+1} &= 0 . \end{aligned}$$

Let us consider now A, B, C as cycles of Q_n, and let us denote by a, b, c their respective homology classes in $H_n(Q_n; Z)$.

LEMMA 2. *If h is odd, then $c = a$. If h is even, then $c = b$.*

Proof. If h is odd, the $h + 1$ equations of (9) can be replaced by the equations $u_j = 0$, $j = 1, \cdots, h + 1$. Hence A and C belong to the same continuous system. If h is even, we can replace the first h equations defining C by $u_j = 0$, $j = 1, \cdots, h$. Hence C and B belong to the same continuous system.

LEMMA 3. *As to the intersection numbers, we have*

(11) *if h is odd, then $(a, a) = 0$, $(b, b) = 0$, $(a, b) = 1$,*

(12) *if h is even, then $(a, a) = 1$, $(b, b) = 1$, $(a, b) = 0$.*

Proof. (1) Let h be *odd*. By Lemma 2 and the first equation of (10), we have $(a, a) = 0$. Similarly, the second equation of (10) and Lemma 2 prove $(a, b) = 1$. In order to prove $(b, b) = 0$, we consider the space

$$B' : \quad v_j = 0 \,, \quad j = 1, \cdots, h \,, \quad u_{h+1} = 0 \,.$$

We claim that B and B' are in the same continuous system. In order to prove this statement, we use Lemma 1 twice. First, we replace the last two equations of (8) by $v_h = 0$, and $u_{h+1} = 0$. Second, in the system obtained by the first step, we replace the first $h - 1$ equations by $v_j = 0$. Now $B \cap B' = \phi$, and this proves $(b, b) = 0$.

(2) Let h be *even*. The proof of (12) is similar to the previous one. The last two equations of (12) are immediate from (10) and Lemma 2. Using Lemma 1, we can find presently a B'', such that $B \cap B''$ be just one point.

LEMMA 4. *Using the previous notations s, a, b, for homology classes,*

$$s = \pm(a - b) \,, \tag{13}$$

the sign depending on the chosen orientation of S^n.

Proof. Let us denote by I the hyperplane $x_{n+2} = 0$, Then, clearly,

$$A \cap I = B \cap I \,.$$

We denote by J this intersection ($J = A \cap B$). Let us consider a pencil of k-planes, $2k + \dim A = 2n + 2$, in general position. If N is a neighborhood of J in B, the k-planes of the pencil project N into a neighborhood M of J in A. Given now a Riemann metric of P_{n+1}, if N is a small enough neighborhood of J, the corresponding points of N, M determine unique geodesic segments. We consider now B as a cycle, whose simplexes are so small that those intersecting J are contained in N. Using the geodesic segments introduced above which start at points of the simplexes of B intersecting J, it is easy to construct a chain E of Q_n, such that

$$A - B + \partial E \tag{14}$$

be a sum of simplexes of $V = Q_n - I$. Hence, s being a generator of $H_n(V; Z)$, (14) will be homologous to a multiple of s. Thus $a - b = ms$ for some integer m. Now $(a - b, a) = m(s, a)$ is ± 1 by Lemma 3, hence $m = \pm 1$.

Proof of the Theorem. (1) Let us suppose that h is *odd*. We use (13) and (11): $(s, s) = (a - b, a - b) = (a, a) - (b, a) - (a, b) + (b, b) = -(b, a) - (a, b) = -2$.

(2) Let us suppose that h is *even*. This time we use (12): $(s, s) = (a, a) + (b, b) = +2$. Hence the proof of (5) is complete.

References

1. É. Cartan, *Sur les propriétés topologiques des quadriques complexes*, Publications Mathématiques de l'Université de Belgrade, **1** (1932), 1-20.

2. I. Fáry, *Cohomologie des variétés algébriques*, Annals of Math. **65** (1957), 21-73.

3. Hodge and Pedoe, *Methods of Algebraic Geometry*, Vol. II, Cambridge University Press, 1954.

4. S. Lefschetz, *L'Analysis Situs et la Géométrie Algébrique*, Paris, Gathier-Villars, 1924.

5. ———, *Géométrie sur les surfaces et les variétés algébriques*, Mémorial des Sciences Mathématiques, Fasc. XL, Paris, Gauthier-Villars, 1929.

6. J. Leray, *Le calcul différentiel et intégral sur une variété analytique complexe* (*Problème de Cauchy*, III), Bulletin de la Soiété Mathématique de France, **87** (1959), 81-180.

7. A. H. Wallace, *Homology Theory on Algebraic Varieties*, New York, Pergamon Press, 1958.

8. O. Zariski, *Algebraic Surfaces*, Berlin, 1935, Springer Verlag.

UNIVERSITY OF CALIFORNIA, BERKELEY

GROUPS WHICH HAVE A FAITHFUL REPRESENTATION OF DEGREE LESS THAN $(p-1/2)$

WALTER FEIT AND JOHN G. THOMPSON

1. Introduction. Let G be a finite group which has a faithful representation over the complex numbers of degree n. H. F. Blichfeldt has shown that if p is a prime such that $p > (2n+1)(n-1)$, then the Sylow p-group of G is an abelian normal subgroup of G [1]. The purpose of this paper is to prove the following refinement of Blichfeldt's result.

THEOREM 1. *Let p be a prime. If the finite group G has a faithful representation of degree n over the complex numbers and if $p > 2n+1$, then the Sylow p-subgroup of G is an abelian normal subgroup of G.*

Using the powerful methods of the theory of modular characters which he developed, R. Brauer was able to prove Theorem 1 in case p^2 does not divide the order of G [2]. In case G is a solvable group, N. Ito proved Theorem 1 [4]. We will use these results in our proof.

Since the group $SL(2, p)$ has a representation of degree $n = (p-1)/2$, the inequality in Theorem 1 is the best possible.

It is easily seen that the following result is equivalent to Theorem 1.

THEOREM 2. *Let A, B be n by n matrices over the complex numbers. If $A^r = I = B^s$, where every prime divisor of rs is strictly greater than $2n+1$, then either $AB = BA$ or the group generated by A and B is infinite.*

For any subset S of a group G, $C_G(S)$, $N_G(S)$, $|S|$ will mean respectively the centralizer, normalizer and number of elements in S. For any complex valued functions ζ, ξ on G we define

$$(\zeta, \xi)_G = \frac{1}{|G|} \sum_G \zeta(x)\overline{\xi(x)} ,$$

and $\|\zeta\|_G^2 = (\zeta, \zeta)_G$. Whenever it is clear from the context which group is involved, the subscript G will be omitted. $H \lhd G$ will mean that H is a normal subgroup of G. For any two subsets A, B of G, $A - B$ will denote the set of all elements in A which are not in B. If a subgroup of a group is the kernel of a representation, then we will also say that it is the kernel of the character of the given representation. All groups

Received November 25, 1960. The first author was partly supported by O. O. R. and an NSF Grant. The second author was partly supported by the Esso Research Foundation.

considered are assumed to be finite.

2. Proof of Theorem 1. We will first prove the following preliminary result.

LEMMA 1. *Assume that the Sylow p-group P of N is a normal subgroup of N. If x is any element of N such that $C_N(x) \cap P = \{1\}$, then $\lambda(x) = 0$ for any irreducible character λ of N which does not contain P in its kernel.*

Proof. Since $|C_N(x)|$ is not divisible by p, it is easily seen that $C_N(x)$ is mapped isomorphically into $C_{N/P}(\bar{x})$, where $\bar{x}$ denotes the image of x in N/P under the natural projection. Let $\mu_1, \mu_2, \cdots$ be all the irreducible characters of N which contain P in their kernel and let $\lambda_1, \lambda_2, \cdots$ be all the other irreducible characters of N. The orthogonality relations yield that

$$\sum_i |\mu_i(x)|^2 = |C_{N/P}(\bar{x})| \geqq |C_N(x)| = \sum_i |\mu_i(x)|^2 + \sum_i |\lambda_i(x)|^2 .$$

This implies the required result.

From now assume that G is a counter example to Theorem 1 of minimal order. We will show that p^2 does not divide $|G|$, then Brauer's theorem may be applied to complete the proof. The proof is given in a series of short steps.

Clearly every subgroup of G satisfies the assumption of Theorem 1, hence we have

(I) *The Sylow p-group of any proper subgroup H of G is an abelian normal subgroup of H.*

Let P be a fixed Sylow p-group of G. Let Z be the center of G.

(II) *P is abelian.*

As P has a faithful representation of degree $n < p$, each irreducible constituent of this representation has degree one. Therefore in completely reduced form, the representation of P consists of diagonal matrices. Consequently these matrices form an abelian group which is isomorphic to P.

(III) *G contains no proper normal subgroup whose index in G is a power of p.*

Suppose this is false. Let H be a normal subgroup of G of minimum order such that $[G:H]$ is a power of p. Let P_0 be a Sylow p-group of H. By (I) $P_0 \lhd H$, hence $P_0 \lhd G$. Thus $C_G(P_0) \lhd G$. If $C_G(P_0) \neq G$, then by (I) and (II), $P \lhd C_G(P_0)$, thus $P \lhd G$ contrary to assumption. Therefore $C_G(P_0) = G$. Burnside's Theorem ([3], p. 203) implies that H contains a normal p-complement which must necessarily be normal in G. The minimal nature of H now yields that p does not divide $|H|$.

If q is any prime dividing $|H|$, then it is a well known consequence of the Sylow theorems that it is possible to find a Sylow q-group Q of H such that $P \subseteqq N(Q)$. Hence PQ is a solvable group which satisfies the hypotheses of Theorem 1. Ito's Theorem [4] now implies that $P \lhd PQ$, thus $Q \subseteqq N(P)$. As q was an arbitrary prime dividing $|H|$, we get that $|H|$ divides $|N(P)|$. Consequently $N(P) = G$, contrary to assumption.

(IV) *Z is the unique maximal normal subgroup of G. G/Z is a noncyclic simple group. $|Z|$ is not divisible by p.*

Let H be a maximal normal subgroup of G, hence G/H is simple. Let P_0 be a Sylow p-group of H. Then by (I) $P_0 \lhd H$, hence $P_0 \lhd G$, thus $C(P_0) \lhd G$. If $C(P_0) \neq G$, then by (I) and (II) $P \lhd C(P_0)$, hence $P \lhd G$ contrary to assumption. Therefore $C(P_0) = G$. If $P_0 \neq \{1\}$, then it is a simple consequence of Grün's Theorem ([3], p. 214) that G contains a proper normal subgroup whose index is a power p. This contradicts (III). Hence $P_0 = \{1\}$ and p does not divide $|H|$.

By (III) $PH \neq G$, hence by (I) $P \lhd PH$. Consequently $PH = P \times H$, and $P \subseteqq C(H) \lhd G$. If $C(H) \neq G$, then (I) yields that $P \lhd C(H)$. Hence once again $P \lhd G$, contrary to assumption. Consequently $C(H) = G$. Therefore $H \subseteqq Z$. As G is not solvable, neither is G/H. Now the maximal nature of H yields that $H = Z$ and suffices to complete the proof.

(V) $P \cap xPx^{-1} = \{1\}$ *unless x is in $N(P)$.*

Let $D = P \cap xPx^{-1}$ be a maximal intersection of Sylow p-groups of G. Then P is not normal in $N(D)$. Hence by (I) $N(D) = G$, or $D \lhd G$. However (IV) now implies that $D \subseteqq Z$. Hence (IV) also yields that $D = \{1\}$ as was to be shown.

Define the subset N_0 of $N(P)$ by

$$N_0 = \{x \mid x \in N(P),\, C(x) \cap P \neq \{1\}\} .$$

Clearly $\{P, Z\} \subseteqq N_0$.

(VI) $N(N_0) = N(P)$. $(N_0 - Z) \cap x(N_0 - Z)x^{-1}$ *is empty unless $x \in N(P)$.*

Clearly $N(P) \subseteqq N(N_0)$. Since P consists of all elements in N_0 whose order is a power of p, it follows that $N(N_0) \subseteq N(P)$.

Suppose $y \in (N_0 - Z) \cap x(N_0 - Z)x^{-1}$. Then y and $x^{-1}yx$ are both contained in $(N_0 - Z)$. Let $P_0 = C(y) \cap P$, $P_1 = C(x^{-1}yx) \cap P$. By assumption $P_0 \neq \{1\} \neq P_1$. It follows from the definitions that P_0 and xP_1x^{-1} are both contained in $C(y)$. Since y is not in Z, $C(y) \neq G$. Hence (I) yields that P_0 and xP_1x^{-1} generate a p-group. Thus by (II) $xP_1x^{-1} \subseteq C(P_0)$. Now (V) implies that $xP_1x^{-1} \subseteq N(P)$. Consequently $xP_1x^{-1} \subseteq P$. By (V), this yields that $x \in N(P)$ as was to be shown.

From now on we will use the following notation:

$$| P | = p^e , \quad | Z | = z , \quad | N(P) | = p^e zt .$$

Let $\chi_0 = 1, \chi_1, \cdots$ be all the irreducible characters of G. Define α_i, β_i, b_i by

$$\chi_{i_{|N(P)}} = \alpha_i + \beta_i , \qquad b_i = \beta_i(1)$$

where α_i is a sum of irreducible characters of $N(P)$, none of which contain P in their kernel and β_i is a character of $N(P)$ which contains P in its kernel.

(VII) *If* $i \neq 0$, *then* $b_i < (1/p^{e/2})\ \chi_i(1)$.

By (VI) $(N_0 - Z)$ has $| G |/p^e zt$ distinct conjugates and no two of them have any elements in common. Since χ_i is a class function on G, this yields that

$$\begin{aligned} 1 = \| \chi_i \|^2 &> \frac{1}{| G |} \frac{| G |}{p^e zt} \Sigma_{(N_0 - Z)} | \chi_i(x) |^2 \\ &= \frac{1}{p^e zt} \{ -\Sigma_Z | \chi_i(x) |^2 + \Sigma_{N_0} | \alpha_i(x) + \beta_i(x) |^2 \} . \end{aligned}$$

If $x \in Z$, then $| \chi_i(x) |^2 = | \chi_i(1) |^2$. As $P \subseteq N_0$, we get that

$$1 > \frac{1}{p^e zt} [- | \chi_i(1) |^2 z + \Sigma_{N_0} \{ | \alpha_i(x) |^2 + \alpha_i(x)\overline{\beta_i(x)} + \overline{\alpha_i(x)}\beta_i(x) \} + \Sigma_{PZ} | \beta_i(x) |^2] .$$

Since P is in the kernel of β_i, we get that $| \beta_i(x) | = b_i$ for $x \in PZ$. Lemma 1 implies that α vanishes on $N(P) - N_0$. Hence

$$1 > \frac{- | \chi_i(1) |^2}{p^e t} + \| \alpha_i \|^2_{N(P)} + (\alpha_i, \beta_i)_{N(P)} + \overline{(\alpha_i, \beta_i)}_{N(P)} + \frac{b_i^2}{t} .$$

By definition $(\alpha_i, \beta_i) = 0$, hence

$$\frac{| \chi_i(1) |^2}{p^e t} > \| \alpha_i \|^2_{N(P)} - 1 + \frac{b_i^2}{t} .$$

By (IV) the normal subgroup generated by P is all of G, hence $\alpha_i \neq 0$.

Therefore $\|\alpha_i\|^2_{N(P)} \geqq 1$. This finally yields that

$$\frac{|\chi_i(1)|^2}{p^e t} > \frac{b_i^2}{t},$$

which is equivalent to the statement to be proved.

(VIII) *If Γ is the character of G induced by the trivial character 1_P of P, then $(\Gamma, \chi_i) = b_i$.*

If λ is an irreducible character of $N(P)$ which does not contain P in its kernel, then λ is not a constituent of the character of $N(P)$ induced by 1_P. Hence by the Frobenius reciprocity theorem $(\lambda_{|P}, 1_P)_P = 0$. Consequently $(\alpha_{i|P}, 1_P)_P = 0$. The Frobenius reciprocity theorem now implies that

$$(\chi_i, \Gamma) = (\chi_{i|P}, 1_P)_P = (\beta_{i|P}, 1_P) = b_i .$$

From now on let χ be an irreducible character of minimum degree greater than one. Define the integers a_i by

$$a_i = (\chi_i, \chi\bar{\chi}) .$$

(IX) $\chi(1) - 1 \leqq \sum_{i\neq 0} a_i b_i$.

By (VIII)

$$\begin{aligned} a_0 b_0 + \sum_{i\neq 0} a_i b_i = (\Gamma, \chi\bar{\chi}) &= \frac{\chi(1)^2}{p^e} + \frac{1}{p^e z t} \Sigma_{P-\{1\}} z t \chi\bar{\chi}(x) \\ &= \frac{1}{p^e} \Sigma_P \chi\bar{\chi}(x) = \|\chi_{|P}\|^2_P . \end{aligned}$$

By (II), $\chi_{|P}$ is a sum of $\chi(1)$ linear characters of P. Consequently

$$a_0 b_0 + \sum_{i\neq 0} a_i b_i \geqq \chi(1) .$$

As χ is irreducible, $a_0 = 1$. Clearly $b_0 = 1$. This yields the desired inequality.

We will now complete the proof of Theorem 1.

It follows from (IX) that

$$\chi(1) - 1 \leqq \sum_{i\neq 0} a_i b_i .$$

(VII) yields that

$$\sum_{i\neq 0} a_i b_i < \frac{1}{p^{e/2}} \sum_{i\neq 0} a_i \chi_i(1) .$$

The definition of the integers a_i implies that

$$\sum_{i\neq 0} a_i\chi_i(1) = \chi(1)^2 - 1 .$$

Combining these inequalities we get that

$$\chi(1) - 1 < \frac{\chi(1)^2 - 1}{p^{e/2}} ,$$

or

$$p^{e/2} < \chi(1) + 1 .$$

By assumption $\chi(1) < (p-1)/2$, hence

$$p^{e/2} < \chi(1) + 1 < p .$$

This implies that $e < 2$. Thus $e \leqq 1$.

R. Brauer's theorem [2] now yields that $P \lhd G$ contrary to assumption. This completes the proof of Theorem 1.

BIBLIOGRAPHY

1. H. F. Blichfeldt, *On the order of linear homogeneous groups*, Trans. Amer. Math. Soc., **4** (1903), 387-397.
2. R. Brauer, *On groups whose order contains a prime number to the first power* II, Amer. J. of Math., **64** (1942), 421-440.
3. M. Hall, *The Theory of Groups*, The MacMillan Company, New York, 1959.
4. N. Ito, *On a theorem of H. F. Blichfeldt*, Nagoya Math. J., **5** (1954), 75-77.

CORNELL UNIVERSITY AND THE UNIVERSITY OF CHICAGO
HARVARD UNIVERSITY AND THE UNIVERSITY OF CHICAGO

MEAN CROSS-SECTION MEASURES OF HARMONIC MEANS OF CONVEX BODIES

WILLIAM J. FIREY

1. In [2] the notion of p-dot means of two convex bodies in Euclidean n-space was introduced and certain properties of these means investigated. For $p = 1$, the mean is more appropriately called the harmonic mean; here we restrict the discussion to this case. The harmonic mean of two convex bodies K_0 and K_1, which will always be assumed to share a common interior point Q, is defined as follows. Let $\hat{K}$ denote the polar reciprocal of K with respect to the unit sphere E centred at Q; let $(1-\vartheta)\hat{K}_0 + \vartheta\hat{K}_1$, with $0 \leqq \vartheta \leqq 1$, be the usual arithmetic or Minkowski mean of $\hat{K}_0$ and $\hat{K}_1$. The harmonic mean of K_0, K_1 is the convex body $[(1-\vartheta)\hat{K}_0 + \vartheta\hat{K}_1]^\wedge$. In more analytic terms, if $F_i(x)$ are the distance functions with respect to Q of K_i, for $i = 0, 1$, then the body whose distance function with respect to Q is $(1-\vartheta)F_0(x) + \vartheta F_1(x)$ is the harmonic mean of K_0 and K_1.

In the paper mentioned, a dual Brunn-Minkowski theorem was established, namely

$$V^{1/n}([(1-\vartheta)\hat{K}_0 + \vartheta\hat{K}_1]^\wedge) \leqq 1\Big/\left[\frac{(1-\vartheta)}{V^{1/n}(K_0)} + \frac{\vartheta}{V^{1/n}(K_1)}\right] \tag{1}$$

where $V(K)$ means the volume of K. There is equality if and only if K_0 and K_1 are homothetic with the centre of magnification at Q.

Here we develop a more inclusive theorem regarding the behaviour of each mean cross-section measure, ("Quermassintegral") $W_\nu(K)$, $\nu = 0, 1, \cdots, n-1$, cf. [1]. The result is

$$W_\nu^{1/(n-\nu)}([(1-\vartheta)\hat{K}_0 + \vartheta\hat{K}_1]^\wedge) \leqq 1\Big/\left[\frac{(1-\vartheta)}{W_\nu^{1/(n-\nu)}(K_0)} + \frac{\vartheta}{W_\nu^{1/(n-\nu)}(K_1)}\right]. \tag{2}$$

The cases of equality are just those of the dual Brunn-Minkowski theorem, ($\nu = 0$).

2. We first list some preliminary items used in the proof of (2). We shall use Minkowski's inequality in the form

$$\int[(1-\vartheta)f_0^p + \vartheta f_1^p]^{1/p}dx \leqq \left[(1-\vartheta)\left(\int f_0 dx\right)^p + \vartheta\left(\int f_1 dx\right)^p\right]^{1/p}. \tag{3}$$

Here the functions f_i are assumed to be positive and continuous over the closed and bounded domain of integration common to all the integrals,

Received September 29, 1960.

and, for our puposes, p satisfies $-1 \leqq p < 0$. There is equality if and only if $f_0(x) \equiv \lambda f_1(x)$ for some constant λ. See [3], Theorem 201, coupled with the remark preceding Theorem 200.

Our second tool, which we shall refer to as the projection lemma, was established in [2]. Let K^* denote the projection of K onto a fixed, m-dimensional, linear subspace E_m through Q for $1 \leqq m < n$. We have

$$(4) \qquad [(1-\vartheta)\hat{K}_0^* + \vartheta\hat{K}_1^*]\hat{} \supseteqq \{[(1-\vartheta)\hat{K}_0 + \vartheta\hat{K}_1]\hat{}\}^* .$$

Since E_m contains Q and the polar reciprocation is with respect to sphere E centred at Q, in forming $\hat{K}^*$ the order of operations is immaterial. This result is proved by a polar reciprocation argument from

$$(1-\vartheta)(\hat{K}_0 \cap E_m) + \vartheta(\hat{K}_1 \cap E_m) \subseteqq [(1-\vartheta)\hat{K}_0 + \vartheta\hat{K}_1] \cap E_m .$$

There is equality in either inclusion if K_0 and K_1 are homothetic with centre of magnification at Q.

The dual Brunn-Minkowski theorem (1) will be used.

Finally we shall make use of Kubota's formula and some of its consequences. This material is covered in [1]. An $(n-\nu)$ dimensional cross-section measure ("Quermass") of K is the $(n-\nu)$ dimensional volume of that convex body which is the vertical projection of K onto an $E_{n-\nu}$. The mean cross-section measures are usually defined as the coefficients in Steiner's polynomial which describes $V(K+\lambda E)$, that is

$$(5) \qquad V(K+\lambda E) = \sum_{\nu=0}^{n} \binom{n}{\nu} W_\nu(K)\lambda^\nu .$$

If we denote the $(\nu-1)^{\text{th}}$ mean cross-section measure of the projection of K onto that E_{n-1} through Q which is orthogonal to the vector u_1 by $W'_{\nu-1}(K, u_1)$, then Kubota's formula is

$$W_\nu(K) = \frac{1}{\kappa_{n-1}}\int_{\Omega_n} W'_{\nu-1}(K, u_1)d\omega_n , \qquad \nu = 1, 2, \cdots, \nu - 1 .$$

Here the integration with respect to the direction u_1 is extended over the surface Ω_n of E, $d\omega_n$ is the element of surface area on Ω_n and κ_{n-1} is the volume of the $n-1$ dimensional unit sphere.

Kubota's formula can be applied to the mean cross-section measure $W'_{\nu-1}(K, u_1)$ for fixed u_1:

$$W'_{\nu-1}(K, u_1) = \frac{1}{\kappa_{n-2}}\int_{\Omega_{n-1}} W''_{\nu-2}(K, u_1, u_2)d\omega_{n-1}$$

where $W''_{\nu-2}$ is the $(\nu-2)$th mean cross-section measure of the projection of κ onto the E_{n-2} through Q orthogonal to u_1 and u_2 with u_2 orthogonal to u_1. After ν such steps we have as the extended form of Kubota's formula:

$$W_\nu(K) = \frac{1}{\kappa_{n-1}\kappa_{n-2}\cdots\kappa_{n-\nu}}\int_{\Omega_n}\int_{\Omega_{n-1}}\cdots\int_{\Omega_{n-\nu}} W_0^{(\nu)}(K, u_1, u_2, \cdots, u_\nu)d\omega_{n-\nu}\cdots d\omega_{n-1}d\omega_n .$$

Each vector u_p is orthogonal to u_q for $q < p$ and $W_0^{(\nu)}(K, u_1, u_2, \cdots, u_\nu)$ is the 0th mean cross-section measure of the projection of K onto that $E_{n-\nu}$ through Q which is the orthogonal complement of the subspace spanned by $u_1, u_2, \cdots, u_\nu$.

Steiner's formula (5) with $\lambda = 0$ shows that $W_0(K)$ is the volume of K and so $W_0^{(\nu)}$ is an $(n - \nu)$ dimensional cross-section measure of K. Thus, to within a numerical factor depending on n and ν, $W_\nu(K)$ is the arithmetic mean of the $(n - \nu)$ dimensional cross-section measures.

In § 3 we shall use the following abbreviations: for $d\omega_{n-\nu}\cdots d\omega_{n-1}d\omega_n$ we write $d\bar{\omega}$ with sign of integration and omit reference to the domains of integration; for one $1/\kappa_{n-1}\kappa_{n-2}\cdots\kappa_{n-\nu}$ we write k; finally for $W_0^{(\nu)}(K, u_1, u_2, \cdots, u_\nu)$ we write $\sigma(K^*)$. In this notation the extended Kubota formula reads

$$W(K) = k\int\sigma(K^*)d\bar{\omega} .$$

3. We now prove (2). By the extended form of Kubota's formula

$$(6)\qquad \begin{aligned} W_\nu^{1/(n-\nu)}([(1-\vartheta)\hat{K}_0 + \vartheta\hat{K}_1]^\wedge) &= \Big[k\int\sigma(\{[(1-\vartheta)\hat{K}_0 + \vartheta\hat{K}_1]^\wedge\}^*)d\bar{\omega}\Big]^{1/(n-\nu)} \\ &\leqq \Big[k\int\sigma([(1-\vartheta)\hat{K}_0^* + \vartheta\hat{K}_1^*]^\wedge)d\bar{\omega}\Big]^{1/(n-\nu)} \end{aligned}$$

in virtue of the projection lemma and the set monotonicity of σ i.e., $\sigma(K^*) \leqq \sigma(\bar{K}^*)$ if $K^* \subseteqq \bar{K}^*$ with equality in the latter relation implying that in the former. We now apply (1), in $E_{n-\nu}$, to the integrand to obtain

$$\sigma([(1-\vartheta)\hat{K}_0^* + \vartheta\hat{K}_1^*]^\wedge) \leqq \Big\{1\Big/\Big[\frac{(1-\vartheta)}{\sigma^{1/(n-\nu)}(K_0^*)} + \frac{\vartheta}{\sigma^{1/(n-\nu)}(K_1^*)}\Big]\Big\}^{(n-\nu)} .$$

Here we take advantage of the fact that

$$(\hat{K})^* = (K^*)^\wedge .$$

This gives

$$(7)\qquad \begin{aligned} &W_\nu^{1/(n-\nu)}([(1-\vartheta)\hat{K}_0 + \vartheta\hat{K}_1]^\wedge) \\ &\leqq \Big[k\int\Big\{1\Big/\Big[\frac{(1-\vartheta)}{\sigma^{1/(n-\nu)}(K_0^*)} + \frac{\vartheta}{\sigma^{1/(n-\nu)}(K_1^*)}\Big]\Big\}^{(n-\nu)}d\bar{\omega}\Big]^{1/(n-\nu)} . \end{aligned}$$

There is equality if and only if all the projections K_0^* and K_1^* are homothetic with the centre of magnification at Q. This condition is sufficient for equality in (6); it is necessary and sufficient for (7).

We now use Minkowski's inequality (3) with $p = -1/n-\nu$. This yields

$$W_\nu^{1/(n-\nu)}(([1-\vartheta)\hat{K}_0 + \vartheta\hat{K}_1)$$
$$\leqq 1\Bigg/\left[\frac{(1-\vartheta)}{\left(k\int\sigma(K_0^*)d\bar{\omega}\right)^{1/(n-\nu)}} + \frac{\vartheta}{\left(k\int\sigma(K_1^*)d\bar{\omega}\right)^{1/(n-\nu)}}\right]$$
$$= 1\Bigg/\left[\frac{(1-\vartheta)}{W_\nu^{1/(n-\nu)}(K_0)} + \frac{\vartheta}{W_\nu^{1/(n-\nu)}(K_1)}\right].$$

The necessary and sufficient conditions for equality in (7) are sufficient for equality in (3) since $K_0 = \lambda K_1$ implies $\sigma(K_0^*) = \lambda^{n-\nu}\sigma(K_1^*)$. This establishes (2).

References

1. T. Bonnesen and W. Fenchel, *Konvexe Körper*, Berlin, 1934, reprint N. Y. (1948), 48-50.
2. W. J. Firey, *Polar Means of Convex Bodies and a Dual to the Brunn-Minkowski theorem.* Canadian Math. J., **13** (1961), 444-453.
3. G. Hardy, J. Littlewood, and G. Pólya, *Inequalities*, Cambridge, (1934), 148.

WASHINGTON STATE UNIVERSITY

THE WAVE EQUATION FOR DIFFERENTIAL FORMS

AVNER FRIEDMAN

1. The Problem. Let M be a compact C^∞ Riemannian manifold of dimension N, having a positive definite metric. The operator $\Delta = d\delta + \delta d$ (see [13] for notation) maps p-forms $(0 \leqq p \leqq N)$ into p-forms and it reduces, when $p = 0$, to minus the Laplace-Beltrami operator. Let $c(P)$ be a C^∞ function which is nonpositive for $P \in M$, and consider the Cauchy problem of solving the system

$$\left(L + \frac{\partial^2}{\partial t^2}\right)v \equiv \left(\Delta + c + \frac{\partial^2}{\partial t^2}\right)v = f(P, t) \tag{1.1}$$

$$v(P, 0) = g(P) , \qquad \frac{\partial}{\partial t} v(P, 0) = h(P) , \tag{1.2}$$

where f, g, h are C^∞ forms of degree p. The main purpose of the present paper is to solve the system (1.1), (1.2) by the method of Fourier.

The Cauchy problem for second order self-adjoint hyperbolic equations was solved by Fourier's method by Ladyzhenskaya [8] and more recently (with some improvements) by V. A. Il'in [6]. In [8], other methods are also described, namely: finite differences, Laplace transforms, and analytic approximations using a priori inequalities. Higher order hyperbolic equations were treated by Petrowski [12], Leray [9] and Garding [5].

The Fourier method can be based on the fact that the series

$$\sum_{\lambda_n > 0} \frac{|\varphi_n(x)|^2}{\lambda_n^\alpha}, \sum_{\lambda_n > 0} \frac{|\partial\varphi_n(x)/\partial x|^2}{\lambda_n^{\alpha+1}}, \sum_{\lambda_n > 0} \frac{|\partial^2\varphi_n(x)/\partial x^2|^2}{\lambda_n^{\alpha+2}} \tag{1.3}$$

are uniformly convergent. Here $\{\varphi_n\}$ and $\{\lambda_n\}$ are the sequences of eigenfunctions and eigenvalues of the elliptic operator appearing in the hyperbolic equation. In [6] the convergence of (1.3) is proved for $\alpha = [N/2] + 1$. Our proof of the analogous result for eigenforms is different from that of [6] and yields a better (and sharp) value for α, namely, $\alpha = N/2 + \varepsilon$ for any $\varepsilon > 0$. It is based on asymptotic formulas which we derive for $\sum_{\lambda_n \leqq \lambda} |\partial^j\varphi_n(x)/\partial x^j|^2$ as $\lambda \to \infty$.

In § 2 we recall various definitions and introduce the fundamental solution for $L + \partial/\partial t$ which was constructed by Gaffney [4] in the case $c(P) \equiv 0$. In § 3 we derive some properties of the fundamental solution. These properties are used in § 4 to derive the asymptotic formulas for $\sum_{\lambda_n \leqq \lambda} |\partial^j\varphi_n(x)/\partial x^j|^2$, by which the convergence of the series in (1.3) for any $\alpha > N/2$ follows. In § 5 we solve the problem (1.1), (1.2); first for f, g, h

Received January 13, 1961. Prepared under Contract Nonr 710 (16) (NR 044 004) between the Office of Naval Research and the University of Minnesota.

infinitely differentiable and then under much weaker differentiability assumptions with regard to M, c, f, g, h. In § 6 we briefly treat the Cauchy problem for the parabolic system

$$Lu + \frac{\partial u}{\partial t} = f(P, t) \tag{1.4}$$

$$u(P, 0) = g(P) . \tag{1.5}$$

2. Preliminaries. The first one to use fundamental solutions of the heat equation in the study of the asymptotic distributions of eigenvalues and eigenfunctions was Minakshisundaram [11]. Gaffney [4] extended his method to derive asymptotic formulas for eigenvalues and eigenforms. We shall describe here some well known facts and some of the results of [4] which we will need later on. Slight modifications will be made due to the fact that in [4] $c \equiv 0$.

As is well known, there exists a sequence of eigenvalues $\{\lambda_n\}$ $(0 \leqq \lambda_1 \leqq \cdots \leqq \lambda_k \to \infty$ as $k \to \infty)$ and a sequence of the corresponding eigenforms $\{\omega_n\}$ of degree p $(0 \leqq p \leqq N, p$ is fixed throughout the paper) of L, that is, $L\omega_n = \lambda_n \omega_n$, such that the eigenforms form a complete orthonormal set in $L_p^2(M)$ (square integrable p-forms on M). The $\omega_i(p)$ are C^∞ forms. The fundamental solution $\Theta(P, Q, t)$ of

$$\left(L + \frac{\partial}{\partial t}\right)\omega = 0 \tag{2.1}$$

is a double p-form which is twice differentiable in Q, once differentiable in t, satisfies (2.1) in (Q, t), $Q \in M$, $t > 0$, (for any fixed P) and, for any $P \in M$,

$$\lim_{t \to 0} \int_M \Theta(P, Q, t) * \alpha(Q) = \alpha(P) \tag{2.2}$$

for any L^2 p-form α which is continuous at P. As in [4] one easily derives the expansion (provided Θ is known to exist)

$$\Theta(P, Q, t) = \sum_{i=1}^{\infty} \omega_i(P)\omega_i(Q)e^{-\lambda_i t} \tag{2.3}$$

where the series on the right is pointwise convergent for all $P, Q \in M$, $t > 0$ (that is, the series of each component is pointwise convergent).

A p-form α can be written locally as

$$\alpha = \sum_{i_1 < \cdots < i_p} A_{i_1 \cdots i_p} dx^{i_1} \cdots dx^{i_p} = \Sigma' A_I dx^I$$

where $'$ indicates summation on $I = (i_1, \cdots, i_p)$ with $i_1 < \cdots < i_p$. The absolute value of α at P is given by

$$|\alpha(P)| = [\Sigma' A_I(x) A^I(x)]^{1/2}$$

where x is the local coordinate of P. Similarly, for a double p-form having local representation $\alpha(P, Q) = \Sigma' A_{IJ}(x, y) dx^I dy^J$ where y is the local coordinate of Q, we define the absolute value by

$$|\alpha(P, Q)| = [\sum_{I,J}' A_{IJ}(x, y) A^{IJ}(x, y)]^{1/2} .$$

The right "half-norm" is defined by

$$|\alpha||(P) = \left[\int_M |\alpha(P, Q)|^2 dV_Q\right]^{1/2} .$$

Given two double p-forms α and β, a new double p-form is defined by

$$[\alpha, \beta] = [\alpha, \beta](P, Q) = \int_M \alpha(P, W) * \beta(Q, W) .$$

One then verifies:

$$|[\alpha, \beta](P, Q)| \leqq |\alpha||(P)|\beta||(Q) . \tag{2.4}$$

The following inequalities are immediate:

$$|\alpha + \beta| \leqq |\alpha| + |\beta|, |\alpha + \beta|| \leqq |\alpha|| + |\beta|| , \tag{2.5}$$

where α, β are any double p-forms.

In order to construct Θ, one first constructs a parametrix. Gaffney [4] constructs a parametrix by generalizing the method of Minakshisandaram [11], making use of some calculation of Kodaira [7]. Given a point P, let $y = (y^i)$ be normal coordinates about P (with coordinates x^i). A p-form can be written as a vector X with $\binom{N}{p}$ components and then

$$\Delta X = -\Sigma g^{ij}\partial_i\partial_j X + \Sigma A^i \partial_i X + BX \tag{2.6}$$

where (g_{ij}) is the metric tensor, (g^{ij}) is the inverse matrix, $\partial_i = \partial/\partial x^i$, and A^i, B are matrices depending on the g_{ij} and their first two derivatives. If $X = f(r^2) W(x, y)$ where r is the geodesic distance from x to y (each component of X is now a vector so that W is a square matrix), then

$$\Delta_y[f(r^2) W] = f(r^2)\Delta_y W - f'(r^2)\left\{2N - 4K + 4r\frac{\partial}{\partial r}\right\} W - 4r^2 f''(r^2) W , \tag{2.7}$$

where $K = K(x, y)$ is a C^∞ matrix which vanishes for $y = x$.

There exists a C^∞ matrix M satisfying

$$r\frac{\partial}{\partial r} M = KM \ (x \text{ fixed}), \qquad M(x, x) = I \tag{2.8}$$

where I is the identity matrix. Using (2.8), (2.7) is simplified to

$$(2.9)\qquad M^{-1}\Delta_y(fMW) = f(M^{-1}\Delta M)_y W - f'\left\{2N + 4r\frac{\partial}{\partial r}\right\}W - 4r^2 f'' W .$$

(2.9) will now be applied with

$$f(r^2, t) = \frac{1}{(4\pi t)^{N/2}} e^{-r^2/4t} \qquad (t > 0 \text{ fixed}) .$$

Setting

$$H_m = \sum_{j=0}^{m} fMU_j t^j, \qquad U_0 = I$$

one then gets

$$\Delta H_\infty = fM\sum_{j=0}^{\infty}\left\{(M^{-1}\Delta M)U_j t^j + \frac{1}{4t}\left(2N + 4r\frac{\partial}{\partial r}\right)U_j t^j - \frac{r^2}{4t^2}U_j t^j\right\} .$$

Calculating also $\partial H_\infty/\partial t$, one then obtains

$$\left(L_y + \frac{\partial}{\partial t}\right)H_\infty = fM\sum_{j=0}^{\infty}\left\{(M^{-1}\Delta M + c)U_j + \left(r\frac{\partial}{\partial r} + j + 1\right)U_{j+1}\right\}t^j$$

which leads to the successive definitions:

$$(2.10)\qquad U_j = -\frac{1}{r^j}\int_0^r (M^{-1}\Delta M + c)U_{j-1}dr\ (1 \leqq j < \infty), \quad \text{where} \quad U_0 = I .$$

We conclude that, for any $m \geqq 0$,

$$(2.11)\qquad \left(L_y + \frac{\partial}{\partial t}\right)H_m = \frac{1}{(4\pi)^{N/2}} e^{-r^2/4t} t^{m-N/2} L_y(MU_m) .$$

H_m is a local parametrix. Note that when P, Q vary in a sufficiently small neighborhood V (contained in one coordinate patch), H_m is defined and is C^∞ in (P, Q, t) if $t > 0$. Let $\eta_\varepsilon(r)$ be a C^∞ function of r which is equal to 1 for $r < \varepsilon$ and is equal to 0 for $r > 2\varepsilon$. If ε is sufficiently small then the support of $\eta_\varepsilon(r)H_m(P, Q, t)$ (where r is the distance from P to Q) as a form in Q lies in V, provided $P \in W$, where W is a given open subset of V, $\bar{W} \subset V$. We can cover the manifold M by a finite number of sets W, call then W_i. Let the H_m corresponding to (the corresponding) V_i be denoted by H_m^i. If $\{\alpha_i\}$ is a C^∞ partition of unity subordinate to $\{W_i\}$, then the support of $\alpha_i(P)\eta_\varepsilon(r)H_m^i(P, Q, t)$ as a form of (P, Q) lies in $W_i \times V_i$ and hence this form is C^∞ in (P, Q, t) if $t > 0$.

The global parametrix is given by

$$(2.12)\qquad \Theta_m(P, Q, t) = \Sigma\alpha_i(P)\eta_\varepsilon(r)H_m^i(P, Q, t) .$$

The fundamental solution should then formally be

$$\Theta(P, Q, t) = \Theta_m(P, Q, t) + \int_0^t [\gamma_m(P, U, t), \Theta_m(Q, U, t - \tau)] d\tau \tag{2.13}$$

where γ_m is defined by

$$\gamma_m(P, Q, t) = \sum_{i=1}^{\infty} (-1)^i \delta_m^i(P, Q, t) \tag{2.14}$$

$$\delta_m^i(P, Q, t) = \int_0^t [\delta_m^{i-1}(P, U, \tau), \delta_m^i(Q, U, t - \tau)] d\tau, \tag{2.15}$$

$$\delta_m^1 = \left(L_y + \frac{\partial}{\partial t}\right)\Theta_m .$$

Using (2.4) and the inequality

$$\left| \int_0^t \alpha(P, Q, \tau) d\tau \right| \leqq \binom{N}{p} \int_0^t |\alpha| \, d\tau , \tag{2.16}$$

Gaffney establishes the uniform convergence of the right side of (2.14) and then proves that Θ, as defined in (2.13), is a fundamental solution, for any $m \geqq 0$, written in matrix form. We shall use the matrix notation of Θ and the usual double form notation for Θ interchangably; the same for Θ_m.

3. **Properties of the fundamental solution.** We denote by $\partial_P^h \Theta(P, Q, t)$ an hth derivative of Θ with respect to the coordinates of P, in a given coordinate system. If $h = (h_1, \cdots, h_N)$, set $|h| = h_1 + \cdots + h_N$. From the formulas defining Θ it is clear that $\partial_P^h \Theta(P, Q, t)$ exists and is continuous (in fact C^∞) in $P, Q \in M$ and $t > 0$. Let

$$\partial_P^h \Theta(P, Q, t) \sim \sum_{i=1}^{\infty} B_i(P, t) \omega_i(Q) \tag{3.1}$$

be the Fourier expansion of $\partial_P^h \Theta$, for (P, t) fixed. Then (recalling (2.3))

$$\begin{aligned} B_i(P, t) &= \int_M \partial_P^h \Theta(P, U, t) * \omega_i(U) = \partial_P^h \int_M \Theta(P, U, t) * \omega_i(U) \\ &= \partial^h \omega_i(P) e^{-\lambda_i t} , \end{aligned} \tag{3.2}$$

where ∂_P^h is abbreviated by ∂^h when there is no confusion.

By the (easily verified) Parseval's equality we get

$$\begin{aligned} \psi(P, Q, t) &\equiv \left[\partial_P^h \Theta\left(P, U, \frac{t}{2}\right), \partial_Q^h \Theta\left(Q, U, \frac{t}{2}\right)\right] \\ &= \sum_{i=1}^{\infty} \partial_P^h \omega_i(P) \partial_Q^h \omega_i(Q) e^{-\lambda_i t} \end{aligned} \tag{3.3}$$

and the series is pointwise convergent for $P, Q \in M, t > 0$.

We need the following notations. Let α be a double p-form. If it is locally represented by $\Sigma' A_{IJ} dx^I dy^J$, then we set

$$[\alpha(P, P)] = \Sigma' A_I^I .$$

If β is also a double p-form, then we define $[[\alpha(P, U), \beta(P, U)]_U]$ to be $[\gamma(P, P)]$ where $\gamma(P, Q) = [\alpha(P, U), \beta(Q, U)]$.

Using (2.13) and the definition of ψ in (3.3) we have

$$\begin{aligned}(3.4)\quad \sum_{i=1}^{\infty} | \partial^h \omega_i(P) |^2 e^{-\lambda_i t} &= [\psi(P, P, t)] \\ &= \Big[\Big[\partial_P^h \Theta_m\Big(P, U, \frac{t}{2}\Big), \partial_P^h \Theta_m\Big(P, U, \frac{t}{2}\Big)\Big]_U\Big] \\ &+ 2\Big[\Big[\int_0^{t/2}\Big[\partial_P^h \gamma_m(P, W, \tau), \Theta_m\Big(U, W, \frac{t}{2} - \tau\Big)\Big] d\tau, \partial_P^h \Theta_m\Big(P, U, \frac{t}{2}\Big)\Big]_U\Big] \\ &+ \Big[\Big[\int_0^{t/2}\Big[\partial_P^h \gamma_m(P, W, \tau), \Theta_m\Big(U, W, \frac{t}{2} - \tau\Big)\Big] d\tau, \\ &\qquad \int_0^{t/2}\Big[\partial_P^h \gamma_m(P, W, \tau), \Theta_m\Big(U, W, \frac{t}{2} - \tau\Big)\Big] d\tau\Big]_U\Big] \\ &\equiv J_1(P, t) + 2J_2(P, t) + J_3(P, t) .\end{aligned}$$

We proceed to estimate the J_i. We shall make use of the inequality [4]

$$(3.5)\qquad [\alpha(P, P)] \leqq \binom{N}{p} | \alpha(P, P) |^2 ,$$

and of the inequality [1]

$$\begin{aligned}(3.6)\quad &\int_0^t \int_{-\infty}^{\infty} \frac{\exp\{-\lambda | x - z |^2/(t - \tau)\}}{(t - \tau)^{\mu}} \frac{\exp\{-\lambda | z - y |^2/\tau\}}{\tau^{\nu}} dz d\tau \\ &\leqq \text{const.} \frac{\exp\{-\lambda | x - y |^2/t\}}{t^{\mu+\nu-1-N/2}}\end{aligned}$$

where $dz = dz^1 \cdots dz^N$ and $\lambda > 0$, $\mu < N/2 + 1$, $\nu < N/2 + 1$. The following, easily verified, inequality will also be used:

$$\begin{aligned}(3.7)\quad &\int_{-\infty}^{\infty} \exp\{-\lambda | x - z |^2/t\} \exp\{-\lambda | z - y |^2/t\} dz \\ &\leqq \text{const.} \exp\{-\mu | x - y |^2/t\} t^{N/2}\end{aligned}$$

where $dz = dz^1 \cdots dz^N$ and $\lambda > \mu > 0$. We shall denote by A_j constants which (unless otherwise stated) may depend only on h and on the manifold M.

Using (3.6) one can prove by induction on i that

$$(3.8)\qquad | \partial_P^h \delta_m^i(P, U, t) | \leqq \frac{A_1^{i+1}}{i!} t^{i(m+1-|h|/2)-1-N/2} e^{-r^2/5t} .$$

The case $i = 1$ follows by (2.11), (2.12). (In deriving (3.8) we also use the elementality inequality $\lambda e^{-\alpha\lambda} \leqq \text{const. } e^{-\delta\lambda}$ for all $\lambda > 0$, where α, δ are constants and $\alpha > \delta \geqq 0$.) In (3.8) it is understood that t° (if it occurs) must be replaced by -$\log t$. From now on we take m such that

$$m + 1 - \frac{|h|}{2} > 0 .$$

Using the definition (2.14) we then conclude from (3.8) that

$$(3.9) \qquad |\partial_P^h \gamma_m(P, Q, t)| \leqq A_2 e^{-r^2/5t} t^{m-(|h|+N)/2} .$$

Next, from the definition of Θ_m one derives

$$(3.10) \qquad |\partial_P^h \Theta_m(P, Q, t)| \leqq A_3 e^{-r^2/5t} t^{-(|h|+N)/2} .$$

Combining (3.9) and (3.10) ($h = 0$) and applying (3.6), we get

$$(3.11) \qquad \left|\int_0^{t/2}\left[\partial_P^h \gamma_m(P, W, \tau), \Theta_m\left(U, W, \frac{t}{2} - \tau\right)\right]d\tau\right| \leqq A_4 e^{-2r^2/5t} t^{m+1-(|h|+N)/2} .$$

Using (3.10), (3.11) one easily derives, applying (3.7),

$$(3.12) \qquad J_2(P, t) \leqq A_5 t^{m+1-|h|-N/2} .$$

Similary one gets

$$(3.13) \qquad J_3(P, t) \leqq A_6 t^{2(m+1)-|h|-N/2} .$$

Evaluation of $J_1(P, t)$. From the construction of Θ_m it follows that for every sufficiently small neighborhood V we may take it to be of the form

$$(3.14) \qquad \Theta_m(P, U, t) = H_m(P, U, t) + R_m(P, U, t) \quad \text{for all} \quad P \in V$$

where H_m is constructed in § 2 and where, for some $\alpha' > 0$,

$$(3.15) \qquad |\partial_P^h R_m(P, U, t)| \leqq A_7 e^{-\alpha'/t} t^{|h|+N/2} \leqq A_8 t^\zeta$$

for any $\zeta > 0$. A_8 depends also on ζ. Next,

$$(3.16) \qquad \partial_P^h H_m(P, U, t) = \sum_{j=0}^{m} t^j \sum_{|\nu|=0}^{|h|} \binom{h}{\nu} \partial_P^\nu f \partial_P^{h-\nu}(MU_j)$$

where $\binom{h}{\nu} = \binom{h_1}{\nu_1} \cdots \binom{h_N}{\nu_N}$. It is easily seen that

$$(3.17) \qquad \partial_P^\nu f(r^2, t) = \sum_{|\mu|=0}^{\nu_0} H_{\nu\mu}\left(\frac{y-x}{\sqrt{t}}\right) f(r^2, t) t^{|\nu|/2+|\mu|/2}$$

where y^i, x^i are the coordinates of U, P respectively, and $H_{\nu\mu}(z)$ is a polynomial in $z = (z^1, \cdots, z^N)$ with C^∞ coefficients which, for $H_{\nu 0}$, are functions of x only. Substituting (3.17) into (3.16) and recalling that $M(P, U)\nu_2$ becomes (δ_I^J) at $P = U$, we obtain

$$\partial_P^h H_m(P, U, t) = H_{h0}\Big(\frac{y-x}{\sqrt{t}}\Big) f(r^2, t) t^{-|h|/2} Y + S_h(P, U, t) \tag{3.18}$$

where Y is the matrix (δ_I^J) and

$$|S_h(P, U, t)| \leqq A_9 e^{-r^2/2t} t^{(1-|h|-N)/2}\,. \tag{3.19}$$

Combining (3.14), (3.15), (3.18), (3.19) we conclude that

$$\partial_P^h \Theta_m(P, U, t) = H_{h0}\Big(\frac{y-x}{\sqrt{t}}\Big) f(r^2, t) t^{-|h|/2} Y + T_h(P, U, t) \tag{3.20}$$

and

$$|T_h(P, U, t)| \leqq A_{10} t^{(1-|h|-N)/2}\,.$$

Using the definition of J_1, and substituting (3.20) in the part of the integral $[\partial_P^h \Theta_m(P, U, t/2), \partial_P^h \Theta_m(P, U, t/2)]_U$ taken over a coordinate patch V_0 containing $\bar{V}$: $y^i - x^i = \xi^i \sqrt{t}$, we find that

$$J_1(P, t) = (C_h(P) + B_0(P, t)) t^{-|h|-N/2} \tag{3.21}$$

where $C_h(P)$ is a continuous function of P, and $|B_0(P, t)| \leqq A_{11}\sqrt{t}$ for $P \in V, 0 < t \leqq b$, for any $b > 0$. A_{11} depends on b.

Combining the evaluation of J_1 with (3.12), (3.13), we obtain from (3.4),

$$\sum_{i=1}^{\infty} |\partial^h \omega_i(P)|^2 e^{-\lambda_i t} = C_h(P) t^{-|h|-N/2} + D_h(P, t) t^{-|h|-(N-1)/2} \tag{3.22}$$

where $D_h(P, t)$ is a uniformly continuous function of (P, t), $P \in V$ and $0 < t \leqq b$ for any $b > 0$. Thus

$$|D_h(P, t)| \leqq A_{12} \tag{3.23}$$

where A_{12} depends on b.

Note that the A_i, in particular A_{12}, are independent of P which varies in V.

4. Asymptotic formulas. To derive asymptotic formulas from the equation (3.22) we use a Tauberian theorem due to Karamata, specialized to Dirichlet series [14; p. 192]. It states:

Let $a_k \geqq 0$ and $0 \leqq \lambda_1 \leqq \lambda_2 \leqq \cdots \leqq \lambda_n \leqq \cdots$, and assume that the Dirichlet series $f(t) = \sum_{k=1}^{\infty} a_k e^{-\lambda_k t}$ converges for $t > 0$ and satisfies

$$f(t) \sim \frac{A}{t^\gamma} \text{ as } t \searrow 0 \qquad (\gamma \geqq 0) .$$

Then the function $\alpha(x) = \sum_{\lambda_k \leqq x} a_k$ *satisfies*

$$\alpha(x) \sim \frac{Ax^\gamma}{\Gamma(\gamma+1)} \quad \text{as} \quad x \to \infty .$$

Applying it to (3.22) (using (3.23)), we get

$$\sum_{\lambda_i \leqq \lambda} |\partial^h \omega_i(P)|^2 = \frac{C_h(P)}{\Gamma(|h|+1+N/2)} \lambda^{|h|+N/2}[1+o(1)] \; (\lambda \to \infty) \tag{4.1}$$

and $o(1) \to 0$ as $\lambda \to \infty$, uniformly in $P \in V$.

Let $\lambda_1 = \cdots = \lambda_{q-1} = 0, \lambda_q > 0$. Using the asymptotic formula (4.1) we shall prove:

THEOREM 1. *For any h and for any $\varepsilon > 0$, the series*

$$\sum_{i=q}^{\infty} \frac{|\partial^h \omega_i(P)|^2}{\lambda_i^{N/2+|h|+\varepsilon}} \tag{4.2}$$

is uniformly convergent in $P \in M$.

Proof. We introduce the function

$$B(P, \lambda) \equiv \sum_{\lambda_q \leqq \lambda_i \leqq \lambda} |\partial^h \omega_i(P)|^2 .$$

Then, we can write the series (4.2) in the form

$$\int_{\lambda'}^{\infty} \frac{dB(P,\lambda)}{\lambda^{N/2+|h|+\varepsilon}} \quad \text{for any} \quad 0 < \lambda' < \lambda_q .$$

Integrating by parts we get

$$\lim_{\mu\to\infty} \left[\frac{B(P,\lambda)}{\lambda^{N/2+|h|+\varepsilon}}\right]_{\lambda=\lambda'}^{\lambda=\mu} - \left(\frac{N}{2} + |h| + \varepsilon\right) \int_{\lambda'}^{\infty} \frac{B(P,\lambda)}{\lambda^{N/2+|h|+\varepsilon+1}} d\lambda . \tag{4.3}$$

Since, by (4.1), $B(P, \lambda) \leqq A_{13}\lambda^{|h|+N/2}$ and since $B(P, \lambda') = 0$, the first term in (4.3) vanishes. The integral in (4.3) converges uniformly in P in view of the bound on $B(P,\lambda)$ just given. The proof of Theorem 1 is thereby completed.

5. Solution of the system (1.1), (1.2). We first derive the formal solution. Substituting

$$g(P) = \sum_{n=1}^{\infty} g_n \omega_n(P), \; h(P) = \sum_{n=1}^{\infty} h_n \omega_n(P), \; f(P,t) = \sum_{n=1}^{\infty} f_n(t) \omega_n(P) \tag{5.1}$$

$$(5.2)\qquad v(P, t) = \sum_{n=1}^{\infty} v_n(t)\omega_n(P)$$

into (1.1), (1.2) we arrive at the equations

$$(5.3)\qquad v_n''(t) + \lambda_n v_n(t) = f_n(t)$$

$$(5.4)\qquad v_n(0) = g_n,\ v_n'(0) = h_n\ .$$

If $\lambda_n = 0$ the solution is

$$v_n(t) = g_n + h_n t + \int_0^t f(\tau)(t-\tau)d\tau\ .$$

If $\lambda_n > 0$ the solution is

$$v_n(t) = g_n \cos\sqrt{\lambda_n}\, t + \frac{h_n}{\sqrt{\lambda_n}} \sin\sqrt{\lambda_n}\, t + \frac{1}{\sqrt{\lambda_n}}\int_0^t f_n(\tau)\sin\sqrt{\lambda_n}\,(t-\tau)d\tau\ .$$

Hence, the formal solution of (1.1), (1.2) is

$$(5.5)\qquad \begin{aligned} v(P, t) &= \sum_{n=1}^{\infty} g_n\omega_n(P)\cos\sqrt{\lambda_n}\, t + \sum_{n=1}^{q-1} h_n\omega_n(P)t \\ &+ \sum_{n=q}^{\infty}\frac{h_n}{\sqrt{\lambda_n}}\omega_n(P)\sin\sqrt{\lambda_n}\, t + \sum_{n=1}^{q-1}\omega_n(P)\int_0^t f_n(\tau)(t-\tau)d\tau \\ &+ \sum_{n=q}^{\infty}\frac{1}{\sqrt{\lambda_n}}\omega_n(P)\int_0^t f_n(\tau)\sin\sqrt{\lambda_n}\,(t-\tau)d\tau\ . \end{aligned}$$

To prove that the formal solution is a genuine one we observe that if $\lambda_n > 0$

$$(5.6)\qquad g_n = \int_M g(Q)*\omega_n(Q) = \frac{1}{\lambda_n^m}\int_M L^m g(Q)*\omega_n(Q)$$

for any positive integer m. Applying Bessel's inequality, we get

$$(5.7)\qquad \sum_{n=1}^{\infty}\lambda_n^{2m}g_n^2 \leqq \int_M L^m g(Q)*L^m g(Q) = \|\, L^m g\,\|^2\ .$$

Similarly,

$$(5.8)\qquad \sum_{n=1}^{\infty}\lambda_n^{2m}h_n^2 \leqq \|\, L^m h\,\|^2,\ \sum_{n=1}^{\infty}\lambda_n^{2m}(f_n(t))^2 \leqq \|\, L^m f(\cdot, t)\,\|^2\ .$$

It will be enough to show that the part of the first series on the right side of (5.5), where summation is on $\lambda_n > 0$, when differentiated term-by-term twice with respect to P is uniformly convergent in $P \in M$, $0 \leqq t \leqq b$, for any $b > 0$. Now the series obtained is majorized by

$$\Sigma\,|\, g_n\,|\,|\,\partial^2 w_n(P)\,| \leqq \Sigma\lambda_n^k\,|\, g_n\,|\frac{|\,\partial^2\omega_n(P)\,|}{\lambda_n^k} \leqq \Sigma\lambda_n^{2k}g_n^2\Sigma\frac{|\,\partial^2\omega_n(P)\,|^2}{\lambda_n^{2k}}\ .$$

Hence that series is uniformly convergent if $k > N/2 + 1$.

It is clear that each series in (5.5) can actually be differentiated term-by-term any number of times and the resulting series is uniformly convergent.

By a solution of (1.1), (1.2) we mean a p-form which is (a) twice continuously differentiable in (P, t) for $P \in M, t > 0$ (b) once continuously differentiable in t for $P \in M, t \geqq 0$ and (c) satisfies (1.1), (1.2).

The uniqueness of the solution can be proved as for the classical wave equation. Assuming $g \equiv 0$, $h \equiv 0$, $f \equiv 0$ and using the rule $\int du^*\omega = \int u^*\delta\omega$ one finds that if u is a solution then

$$\frac{\partial}{\partial t}\int_M [u_t * u_t + \delta u * \delta u + du * du - cu * u] = 0 .$$

Since the integral vanishes for $t = 0$, it vanishes for all $t > 0$. Since the integrand is nonnegative, $u_t * u_t \equiv 0$, which implies $u_t \equiv 0$ and hence, $u \equiv 0$.

We have thus completed the proof of the following theorem.

THEOREM 2. *Let g, h be C^∞ p-forms and let f be a C^∞ p-form such that $\partial_P^\lambda f$ is continuous in (P, t), for any λ. Then the Cauchy problem* (1.1), (1.2) *has one and only one solution. The solution is a C^∞ p-form and is given by* (5.5).

The assumption that the manifold M is C^∞ can be weakened. Indeed, the theory of differential forms used above remains valid under the assumption that the metric tensor is C^5 (Gaffney [3]; see also Friedrichs [2]). The assumptions on f, g, h can also be weakened without any modification of the preceding proof of Theorem 2.

We need the assumptions:

(A) The metric tensor g_{ij} belongs to $C^{[N/2]+2}$ and to C^5, and c belongs to $C^{[N/2]+1}$ (recall that $c \leqq 0$).

(B) The form g belongs to $C^{[N/2]+3}$ and $L^{[(N+4)/4]}g$ belongs to C^1.

(C) The form h belongs to $C^{[N/2]+2}$ and $L^{[(N+2)/2]}h$ belongs to C^1.

(D) The form f and its first $[N/2] + 2$ p-derivatives are continuous for $P \in M$, $0 \leqq t \leqq b$ (for any $b > 0$); $L^{[(N+2)/2]}f$ and its first p-derivatives are continuous for $P \in M$, $0 \leqq t \leqq b$.

THEOREM 2′. *Under the assumptions* (A) — (D), *there exists one and only one solution of the Cauchy problem* (1.1), (1.2). *It is given by* (5.5).

The assertion of Theorem 2′ remains valid if we further weaken the assumptions (A) — (D) by replacing the classes of continuous derivatives C^q by classes of "strong" derivatives W_2^q (see [6]), assuming that $g_{ij} \in C^5$.

6. The heat equation. The method of § 5 can easily be extended to solve the system (1.4), (1.5). The formal solution is

$$u(P, t) = \sum_{n=1}^{\infty} g_n \omega_n(P) e^{-\lambda_n t} + \sum_{n=1}^{\infty} \omega_n(P) \int_0^t f_n(\tau) e^{-\lambda_n (t-\tau)} d\tau \,. \tag{6.1}$$

We shall need the assumptions:

(A′) g_{ij} belong to $C^{[N/2]+1}$ and to C^5, and e belongs to $C^{[N/2]}$.

(B′) The form g belongs to $C^{[N/2]+1}$ and $L^{[N/4]}g$ belongs to C^1.

THEOREM 3. *Under the assumption* (A′), (B′), (D) *there exists a unique solution of the system* (1.4), (1.5). *It is given by* (6.1).

REMARK 1. The assumption $c \leqq 0$ is not needed for the validity of Theorem 3 since it can be achieved by a transformation $u = e^{\alpha t} u$ for any constant $\alpha \geqq c$.

REMARK 2. Assuming $c \leqq 0, f \equiv 0$, we can rewrite (6.1) as an operator equation

$$T_t = H + \sum_{k=1}^{\infty} e^{-\mu_k t} H_k \tag{6.2}$$

where $\{\mu_k\}$ is the sequence $\{\lambda_j\}$ taken without multiplicities, H_k is the projection into the space of eigenforms corresponding to μ_k, H corresponds to $\mu_0 = 0$, and T_t is the operator which maps g into the solution u, that is, $u(P, t) = T_t g(P)$. Formula (6.2) was derived, in a different way (for $c \equiv 0$) by Milgram and Rosenbloom [10].

REFERENCES

1. F. G. Dressel, *The fundamental solution of the parabolic equation II*, Duke Math. J., **13** (1946), 61-70.
2. K. O. Friedichs, *Differential forms on Riemannian manifolds*, Comm. Pure Appl. Math., **8** (1955), 551-590.
3. M. P. Gaffney, *Hilbert space methods in the theory of harmonic integrals*, Trans. Amer. Math. Soc., **78** (1955), 426-444.
4. M. P. Gaffney, *Asymptotic distributions associated with the Laplacian forms*, Comm. Pure Appl. Math., **11** (1958), 535-545.
5. L. Garding, *Cauchy's Problem for hyperbolic equations*, University of Chicago Notes, 1957.
6. V. A. Il'in, *On solving mixed problems for hyperbolic and parabolic equations*, Uspehi Math. Nauk, **15**, no. 2 (1960), 97-154.
7. K. Kodaira, *Harmonic fields in Riemannian manifolds, generalized potential theory*, Ann. of Math., **50** (1949), 587-665.
8. O. A. Ladyzhenskaya, *The mixed problem for hyperbolic equations*, Gosudarvstv. Izdat.

Tehn.-Teor. Lit., Moscow, 1953.
9. J. Leray, *Lectures on hyperbolic equations with variable coefficients*, Princeton, Inst. for Advanced Study, 1952.
10. A. N. Milgram and P. C. Rosenbloom, *Harmonic forms and heat conduction I: Closed Riemannian manifolds*, Proc. Nat. Acad. Sci., **37** (1951), 180-184.
11. S. Minakshisundaram, *Eigenfunctions on Riemannian manifolds*, J. Indian Math. Soc., **17** (1953), 159-165.
12. I. G. Petrowski, *Über des Cauchysche Problem für Systeme von partiellen Differentialgleichungen*, Rec. Math. (Moscow) N.S., **2** (1937), 814-868.
13. G. De Rham, *Variétés Différentiables*, Hermann, Paris, 1955.
14. D. Widder, *The Laplace Transform*, Princeton Univ. Press, Princeton, 1946.

UNIVERSITY OF MINNESOTA
INSTITUTE OF TECHNOLOGY
MINNEAPOLIS, MINNESOTA

BASES OF TENSOR PRODUCTS OF BANACH SPACES

B. R. GELBAUM AND J. GIL DE LAMADRID

1. Introduction. In this note we use the conventions and notations of Schatten [4] with the exception that we use B' to indicate the dual (conjugate) space of a Banach space B and $\langle x, x' \rangle$ as the action of an element x and a functional x' on each other. Schatten defines the tensor product $B_1 \otimes {}_\alpha B_2$ as the completion of the algebraic tensor product $B_1 \otimes B_2$ of two Banach spaces B_1 and B_2, on which the cross norm α has been imposed. We discuss the proposition, "If B_1 and B_2 have Schauder bases, then $B_1 \otimes {}_\alpha B_2$ has a Schauder basis." We prove this for $\alpha = \gamma$ ($B_1 \otimes {}_\gamma B_2$ is the trace class of transformations of B_1' into B_2). We also prove it for $\alpha = \lambda$, ($B_1 \otimes {}_\lambda B_2$ is the class of all completely continuous linear transformations of B_1' into B_2) in the case in which the bases of B_1 and B_2 satisfy an "isometry condition". This condition is not very restrictive. We know of no instance in which it is not satisfied. Next we show that unconditional bases of B_1 and B_2 do not necessarily yield an unconditional basis for the tensor product, even in the nicest conceivable infinite dimensional case, that in which $B_1 = B_2 =$ Hilbert space, and the bases are orthonormal and identical.

We recall certain facts about Schauder bases, and set some general notation that we use throughout the paper. We usually work with a biorthogonal set $\Omega = \{x_i, x_i'\}_i$ associated with a Banach space B, so that $\chi = \{x_i\}_i$ is a basis for B with coefficients supplied by the corresponding sequence of functionals $\chi' = \{x_i'\}_i$. We will have to do with the closed linear manifold B^Ω of B' generated by the elements of χ'. Since B and B^Ω are in duality it is possible to embed B in $(B^\Omega)'$ by the same formula that effects the embedding of B in B''. We denote by ${}_nP_m$ the projection of B defined by ${}_nP_m x = \sum_{i=n}^{m} \langle x, x_i' \rangle x_i$. The double sequence $\{{}_nP_m\}_{n,m}$ is uniformly bounded. We denote by T' the transpose of any transformation T. The following lemma, given without proof, is but a trivial strengthening of [2, p. 18, Theorem 1].

LEMMA 1. *Let E be a dense vector subspace of B, Ω a biorthogonal set of B such that $\chi \subset E$, the vector space spanned by χ is dense in E and the sequence $\{{}_nP_m\}_{n,m}$ is uniformly bounded on E. Then Ω defines a basis for B.*

2. The tensor product of two biorthogonal sets. Let $\Omega_1 = \{x_i, x_i'\}_i$ be a biorthogonal set of B_1 and $\Omega_2 = \{y_i, y_i'\}_i$ a biorthogonal set of B_2.

Received November 15, 1960. This research was supported in part by NSF research grant No. NSFG 14137 and in part by NSF research grant No. NSFG 11048.

The elements $x'_i \otimes y'_j$ can be considered as belonging to $(B_1 \otimes_\alpha B_2)'$ for any cross norm α [4 p. 43], and $\{x_i \otimes y_j, x'_i \otimes y'_j\}_{i,j}$ is clearly a biorthogonal set. We enumerate it, not by the diagonal method, i.e., as in the usual proof that the rationals are denumerable, but as follows: In the table

$$\begin{array}{llll} x_1 \otimes y_1 & x_1 \otimes y_2 & x_1 \otimes y_3 & \cdots\cdots \\ x_2 \otimes y_1 & x_2 \otimes y_2 & x_2 \otimes y_3 & \cdots\cdots \\ x_3 \otimes y_1 & x_3 \otimes y_2 & x_3 \otimes y_3 & \cdots\cdots \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \end{array}$$

we simply order the elements by listing the entries on the two inner sides of each successive upper left hand block to obtain $x_1 \otimes y_1, x_1 \otimes y_2, x_2 \otimes y_2, x_2 \otimes y_1, x_1 \otimes y_3, x_2 \otimes y_3, x_3 \otimes y_3, x_3 \otimes y_2, x_3 \otimes y_1, \cdots, x_1 \otimes y_k, x_2 \otimes y_k \cdots x_k \otimes y_k, x_k \otimes y_{k-1}, \cdots, x_k \otimes y_2, x_k \otimes y_1, \cdots$. This double sequence with the given order is called the tensor product of $\chi_1 = \{x_i\}_i$ and $\chi_2 = \{y_j\}_j$ and is denoted by $\chi_1 \otimes \chi_2$. Similarly $\chi'_1 \otimes \chi'_2$ denotes the set $\{x'_i \otimes y'_j\}_{i,j}$ with the corresponding order. The biorthogonal set formed by $\chi_1 \otimes \chi_2$ and $\chi'_1 \otimes \chi'_2$ is called the tensor product of Ω_1 and Ω_2 and denoted by $\Omega_1 \otimes \Omega_2$.

THEOREM 1. *If Ω_1 defines a basis for B_1 and Ω_2 defines a basis for B_2, then $\Omega_1 \otimes \Omega_2$ defines a basis for $B_1 \otimes_\gamma B_2$.*

Proof. We show that the vector space spanned by $\chi_1 \otimes \chi_2$ is dense in $B_1 \otimes B_2$. To see this let ${}_nP^i_m$ be the ${}_nP_m$ defined in § 1 for Ω_i, and define

$$\text{(1)}\quad A_m = x \otimes y - \sum_{k,j=1}^{m} \langle x, x'_k\rangle\langle y, y'_j\rangle x_k \otimes y_j = x \otimes y - [{}_1P^1_m x] \otimes [{}_1P^2_m y]$$
$$= x \otimes [y - {}_1P^2_m y] + [x - {}_1P^1_m x] \otimes {}_1P^2_m y\ .$$

Then

$$\text{(2)}\qquad \gamma(A_m) \leqq ||x||\,||y - {}_1P^2_m y|| + ||x - {}_1P^1_m x||\,||{}_1P^2_m y||\ .$$

The right hand side of (2) tends to zero with m^{-1}. This argument extends by linearity to sums of elements of the form $x \otimes y$.

Let now T_q be the ${}_1P_q$ defined in § 1 corresponding to $\Omega_1 \otimes \Omega_2$. It remains to show that $\{T_q\}_q$ is uniformly bounded. It is easy to show that each T_q has one of the following three forms: ${}_1P^1_n \otimes {}_1P^2_n$, ${}_1P^1_n \otimes {}_1P^2_n + {}_{n+1}P^1_{n+1} \otimes {}_1P^2_m$, ${}_1P^1_m \otimes {}_{n+1}P^2_{n+1}$. Hence, it suffices to show that $\{{}_nP^1_m \otimes {}_qP^2_r\}_{q,r\,n,m}$ is uniformly bounded. Let M be a common bound for all ${}_nP^1_m$ and ${}_qP^2_r$. For $\Sigma x \otimes y \in B_1 \otimes B_2$

$$\text{(3)}\qquad \gamma[{}_nP^1_m \otimes {}_qP^2_r(\Sigma x \otimes y)] = \gamma[\Sigma({}_nP^1_m x) \otimes ({}_qP^2_r y)]$$
$$\leqq (\textstyle\sum ||x||\,||y||)M^2\ .$$

Since (3) holds for any representation $\Sigma x \otimes y$ of a given tensor product element, we may replace in it the sum $\Sigma \|x\| \|y\|$ by $\gamma(\Sigma x \otimes y)$, thereby proving our assertion. From Lemma 1, we can conclude that $\Omega_1 \otimes \Omega_2$ defines a basis for $B_1 \otimes_\gamma B_2$.

3. The space of completely continuous transformations. We recall that there is a canonical imbedding of B, with a biorthogonal set Ω defining a basis of B, into $(B^\Omega)'$. The norm of the image of an element $x \in B$ is less than or equal to $\|x\|$. We say that Ω satisfies the condition of isometry if the imbedding is actually an isometery. For such an Ω, $(B^\Omega)^\Omega = B$, isometrically. We state first the following corollary of Theorem 1.

COROLLARY 1. *If Ω_k is a biorthogonal set defining a basis for $B_k, k = 1, 2$, then $\Omega_1 \otimes \Omega_2$ defines a basis for $B_1^{\Omega_1} \otimes_\lambda B_2^{\Omega_2}$.*

Proof. Each $x'_i \otimes y'_j$ is an element of $B_1^{\Omega_1} \otimes B_2^{\Omega_2}$ which, as a subset of $B'_1 \otimes_\lambda B'_2$, can be imbedded isometrically in $(B_1 \otimes_\gamma B_2)'$ [4, *p.* 47, Theorem 3.2]. What is more, the vector space spanned by $\{x'_i \otimes y'_j\}_{i,j}$ is dense, with respect to λ, in $B_1^{\Omega_1} \otimes B_2^{\Omega_2}$, hence in $B_1^{\Omega_1} \otimes_\lambda B_2^{\Omega_2}$. This is true because

$$(4) \qquad \lambda[x' \otimes y' - (\sum_{i=1}^n \langle x_i, x' \rangle x'_i) \otimes (\sum_{i=1}^n \langle y_i, y' \rangle y'_i)] \leqq \gamma\,[x' \otimes y' - (\sum_{i=1}^n \langle x_i, x' \rangle x'_i) \otimes (\sum_{i=1}^n \langle y_i, y' \rangle y'_i)]$$

and the latter quantity tends to 0. Hence $B_1^{\Omega_1} \otimes_\lambda B_2^{\Omega_2} = (B_1 \otimes_\gamma B_2)^{\Omega_1 \otimes \Omega_2}$. Our result is a consequence of this.

The next theorem follows easily from this corollary.

THEOREM 2. *If both Ω_1 and Ω_2 satisfy the condition of isometry $\Omega_1 \otimes \Omega_2$ defines a basis for $B_1 \otimes_\lambda B_2$.*

Proof. If in Corollary 1 we replace B_1 by $B_1^{\Omega_1}$ and B_2 by $B_2^{\Omega_2}$, we conclude that $\Omega_1 \otimes \Omega_2$ defines a basis for $(B_1^{\Omega_1})^{\Omega_1} \otimes_\lambda (B_2^{\Omega_2})^{\Omega_2}$. When the condition of isometry is satisfied the last tensor product can be identified with $B_1 \otimes_\lambda B_2$, owing to the relations $B_k = (B_k^{\Omega_k})^{\Omega_k}$ for $k = 1, 2$, and the universal character of λ, [4, p. 35, Lemma 2.12].

Theorem 2 can be considered as a sharpening of the well known fact that if B_1 and B_2 have bases, then every completely continuous linear transformation of B'_1 into B_2 can be uniformly approximated by finite dimensional linear transformations. Our theorem goes further to state that if Ω_1 and Ω_2 satisfy the condition of isometry, the space of all completely continuous linear transformations of B'_1 into B_2 has a basis consisting of one-dimensional linear transformations.

The condition of isometry deserves some explanation. It is satisfied by a large class of bases, which includes every base for which

$$B^{\Omega} = B' .[1] \tag{5}$$

The equation (5) holds always for reflexive spaces. It also holds for certain bases of non-reflexive spaces.

A non-reflexive example of (5) is exhibited in [2, p. 188, Example 1], involving the usual basis of c_0, $x_i = \{\delta_j^i\}_j$, with $x_i' = \{\delta_j^i\}_j \in l^1$. An example of the condition of isometry, in the absence (5), is obtained from this first example, by setting [2, p. 188, Example 2] $y_1 = x_1$, and $y_i = x_i - x_{i-1} + \cdots + (-1)^{i-1}x_1$, for $i > 1$, and $y_i' = x_i' + x_{i+1}'$. For $\Omega = \{y_i, y_i'\}_i$, $x_1' \in B' \backslash B^{\Omega}$. Ω satisfies the condition of isometry for, if $x \in c_0$, then

$$\|\sum_{k=1}^{n} \langle x, y_k'\rangle y_k\| = \|\sum_{k=1}^{n} \langle x, x_k'\rangle x_k\| \leqq \|x\|.$$

The conclusion is now a consequence of the following theorem and its corollary.

THEOREM 3. *If for every $x' \in B'$, $\|{}_1P_n'x'\| \to \|x'\|$, then Ω satisfies the condition of isometry.*

Proof. Let $x_0 \in B$ and $x_0' \in B'$ such that $\|x_0'\| = 1$ and $\langle x_0, x_0'\rangle = \|x_0\|$. Then

$$\lim_{n\to\infty} \frac{\langle x_0, {}_1P_n'x_0'\rangle}{\|{}_1P_n'x_0'\|} = \|x_0\| , \qquad \text{Q.E.D.}$$

COROLLARY 2. *If $\|{}_1P_n\| \leqq 1$ for every n, then Ω satisfies the condition of isometry.*

Proof. We show the above hypothesis implies the hypothesis of Theorem 3. To see this, let $x_0' \in B'$, and $\varepsilon > 0$. There is $x_0 \in B$ so that $\|x_0\| = 1$ and $\langle x_0, x_0'\rangle > \|x_0\| - \varepsilon/2$ and an integer $N > 0$ so that

$$\begin{aligned}\|x_0'\| \geqq \|{}_1P_n'x_0'\| \geqq \langle x_0, {}_1P_n'x_0'\rangle = \langle {}_1P_nx_0, x_0'\rangle > \langle x_0, x_0'\rangle - \varepsilon/2 \\ > \|x_0'\| - \varepsilon , \qquad \text{Q.E.D.}\end{aligned}$$

As we have seen, the two biorthogonal sets described above for c_0 satisfy the hypothesis of Corollary 1.

An example of the isometry condition in which B' is not separable is furnished by Schauder's basis for $C([0, 1])$, given by the biorthogonal system $\Omega = \{x_i, x_i'\}_i$ described in [1, p. 69]. We consider [0, 1] imbedded

[1] This equation may be described by saying that $\{x_i'\}_i$ is a *retrobasis for B'*, [2, p. 188, Definition 1].

in B' and treat its points as functionals. The space B^{Ω} of this example contains the set D of all dyadic fractions. Consequently Ω satisfies the condition of isometry, since, for $f \in B$, $\|f\| = \sup_{d \in D}|f(d)|$.

We know of no biorthogonal set defining a basis which does not satisfy the condition of isometry. Neither do we know if $B_1 \otimes_\alpha B_2$ has a basis for an arbitrary cross norm α, even if B_1 and B_2 have bases. It is clear that for any element of $B_1 \otimes B_2$, the formal expansion of Theorem 1 converges to that element with respect to α, since it does with respect to $\gamma \geqq \alpha$. The difficulty lies in establishing that the set $\{{}_pP^1_q \otimes {}_rP^2_s\}_{\substack{p,q\\r,s}}$ is uniformly bounded with respect to α.

4. Hilbert spaces and unconditional bases. The problem of approximation of compact operators by finite dimensional operators in a Banach space, can, after elaborate rearrangement, lead to the following question: Can there exist a matrix $C = (c_{ij})_{i,j=1}^{\infty}$ satisfying the following conditions:

(a) For some $a_i \geqq 0$, $\sum_{i=1}^{\infty} a_i^2 < \infty$, $|c_{ij}| \leqq a_i a_j$;

(b) $C^2 = 0$;

(c) $\sum_{i=1}^{\infty} c_{ii} = 1$?

Of course, (b) and (c) are incompatible if C is in the trace class. Thus there arises the question: Does (a) imply that C is in the trace class? To this we can give a definite negative answer via the following theorems.

Therem 4. *Let $\Omega = \{x_i, x_i'\}_i$, $x_i = \{\delta^i_j\}_j$, $x_i' = \{\delta^i_j\}_j$ be the canonical orthonormal basis in l_2. Then $\Omega \otimes \Omega$ defines an unconditional basis in $l_2 \otimes_\gamma l_2$ if and only of condition* (a) *implies C is in the trace class.*

Proof. Let $\Omega \otimes \Omega$ define an unconditional basis for $l_2 \otimes_\gamma l_2$. Then we note that (a) may be rephrased by stating: $c_{ij} = \varepsilon_{ij} a_i a_j$, $|\varepsilon_{ij}| \leqq 1$. Since $l_2 \otimes_\gamma l_2$ is precisely the trace class of operators [4] it follows that $\sum_{i,j=1}^{\infty} \varepsilon_{ij}\, a_i a_j (x_i \otimes x_j)$ exists in $l_2 \otimes_\gamma l_2$ and is therefore in the trace class.

On the other hand, if (a) implies that C is in the trace class, then for $a \otimes a$ in $l_2 \otimes_\gamma l_2$ $(a = (a_1, a_2, \cdots))$, $a \otimes a = \sum_{i,j=1}^{\infty} a_i a_j (x_i \otimes x_j)$. If $B = (\varepsilon_{ij} a_i a_j)$ is in the trace class, then B has an expansion $\sum_{i,j=1}^{\infty} \varepsilon_{ij} a_i a_j (x_i \otimes x_j)$, which shows $\Omega \otimes \Omega$ defines an unconditional basis for $l_2 \otimes_\gamma l_2$.

THEOREM 5. *$\Omega \otimes \Omega$ does not define an unconditional basis for $l_2 \otimes_\gamma l_2$.*

Proof. Let $A_1 = (a_{ij})$ be a 2×2 matrix with $a_{11} = a_{12} = a_{22} = -a_{21} = 1$, and A_n the $2^n \times 2^n$ matrix (A_{ij}) $i, j = 1, 2$, with $A_{11} = A_{12} = A_{22} = -A_{21} = A_{n-1}$. Let B be the direct sum of the matrices $\{1/2^{n/2} A_n\}_n$. Then a direct computation reveals that B is unitary. Let $B = (b_{ij})$, and let

$C = (|b_{ij}|)$. If $\Omega \otimes \Omega$ were an unconditional basis for $l_2 \otimes_\gamma l_2$, then for B, regarded as a member of $(l_2 \otimes_\gamma l_2)'$ [4, p. 47, Theorem 3.2] and arbitrary $u \otimes v$ in $l_2 \otimes_\gamma l_2$, $\sum_{i,j=1}^{\infty} u_i v_j \langle x_i, Bx_j \rangle$ would converge unconditionally, i.e. $\sum_{i,j=1}^{\infty} u_k v_j |b_{ji}|$ would converge. In particular, let $u = v$, where u is given by the vector: $\sum_{n=1}^{\infty} (1/n) x_n$, $(\sqrt{2^n}) x_n = (\underbrace{0, 0, \cdots 0}_{2(2^{n-1}-1)}, \overbrace{1, 1, \cdots, 1}^{2^n}, 0, 0, \cdots)$. A simple verification shows that u exists in l_2. On the other hand, more calculation shows $\sum_{i,j=1}^{\infty} |b_{ij}| u_i u_j = \infty$. The contradiction implies the theorem.

Theorem 5 remains valid when γ is replaced by λ, since $l_2 \otimes_\gamma l_2 = (l_2 \otimes_\lambda l_2)'$, and unconditionality of $\Omega \otimes \Omega$ in $l_2 \otimes_\lambda l_2$ implies the same in $l_2 \otimes_\gamma l_2$.

NOTE. We owe to the referee the remark that a space B with a biorthogonal set Ω which defines a basis for B can always be renormed, preserving the topology of B [1, Theorem 1, p. 67], in such a way that Ω satisfies the condition of isometry (section 3) with respect to the resulting norm of B and the corresponding norm of B'. This makes possible the following completely general form of Theorem 2.

THEOREM 2'. *If Ω_i defines a basis for B_i, for $i = 1, 2$, then $\Omega_1 \otimes \Omega_2$ defines a basis for $B_1 \otimes_\lambda B_2$.*

Proof. Renorm B_1 and B_2 as indicated above. Then, if λ' denotes the operator norm with respect to the new norms of B_1 and B_2, $B_1 \otimes_{\lambda'} B_2$ has a basis defined by $\Omega_1 \otimes \Omega_2$ (Theorem 2). But $B_1 \otimes_{\lambda'} B_2 = B_1 \otimes_\lambda B_2$ both point-set-wise and topologically. Hence our conclusion.

BIBLIOGRAPHY

1. M. M. Day, *Normed linear spaces*, Ergebnisse der Math. und ihre Grenzgebiete, N. F. 28, 1958.
2. B. R. Gelbaum, *Expansions in Banach Spaces*, Duke Math. J., **17** (1950), 187-196.
3. A. Grothendieck, *Produits tensoriels topologiques et espaces nucléaires*, Mem. Amer. Math. Soc. No. 16, (1955).
4. R. Schatten, *A theory of cross spaces*, Ann. Math. Studies. No. 26, Princeton Univ. Press, Princeton 1950.

INSTITUTE FOR ADVANCED STUDY,
UNIVERSITY OF MINNESOTA

INFINITELY DIVISIBLE PROBABILITIES ON THE HYPERBOLIC PLANE

R. K. GETOOR

1. Introduction. This paper may be regarded from two points of view. First of all it presents the theory of infinitely divisible radially symmetric probability measures on the hyperbolic plane and the naturally associated limit theorems. This point of view provided the motivation for the present paper and is explained in some detail in § 2 and 3. However, just as the analagous theory in the Euclidean case may be viewed as a chapter in the theory of Fourier transforms, so may the present theory be viewed as a chapter in the theory of Legendre transforms. That is, by using the harmonic analysis described in §§ 2 and 3 one can set up a one-to-one correspondence between the radially symmetric probability measures, μ, on the hyperbolic plane and certain functions of a complex variable $\phi(z)$ in such a way that the convolution of μ_1 and μ_2 on the hyperbolic plane corresponds to the pointwise product of their "transforms" ϕ_1 and ϕ_2. Since μ is radially symmetric it is completely specified by a distribution function $F(\lambda)$ on $\lambda \geqq 0$ and the correspondence between ϕ and μ (or F) is given by

$$\phi(z) = \int_0^\infty K(z, \lambda) dF(\lambda) \tag{1.1}$$

where $K(z, \lambda)$ is a certain Legendre function given by (4.9). The convolution of μ_1 and μ_2, at least in the case where F_1 and F_2 have densities, is written down explicitly in (3.9).

This second point of view is adopted for the most part beginning in § 4 and so the majority of the paper (sections 4–10) deals with certain problems in the theory of the Legendre transform (1.1). The tools we use are those of classical analysis, but the problems treated are motivated by probability theory. The main results of the paper are contained in §§ 7 and 8. In § 10 Gaussian and stable distributions are defined within the present context. Finally in § 11 we indicate the extensions of these ideas to a wider class of Legendre transforms which includes the theory of radially symmetric probability measures on the higher dimensional hyperbolic spaces as special cases.

We would like to thank Professor H. P. McKean who first introduced us to the material in § 2, and who expressed interest and encouragement when the present paper was in its formative stages.

Received January 10, 1961. This research was supported, in part, by the National Science Foundation.

2. General Remarks. The present section is devoted to a general situation that we will specialize to the hyperbolic plane in the next section. We follow, more or less, the expository article of Godement [4]. Let G be a locally compact second countable topological group and let K be a compact subgroup. Let x, y, z denote elements of G and u, v elements of K. We define two equivalence relations on G as follows:

$$x \sim y \Longleftrightarrow xy^{-1} \in K \tag{2.1}$$

$$x \approx y \Longleftrightarrow \text{ there exist } \; u, v \in K \; \text{ such that } \; x = uyv \; . \tag{2.2}$$

Thus $H = G/\sim$ is the space of right cosets and $R = G/\approx$ is the space of double cosets. We give H and R the usual quotient space topology. Let dx be *right invariant* Haar measure on G, then dx induces measures dh and dr on H and R which are invariant under the (right) action of G. In order to avoid notational complications it will be convenient to regard all functions as being defined on G. Thus the statement $f \in L_1(R)$ will mean $f(x) = f(y)$ if $x \approx y$ and $\int |f(x)| dx < \infty$ with the obvious conventions for functions defined almost everywhere. Thus we have

$$L_p(R) \subset L_p(H) \subset L_p(G) \tag{2.3}$$

for each $p > 0$.

If f and g are in $L_1(G)$ we define their convolution

$$f*g.(x) = \int f(xy^{-1})g(y)dy \; . \tag{2.4}$$

It is well known that $L_1(G)$ is a Banach algebra and it is immediate that $L_1(R)$ and $L_1(H)$ are closed sub-algebras. The basic assumption of [4] is that $L_1(R)$ be *commutative*. Selberg [6] has shown that if H is a symmetric (or more generally, weakly symmetric) space then $L_1(R)$ is commutative. For us the following simple sufficient condition (whose proof is a routine calculation and is therefore omitted) will suffice.

THEOREM 2.1[1]. *If $x \approx x^{-1}$ for all x in G, then $L_1(R)$ is commutative.*

In the remainder of this section we will assume that $L_1(R)$ is commutative. If α is a multiplicative linear functional on $L_1(R)$ then

$$\alpha(f) = \int p_\alpha(x) f(x) dx \tag{2.5}$$

where p_α is in $L_\infty(R)$. It is easy to see that α defined by (2.5) is a multiplicative linear functional if and only if

$$p_\alpha(x)p_\alpha(y) = \int_K p_\alpha(xuy)du \tag{2.6}$$

[1] This remark is due to H. P. McKean.

for almost all x, y. Here du is *normalized* Haar measure on K. Moreover (2.6) implies that p_α is equal almost everywhere to a continuous function and thus the multiplicative linear functionals on $L_1(R)$ may be identified with the bounded continuous functions on R satisfying (2.6). Such functions are called *spherical* functions on R.

Let C_∞ be the continuous functions on G with compact support. If $f \in C_\infty$ we define $\tilde{f}(x) = \bar{f}(x^{-1})$. A (signed) Radon measure μ on G is said to be of *positive type* if

$$\mu(\tilde{f} * f) \geqq 0 \tag{2.7}$$

for all $f \in C_\infty$. A continuous function p on G is of positive type if the measure $p(x)dx$ is. Let $\hat{R}$ be the totality of all spherical functions on R which are of positive type. For $f \in L_1(R)$ we define

$$\hat{f}(p) = \int f(x)p(x)dx \quad \text{for} \quad p \in \hat{R}\,, \tag{2.8}$$

then, at least if G is unimodular, one can develop a complete theory of harmonic analysis including a Plancherel theorem. For details see [4]. Since we won't need this general theory we will terminate our general discussion at this point.

3. **The Hyperbolic Plane.** Let D be the interior of the unit disc in the complex plane, i.e., $D = \{z : z \text{ complex}, |z| < 1\}$. The set D furnished with the Riemannian metric

$$ds^2 = 4(1 - r^2)^{-2}[dx^2 + dy^2] \tag{3.1}$$

where $r^2 = |z|^2 = x^2 + y^2$ will be called the hyperbolic plane. The geodesic joining z_1 and z_2 is the unique circle through them cutting the circle $|z| = 1$ orthogonally. The hyperbolic distance $\zeta(z_1, z_2)$ between z_1 and z_2 is given by

$$th\zeta/2 = |z_1 - z_2| \cdot |1 - \bar{z}_1 z_2|^{-1} \tag{3.2}$$

where "th" denotes the hyperbolic tangent, similarly "ch" and "sh" will denote the hyperbolic cosine and sine. See [2] for a discussion of the hyperbolic plane including the above facts.

If u and b are complex numbers with $|u| = 1$ and $|b| < 1$ we define the hyperbolic motion (u, b) as follows:

$$(u, b) : z \to u(z - b)(1 - \bar{b}z)^{-1}\,. \tag{3.3}$$

It is easy to check that (3.3) maps D onto D and preserves the hyperbolic distance (3.2). Let G be the totality of all such motions with the obvious topology. Clearly G is a topological group satisfying the hypotheses of § 2. The multiplication in G is composition, i.e.,

$$(3.4) \qquad (u_1, b_1)(u_2, b_2) = \left(\frac{u_1(u_2 + b_1\bar{b}_2)}{u_2(\bar{u}_2 + \bar{b}_1 b_2)}, \frac{u_2 b_2 + b_1}{u_2 + b_1\bar{b}_2}\right)$$

and

$$(3.5) \qquad (u, b)^{-1} = (\bar{u}, -ub) .$$

Let K be the compact subgroup consisting of motions $(u, 0)$, then K is just the circle group. It is straight forward to check that the equivalence relations defined in (2.1) and (2.2) become

$$(3.6) \qquad (u_1, b_1) \sim (u_2, b_2) \Longleftrightarrow b_1 = b_2$$

$$(3.7) \qquad (u_1, b_1) \approx (u_2, b_2) \Longleftrightarrow |b_1| = |b_2| .$$

Thus H can be identified with D and R with the half-open segment $0 \leqq r < 1$. Moreover Haar measure in G can be chosen so that the invariant measure induced in $H = D$ is just that associated with the Riemannian metric (3.1).

It is convenient to introduce geodesic polar coordinates (ζ, θ) in D with pole at $z = 0$. Here ζ is the hyperbolic distance and the coordinates of the point $z = re^{i\theta}$ with $0 \leqq \theta < 2\pi$ are $(\log 1 + r/1 - r, \theta)$. In terms of these coordinates we have $ds^2 = d\zeta^2 + sh^2\zeta d\theta^2$ and the corresponding volume element is given by

$$(3.8) \qquad sh\zeta \, d\zeta \, d\theta .$$

We now regard R as the half-line $0 \leqq \zeta < \infty$ and write $f(ch\zeta)$ for the generic function defined on R. Theorem 2.1, (3.5), and (3.7) imply that $L_1(R)$ is commutative and routine calculations show that if $f, g \in L_1(R)$, then their convolution is given by

$$(3.9) \qquad f*g.(ch\zeta) = \int_0^\infty \int_0^{2\pi} f(ch\lambda)g[ch\zeta ch\lambda - sh\zeta sh\lambda \cos\theta]sh\lambda d\theta d\lambda$$
$$= \int_0^\infty \int_0^{2\pi} f[ch\zeta ch\lambda - sh\zeta sh\lambda \cos\theta]g(ch\lambda)sh\lambda d\theta d\lambda .$$

Moreover the defining equation (2.6) for spherical functions becomes

$$(3.10) \qquad p(ch\lambda)p(ch\zeta) = \frac{1}{2\pi}\int_0^{2\pi} p[ch\lambda ch\zeta - sh\lambda sh\zeta \cos\theta]d\theta .$$

From (3.10) one can show that p is a solution of the Legendre differential equation and since $p(1) = 1$, it follows that the solutions of (3.10) are $P_\nu(ch\zeta)$ where P_ν denotes the Legendre function of the first kind. See [3]. Equation (3.10) is then a simple consequence of the usual addition theorem for Legendre functions [3].

Since the spherical functions are bounded we must have $-1 \leqq Re(\nu) \leqq 0$. Finally it is not too difficult to see that $P_\nu(ch\zeta)$ is of positive type if and only if $Re(\nu) = -1/2$. Thus we have identified the spherical functions and the spherical functions of positive type for the hyperbolic plane. See also [4].

The fact that the two integrals in (3.9) are equal is, of course, a consequence of the general theory of § 2. However, one can see this directly as follows. In the first integral in (3.9) we regard (λ, θ) as geodesic polar coordinates for the hyperbolic plane D with pole at 0. Let z_0 be the point whose coordinates are $(\zeta, 0)$ and let (μ, φ) be geodesic polar coordinates for D with pole at z_0 with the same polar axis. Using (3.2) it is not difficult to see that relationship between the coordinates (λ, θ) and (μ, φ) of a point z is

$$\begin{aligned} ch\mu &= ch\lambda ch\zeta - sh\lambda sh\zeta \cos\theta \\ ch\lambda &= ch\mu ch\zeta + sh\mu sh\zeta \cos\varphi \,. \end{aligned} \tag{3.11}$$

Thus regarding the first integral in (3.9) as an integral over the hyperbolic plane we see that it is equal to (the volume elements are $sh\lambda d\lambda d\theta$ and $sh\mu d\mu d\varphi$, see (3.8))

$$\int_0^\infty \int_0^{2\pi} f[ch\mu ch\zeta + sh\mu sh\zeta \cos\varphi] g(ch\mu) sh\mu d\varphi d\mu$$

and this is obviously equal to the second integral in (3.9). The relation expressed in (3.11) is just the law of cosines for hyperbolic trigonometry.

4. Preliminaries on Legendre Functions. We intend to study integral transforms of the form

$$\varphi(\nu) = \int_0^\infty P_\nu(ch\lambda) dF(\lambda) \tag{4.1}$$

where F is a bounded monotone nondecreasing function and $P_\nu(ch\lambda)$ is the usual Legendre function. In this section we gather together some facts about the kernels $P_\nu(ch\lambda)$ that we will need in the sequel.

Combining formula (3) of § 3.2 and formula (22) of § 2.1 of [3] we see that

$$P_\nu(ch\lambda) = \left(\frac{1 + ch\lambda}{2}\right)^\nu F\left(-\nu, -\nu; 1; \frac{ch\lambda - 1}{ch\lambda + 1}\right) \tag{4.2}$$

for all complex ν provided $0 \leqq \lambda < \infty$, where F is the usual hypergeometric function. In all statements to follow λ is a nonnegative real number. It is immediate from (4.2) that for each fixed λ the function $P_\nu(ch\lambda)$ is an entire function of the complex variable ν. Also

$$\begin{aligned} P_\nu(1) &= 1 \\ P_0(ch\lambda) = P_{-1}(ch\lambda) &= 1 \,. \end{aligned} \tag{4.3}$$

From formula (9) of § 3.7 of [3] we have

$$P_\nu(ch\lambda) = -\frac{\sqrt{2}}{\pi}\sin\nu\pi\int_0^\infty \frac{ch\left(\nu+\frac{1}{2}\right)t\,dt}{(ch\lambda+cht)^{1/2}} \tag{4.4}$$

provided $-1 < Re(\nu) < 0$ and in particular for real x we have

$$P_{-\frac{1}{2}+ix}(ch\lambda) = \frac{\sqrt{2}}{\pi}ch\pi x\int_0^\infty \frac{\cos xt\,dt}{(ch\lambda+cht)^{1/2}}\,. \tag{4.5}$$

Using (4.2) above and the standard integral representation for the hypergeometric function (§ 2.1.3 of [3]) we find that

$$P_\nu(ch\lambda) = \left(\frac{1+ch\lambda}{2}\right)^\nu [\Gamma(-\nu)\Gamma(1+\nu)]^{-1} \int_0^1 t^{-\nu-1}(1-t)^\nu\left(1-\frac{ch\lambda-1}{ch\lambda+1}t\right)^\nu dt \tag{4.6}$$

provided $-1 < Re\nu < 0$. For ν real with $-1 < \nu < 0$ it is immediate from (4.6) that

$$0 \leqq P_\nu(ch\lambda) \leqq 1 \tag{4.7}$$

and since $P_\nu(ch\lambda)$ is a continuous function of ν the inequality (4.7) must hold for all ν in the interval $-1 \leqq \nu \leqq 0$. On the other hand using formula (14) of § 3.7 of [3] we have

$$|P_\nu(ch\lambda)| \leqq \frac{1}{\pi}\int_0^\pi [ch\lambda + sh\lambda\cos t]^{Re\nu}dt = P_{Re\nu}(ch\lambda)\,,$$

and combining this with (4.7) we obtain

$$|P_\nu(ch\lambda)| \leqq 1 \tag{4.8}$$

provided $-1 \leqq Re\nu \leqq 0$.

Let $z = x + iy$ be a complex variable and define the function

$$K(z, \lambda) = P_{-\frac{1}{2}+iz}(ch\lambda) \tag{4.9}$$

for $\lambda \geqq 0$ and $-1/2 \leqq y \leqq 1/2$. For each fixed λ, $K(z, \lambda)$ is an analytic function of z in the strip $-1/2 < y < 1/2$ and is continuous in the closed strip $-1/2 \leqq y \leqq 1/2$. The properties (4.3) and (4.8) become

$$K(z, 0) = 1,\ K\left(-\frac{i}{2}, \lambda\right) = K\left(\frac{i}{2}, \lambda\right) = 1 \qquad |K(z, \lambda)| \leqq 1\,. \tag{4.10}$$

Moreover $K(x, \lambda)$ is given by the right hand side of (4.5).

5. Uniqueness and Continuity. Let $\mathscr{F}$ be the collection of all bounded monotone nondecreasing real valued functions defined on $0 \leqq \lambda < \infty$ and normalized so that

$$\text{(5.1)}\quad \text{(i)} \qquad F(0) = 0$$
$$\text{(ii)} \qquad F(\lambda) \to F(\mu) \quad \text{as} \quad \lambda \downarrow \mu \quad \text{for all} \quad \mu > 0 \, .$$

Note that $F(0+)$ need not be zero. Let $\mathscr{F}_0$ be those $F \in \mathscr{F}$ which satisfy

$$\text{(5.2)} \qquad F(\infty) = \lim_{\lambda \to \infty} F(\lambda) = 1 \, .$$

All integrals are to be in the Lebesgue-Stieltjes sense. Integrals over $0 \leqq \lambda < \infty$ will be written $\int_0^\infty$, while integrals over $0 < \lambda < \infty$ will be written $\int_{0+}^\infty$.

It $\{F_n\}$ is a sequence in $\mathscr{F}$ and $F \in \mathscr{F}$, then we say that F_n *converges weakly* to F (written $F_n \to F$) provided

$$\text{(5.3)} \qquad \int_0^\infty f dF_n \to \int_0^\infty f dF$$

for all continuous f with compact support. We say that F_n is *Bernoulli convergent* to F provided (5.3) holds for all bounded continuous f. It is obvious that if each $F_n \in \mathscr{F}_0$ and $F_n \Rightarrow F$, then $F \in \mathscr{F}_0$. ($F_n \Rightarrow F$ means F_n is Bernoulli convergent to F.)

If $K(z, \lambda)$ is the kernel defined in (4.9) we define the transform of $F \in \mathscr{F}$ by

$$\text{(5.4)} \qquad \varphi(z) = \int_0^\infty K(z, \lambda) dF(\lambda) \, .$$

It is immediate that φ is bounded in absolute value and continuous in the strip $-1/2 \leqq y \leqq 1/2$, and is analytic in $-1/2 < y < 1/2$. In particular $\varphi(x)$ is a real valued even function of x, and $F \in \mathscr{F}_0$ if and only if $\varphi(-i/2) = 1$. Of course, the values of φ on the real axis completely determine φ in the strip $-1/2 \leqq y \leqq 1/2$. We now show that F is uniquely determined by φ.

THEOREM 5.1 *If φ is the transform of F, then $\varphi(x)$ uniquely determines F.*

Proof. It suffices to prove that if F is of bounded variation (not necessarily monotone) and if the integral in (5.4) vanishes for all x then F is identically zero. Using the representation (4.5) for $K(x, \lambda)$ we have

$$\text{(5.5)} \qquad \int_0^\infty h(t) \cos xt dt = 0 \text{ for all } x$$

where

$$h(t) = H(cht) = \int_0^\infty [ch\lambda + cht]^{-\frac{1}{2}} dF(\lambda) . \tag{5.6}$$

The interchange of order of integration is justified since

$$|h(t)| \leqq Var(F)[1 + cht]^{-\frac{1}{2}} .$$

Moreover h is continuous and the above inequality implies $h \in L_1$, hence (5.5) coupled with the uniqueness theorem for Fourier integrals yields $h(t) = 0$ for all $t \geqq 0$. Thus if we define $G(u) = F(ch^{-1}(u))$ for $u \geqq 1$ and $G(u) = 0$ for $0 \leqq u < 1$ we have

$$H(t) = \int_0^\infty (u + t)^{-\frac{1}{2}} dG(u) = 0$$

for all $t \geqq 1$. But

$$\begin{aligned} H(t) &= \pi^{-\frac{1}{2}} \int_0^\infty \int_0^\infty e^{-\sigma(u+t)} \sigma^{-\frac{1}{2}} d\sigma dG(u) \\ &= \pi^{-\frac{1}{2}} \int_0^\infty e^{-\sigma t} \sigma^{-\frac{1}{2}} g(\sigma) d\sigma \end{aligned}$$

where $g(\sigma)$ is the Laplace-Stieltjes transform of G. Since $H(t) = 0$ for $t \geqq 1$ we see that $\sigma^{-\frac{1}{2}} g(\sigma) = 0$ for almost all σ which in turn implies that G, and hence F, is zero. Here we have used the uniqueness theorem [7] for Laplace transforms twice.

In the present work the following rather weak continuity theorem will suffice.

THEOREM 5.2. (i) *If $F_n \to F$ and $F_n(\infty) \leqq M$ then $\varphi_n(x) \to \varphi(x)$.*

(ii) *Let φ_n be the transform of F_n and suppose $\varphi_n(-i/2) = F_n(\infty) \leqq M$, then if $\varphi_n(x) \to \varphi(x)$ there exists an $F \in \mathscr{F}$ such that $F_n \to F$ and φ is the trasform of F.*

Proof. (i) For each x the function $K(x, \cdot)$ is continuous and (4.5) implies that it vanishes at infinity. Thus it is immediate that for each fixed x we have $\varphi_n(x) \to \varphi(x)$.

(ii) Since $F_n(\infty) \leqq M$ the Helly theorem implies the existence of a subsequence $\{F'_n\}$ weakly convergent to F'. If φ' is the transform of F' then (i) implies that $\varphi'(x) = \varphi(x)$ for all x. If the entire sequence $\{F_n\}$ does not converge to F', then there exists another subsequence $\{F''_n\}$ converging to $F'' \neq F'$. But as before $\varphi''(x) = \varphi(x) = \varphi'(x)$ which contradicts the uniqueness theorem. Thus if we let $F = F'$ the proof of (ii) is complete.

REMARK 1. Since the limit φ in (ii) above is a transform it follows that $\varphi(z)$ may be defined as a continuous function on the strip $-1/2 \leqq y \leqq 1/2$ which is analytic in the open strip. If $\varphi(-i/2) = \lim_{n\to\infty} \varphi_n(-i/2)$, then $F_n(\infty) \to F(\infty)$ which implies that $F_n \Rightarrow F$.

REMARK 2. If $F_n \Rightarrow F$ then since $K(z, \lambda)$ is continuous and bounded for $\lambda \geqq 0$ and z in the strip $-1/2 \leqq y \leqq 1/2$, it follows that $\varphi_n(z) \to \varphi(z)$ uniformly on each compact sub strip $-1/2 \leqq y \leqq 1/2$ and $0 \leqq x \leqq x_0$. See, for example, Lemma 1.5.2 (iv) of [1].

6. **Closure Properties.** Let $\mathscr{P}(\mathscr{P}_0)$ be the class of all transforms of functions in $\mathscr{F}(\mathscr{F}_0)$. It is then immediate that if φ and ψ are in $\mathscr{P}$ and $\alpha, \beta \geqq 0$ then $\alpha\varphi + \beta\psi \in \mathscr{P}$, while if $\varphi, \psi \in \mathscr{P}_0$ and $\alpha + \beta = 1$, $\alpha \geqq 0, \beta \geqq 0$, then $\alpha\varphi + \beta\psi \in \mathscr{P}_0$. Moreover Theorem 5.2 (ii) implies that if $\{\varphi_n\}$ is a sequence in $\mathscr{P}_0$ and $\varphi_n(x) \to \varphi(x)$, then $\varphi \in \mathscr{P}$.

The main result of this section is that $\mathscr{P}$ and $\mathscr{P}_0$ are closed under pointwise products. The proof of the following theorem is, of course, motivated by the general discussion of §§ 2 and 3.

THEOREM 6.1. *If φ_1 and φ_2 are in $\mathscr{P}(\mathscr{P}_0)$ then $\varphi_1\varphi_2$ is in $\mathscr{P}(\mathscr{P}_0)$.*

Proof. Let φ_1 and φ_2 be the transforms of F_1 and F_2. We first consider the case in which F_1 and F_2 have continuous densities $f_1(ch\lambda)$ and $f_2(ch\lambda)$ with respect to the measure $sh\lambda d\lambda$. Of course, $f_j(ch\lambda)$ may be unbounded near $\lambda = 0$. Thus

$$\varphi_j(x) = \int_0^\infty K(x, \lambda) f_j(ch\lambda) sh\lambda d\lambda \quad j = 1, 2 \tag{6.1}$$

$$\int_0^\infty f_j(ch\lambda) sh\lambda d\lambda < \infty\ ;\ f_j(ch\lambda) \geqq 0\ ;\ j = 1, 2.$$

For the purposes of the present proof it will be convenient to write

$$p(z, ch\lambda) = K(z, \lambda) = P_{-\frac{1}{2}+iz}(ch\lambda)\ . \tag{6.2}$$

An immediate consequence of the addition theorem for Legendre functions [3] is that

$$p(z, ch\lambda) p(z, ch\mu) = \frac{1}{2\pi}\int_0^{2\pi} p(z, a(\theta)) d\theta \tag{6.3}$$

where

$$a(\theta) = ch\lambda ch\mu - sh\lambda sh\mu \cos\theta\ . \tag{6.4}$$

Therefore

$$\varphi_1(z)\varphi_2(z) = \frac{1}{2\pi}\int_0^\infty f_1(ch\lambda)sh\lambda d\lambda\int_0^\infty\int_0^{2\pi} f_2(ch\mu)p(z, a(\theta))sh\mu d\theta d\mu .$$

If in the inner integrals we make the change of variable described in the last paragraph of § 3 we find (the manipulations are justified since $f_2(ch\mu) \in L_1(sh\mu d\mu)$ and p is bounded)

$$\varphi_1(z)\varphi_2(z) = \int_0^\infty p(z, ch\mu)h(ch\mu)sh\mu d\mu$$

where

$$h(ch\mu) = \frac{1}{2\pi}\int_0^\infty\int_0^{2\pi} f_1(ch\lambda)f_2(a(\theta))sh\lambda d\theta d\lambda .$$

clearly $h \geqq 0$ and

$$F_1(\infty)F_2(\infty) = \varphi_1\Big(-\frac{i}{2}\Big)\varphi_2\Big(-\frac{i}{2}\Big) = \int_0^\infty h(ch\mu)sh\mu d\mu .$$

Thus $\varphi_1\varphi_2$ is in $\mathscr{P}$, and if φ_1, φ_2 are in $\mathscr{P}_0$, then $\varphi_1\varphi_2$ is in $\mathscr{P}_0$.

We now turn to the general case. Let $k(\lambda) = e^{-\lambda}$ if $\lambda \geqq 0$ and $k(\lambda) = 0$ if $\lambda < 0$, and put $k_n(\lambda) = nk(n\lambda)$. Defining $F_1(\lambda)$ and $F_2(\lambda)$ to be zero for $\lambda < 0$, it is clear that

$${}_nf_i(\lambda) = \int_{-\infty}^\infty k_n(\lambda - \mu)dF_i(\mu) = \int_0^\lambda k_n(\lambda - \mu)dF_i(\mu)$$

are continuous functions of $\lambda \geqq 0$. Moreover if we define ${}_nF_i(\lambda) = \int_0^\lambda {}_nf_i(\mu)d\mu$, it follows that ${}_nF_i \in \mathscr{F}$, and ${}_nF_i \in \mathscr{F}_0$ if $F_i \in \mathscr{F}_0$. Here $i = 1, 2$. It is well known [1, Th. 1.5.3] that ${}_nF_i \Rightarrow F_i$ as $n \to \infty$. If ${}_ng_i(ch\lambda) = {}_nf_i(\lambda)$ then ${}_ng_i(ch\lambda)(sh\lambda)^{-1}$ is the density of ${}_nF_i$ with respect to $sh\lambda d\lambda$. Thus if ${}_n\varphi_i$ is the transform of ${}_nF_i$, it follows from what we proved above that $\psi_n(z) = {}_n\varphi_1(z){}_n\varphi_2(z)$ is in $\mathscr{P}$ (or $\mathscr{P}_0$). But by the second remark following theorem 5.2 we have ${}_n\varphi_i(z) \to \varphi_i(z)$ everywhere on the strip $-1/2 \leqq y \leqq 1/2$. Thus $\varphi_1(z)\varphi_2(z) = \lim_{n\to\infty}\psi_n(z)$ is in $\mathscr{P}$, and $\varphi_1\varphi_2 \in \mathscr{P}_0$ if both φ_1 and φ_2 are since $\psi_n(-i/2) \to \varphi_1(-i/2)\varphi_2(-i/2)$. This completes the proof of Theorem 6.1.

The following theorem gives another interesting closure property of $\mathscr{P}$.

THEOREM 6.2. *If $\varphi \in \mathscr{P}$, then $\psi(z) = \exp(t[\varphi(z) - \varphi(-i/2)])$ is in $\mathscr{P}_0$ for all $t > 0$.*

Proof. If $\alpha = \varphi(-i/2) = F(\infty) \geqq 0$, then

$$\psi(z) = e^{-\alpha t}\lim_{n\to\infty}\sum_{k=0}^n \frac{t^k\varphi(z)^k}{k!} .$$

Using Theorem 6.1, the fact that $\mathscr{P}$ is closed under positive linear combinations, and Theorem 5.2 it follows that $\psi \in \mathscr{P}$. (Theorem 5.2 is applicable since

$$\sum_{k=0}^{n} \frac{t^k \varphi\left(-\frac{i}{2}\right)^k}{k!} \leqq e^{\alpha t}$$

for all n.) Moreover $\psi(-i/2) = 1$ and so $\psi \in \mathscr{P}_0$.

7. Limit Theorems. In this section we will consider only the class $\mathscr{P}_0$. We begin by making the following definitions suggested by probability theory.

DEFINITION 1. $\varphi \in \mathscr{P}_0$ is *infinitely divisible* if for each positive integer n there exists a $\psi_n \in \mathscr{P}_0$ such that $\varphi = (\psi_n)^n$.

DEFINITION 2. $\varphi \in \mathscr{P}_0$ is a *generalized limit* if there exist $\varphi_{nk} \in \mathscr{P}_0$ for $n = 1, 2, \cdots$, and $k = 1, 2, \cdots, k_n$ with $k_n \to \infty$ as $n \to \infty$ such that

$$\max_k |\varphi_{nk}(x) - 1| \to 0 \tag{7.1}$$

uniformly on each bounded interval $0 \leqq x \leqq x_0 < \infty$, and

$$\varphi_n(x) = \prod_{k=1}^{k_n} \varphi_{nk}(x) \to \varphi(x) \tag{7.2}$$

for all x.

In this section we will show that φ is infinitely divisible if and only if φ is a generalized limit, and at the same time obtain a canonical form for such φ. In the course of our discussion we will need the following two lemmas which we state here for convenience. The proofs of these lemmas will be given in § 9.

LEMMA 7.1. (i) *There exist constants $\lambda_0 > 0$ and $M_R < \infty$ such that $\lambda^{-2}|1 - K(z, \lambda)| \leqq M_R$ provided $0 \leqq \lambda \leqq \lambda_0$ and $|z| \leqq R$ (of course z is in the strip $-1/2 \leqq y \leqq 1/2$).*

(ii) *$\lambda^{-2}[1 - K(x, \lambda)] \to 1/4(1/4 + x^2)$ as $\lambda \to 0$.*

LEMMA 7.2. *Let $H_T(\lambda) = \dfrac{1}{T}\displaystyle\int_0^T [1 - K(x, \lambda)]dx$, then*

(i) *$0 \leqq H_T \leqq 2$ and $H_T(\lambda) > 0$ for $\lambda > 0$;*
(ii) *$H_T(\lambda) \to 1$ as $\lambda \to \infty$;*
(iii) *$\lambda^{-2}H_T(\lambda) \to C(T) > 0$ as $\lambda \to 0$.*

We begin with the simple half of our main result.

THEOREM 7.1. *If φ is infinitely divisible, then φ is a generalized limit.*

Proof. Letting $k_n = n$ and $\varphi_{nk} = \psi_n$ for all k we see that φ_n (defined in (7.2)) is identical with φ for all n. Thus we need only verify (7.1). But (4.10) implies that $|\varphi(x)| \leqq 1$, and, of course, $\varphi(x)$ and $\psi_n(x)$ are real. Therefore

$$\psi_n(x)^2 = [\varphi(x)^2]^{\frac{1}{n}} \rightarrow \psi(x)$$

where $\psi(x) = 0$ or 1 according as $\varphi(x) = 0$ or $\varphi(x) \neq 0$. But $\varphi(0) > 0$ and hence $\psi(0) = 1$. Moreover $\psi_n^2 \in \mathscr{P}_0$ and thus Theorem 5.2 implies that $\psi \in \mathscr{P}$. In particular $\psi(x)$ is continuous and since ψ can only take on the values 0 and 1 it follows that $\psi(x) = 1$ for all x. Hence $\varphi(x)$ never vanishes and since $\psi_n(x) = \exp[1/n \log \varphi(x)]$ in a neighborhood (depending on n) of 0 we must have $\psi_n(x) = \exp[1/n \log \varphi(x)]$ for all $x \geqq 0$. Therefore $\psi_n(x) \rightarrow 1$ as $n \rightarrow \infty$. This completes the proof of Theorem 7.1.

We now turn our attention to the converse of Theorem 7.1. This will not be established in full generality until § 8. In working with Definition 2 we will adopt the convention that F_{nk} is the element in $\mathscr{P}_0$ whose transform is φ_{nk}, similarly φ_n is the transform of F_n, φ of F. We begin with the following result.

THEOREM 7.2. *Condition* (7.1) *of definition* (2) *is equivalent to*

$$\max_k \int_\varepsilon^\infty dF_{nk}(\lambda) \rightarrow 0 \tag{7.3}$$

as $n \rightarrow \infty$ for all $\varepsilon > 0$.

Proof. Suppose (7.3) holds then

$$\begin{aligned}\max_k |\varphi_{nk}(x) - 1| &\leqq \max_k \int_0^\varepsilon [1 - K(x, \lambda)] dF_{nk}(\lambda) \\ &\quad + \max_k \int_\varepsilon^\infty [1 - K(x, \lambda)] dF_{nk}(\lambda) \\ &\leqq \varepsilon^2 M_R + 2 \max_k \int_\varepsilon^\infty dF_{nk}(\lambda)\end{aligned}$$

provided $\varepsilon < \lambda_0$ and $0 \leqq x \leqq R$ where λ_0 and M_R are defined in Lemma 7.1. Thus (7.1) follows.

Suppose (7.1) holds and (7.3) does not hold, then there exists an $\varepsilon > 0$ and a subsequence n_j such that

$$\max_k \int_\varepsilon^\infty dF_{n_j k} \geqq \eta > 0 \,. \tag{7.4}$$

Let k_j be the value of k at which the maximum in (7.4) is attained ($0 \leqq k_j \leqq k_{n_j}$) and let $G_j = F_{n_j k_j}$. Let ψ_j be the transform of G_j, then (7.1) implies that $\psi_j(x) \to 1$ which is the transform of E (unit mass at the origin). Thus $G_j \to E$ weakly, but this is clearly a contradiction since $\int_\varepsilon^\infty dG_j \geqq \eta > 0$.

THEOREM 7.3. *If φ is generalized limit then $\varphi(x) = \exp[-\psi(x)]$ where*

$$\psi(x) = \int_0^\infty [1 - K(x, \lambda)] \frac{1 + \lambda^2}{\lambda^2} d\Omega(\lambda) \tag{7.5}$$

with $\Omega \in \mathscr{F}$.

Proof. Theorem 6.1 implies that each $\varphi_n \in \mathscr{P}_0$. Hence $F_n(\infty) = 1 = F(\infty)$ and since $\varphi_n(x) \to \varphi(x)$ we see that $F_n \to F$. Combining these facts yields $F_n \Rightarrow F$. Thus $\varphi_n(z) \to \varphi(z)$ uniformly on each strip $-1/2 \leqq y \leqq 1/2$ and $0 \leqq x \leqq x_0$. (Remarks following Theorem 5.2). Also $\varphi(x)$ can not vanish near $x = 0$ since $\varphi \in \mathscr{P}_0$. Let x_0 be the first zero of φ, then $\varphi_n(x) \to \varphi(x)$ uniformly on $0 \leqq x \leqq x_0$. Condition (7.1) implies that $\varphi_{nk}(x)$ doesn't vanish for $0 \leqq x \leqq x_0$ and all k provided that n is sufficiently large (how large depending only on x_0). Thus for $0 \leqq x < x_0$ and n sufficiently large we can write

$$-\log \varphi_{nk}(x) = -\log \left\{1 - \int_0^\infty [1 - K(x, \lambda)] dF_{nk}(\lambda)\right\}.$$

Letting

$$a_{nk}(x) = 1 - \varphi_{nk}(x) = \int_0^\infty [1 - K(x, \lambda)] dF_{nk}(\lambda) \geqq 0 ,$$

it follows from (7.1) that $a_{nk}(x) \to 0$ uniformly on $0 \leqq x \leqq x_0$ uniformly in k as $n \to \infty$. Hence

$$-\log \varphi_n(x) = -\sum_{k=1}^{k_n} \log \varphi_{nk}(x) = \sum_{k=1}^{k_n} \sum_{j=1}^{\infty} (j)^{-1}(a_{nk})^j , \tag{7.6}$$

and letting $n \to \infty$ we obtain

$$-\log \varphi(x) = \lim_{n\to\infty} \left\{\sum_{k=1}^{k_n} a_{nk} + \sum_{k=1}^{k_n} \sum_{j=2}^{\infty} (j)^{-1}(a_{nk})^j\right\} \tag{7.7}$$

provided $0 \leqq x < x_0$.

Since all the terms involved are nonnegative we have for $0 \leqq x < x_0$

$$0 \leqq \sum_{k=1}^{k_n} \sum_{j=2}^{\infty} (j)^{-1}(a_{nk})^j \leqq \max (a_{nk}) \sum_{k=1}^{k_n} \sum_{j=1}^{\infty} (j)^{-1}(a_{nk})^j \tag{7.8}$$

$$\to 0(-\log \varphi(x)) = 0$$

as $n \to \infty$. If we define

$$G_n(\lambda) = \sum_{k=1}^{k_n} F_{nk}(\lambda)$$

and use (7.7) and (7.8) we obtain

$$-\log \varphi(x) = \lim_{n\to\infty} \int_0^\infty [1 - K(x, \lambda) dG_n(\lambda) \tag{7.9}$$

provided $0 \leqq x < x_0$.

We now investigate the behavior of the functions G_n as $n \to \infty$. It is an immediate consequence of Lemma 7.2 that for each $T > 0$ there exists a constant $A(T) > 0$ such that

$$\frac{1}{T}\int_0^T [1 - K(x, \lambda)]dx \geqq A(T)\frac{\lambda^2}{1+\lambda^2}$$

for all $\lambda \geqq 0$. Also from (7.6) and the definition of G_n it is clear that

$$-\log \varphi_n(x) \geqq \int_0^\infty [1 - K(x, \lambda)]dG_n(\lambda)$$

for $0 \leqq x < x_0$ and n sufficiently large. Moreover since $\varphi_n(x) \to \varphi(x)$ uniformly on $0 \leqq x \leqq x_0$ and $\varphi(x)$ is continuous and bounded away from zero on $0 \leqq x \leqq 1/2x_0$ it follows that $\log \varphi_n(x) \to \log \varphi(x)$ uniformly on $0 \leqq x \leqq 1/2x_0$. Thus if $0 < T < 1/2x_0$ and n large enough we have

$$\int_0^\infty \frac{\lambda^2}{1+\lambda^2} dG_n(\lambda) \leqq [TA(T)]^{-1}\int_0^T -\log \varphi_n(x)dx$$
$$\to [TA(T)]^{-1}\int_0^T -\log \varphi(x)dx < \infty .$$

Hence there exists a constant $M < \infty$ such that

$$\int_0^\infty \frac{\lambda^2}{1+\lambda^2} dG_n(\lambda) \leqq M . \tag{7.10}$$

Next we will show that given $\varepsilon > 0$ there exists an R (independent of n) such that

$$\int_R^\infty dG_n(\lambda) \leqq \varepsilon . \tag{7.11}$$

To this end we first note that $\varphi_n(iy) \to \varphi(iy)$ uniformly an $-1/2 \leqq y \leqq 1/2$. Also (4.2) and (4.9) yield $0 < K(iy, \lambda) \leqq 1$ for $\lambda \geqq 0$ and $-1/2 \leqq y \leqq 1/2$ with $K(-i/2, \lambda) = K(i/2, \lambda) = 1$. Thus $\varphi_n(iy)$ and $\varphi(iy)$ are strictly positive on $-1/2 \leqq y \leqq 1/2$. Since $a_{nk}(iy) = 1 - \varphi_{nk}(iy) < 1$ an argument similar to the one leading to (7.6) yields

$$-\log \varphi_n(iy) \geqq \int_0^\infty [1 - K(iy, \lambda)] dG_n(\lambda) \tag{7.12}$$

Let $\eta > 0$ be given, then since $\varphi(-i/2) = 1$ we can choose T such that $-1/2 < T < 1/2$ and

$$\frac{-2}{2T+1}\int_{-\frac{1}{2}}^{T} \log \varphi(iy) dy < \eta \; .$$

Moreover (4.4) and (4.9) imply $K(iy, \lambda) \to 0$ as $\lambda \to \infty$ for each fixed y with $-1/2 < y < 1/2$. Thus we can choose R_0 so that

$$\frac{2}{2T+1}\int_{-\frac{1}{2}}^{T} [1 - K(iy, \lambda)] dy > 1 - \eta$$

for all $\lambda \geqq R_0$. Since $\varphi_n(iy) \to \varphi(iy)$ uniformly on $-1/2 \leqq y \leqq T$ and φ is bounded away from zero there we can choose n_0 so that for $n > n_0$

$$\frac{-2}{2T+1}\int_{-\frac{1}{2}}^{T} \log \varphi_n(iy) dy \leqq \frac{-2}{2T+1}\int_{-\frac{1}{2}}^{T} \log \varphi(iy) dy + \eta \leqq 2\eta \; .$$

Thus for $n > n_0$ we have

$$\begin{aligned}\int_{R_0}^{\infty} (1-\eta) dG_n(\lambda) &\leqq \frac{2}{2T+1}\int_{-\frac{1}{2}}^{T}\int_0^\infty [1 - K(iy, \lambda)] dG_n(\lambda) dy \\ &\leqq \frac{-2}{2T+1}\int_{-\frac{1}{2}}^{T} \log \varphi_n(iy) dy \leqq 2\eta \; ,\end{aligned}$$

or $\int_{R_0}^\infty dG_n \leqq \dfrac{2\eta}{1-\eta}$ if $n > n_0$. It is now evident that given $\varepsilon > 0$ one can choose an R so that (7.11) holds (each G_n being monotone nondecreasing and bounded).

Define

$$\Omega_n(\lambda) = \int_0^\lambda \frac{t^2}{1 + t^2} dG_n(t) \; . \tag{7.13}$$

Then each Ω_n is in $\mathscr{F}$. Using (7.10), the Helly theorem implies the existence of a subsequence (call it Ω_n again) such that $\Omega_n \to \Omega$ with $\Omega \in \mathscr{F}$ and $\Omega\ (\infty) \leqq M$. Moreover from (7.11)

$$\int_R^\infty d\Omega_n(\lambda) \leqq \int_R^\infty dG_n(\lambda) \leqq \varepsilon$$

uniformly in n for R sufficiently large and this easily implies that $\Omega_n \Rightarrow \Omega$. Therefore

$$\int_0^\infty [1 - K(x, \lambda)] \frac{1+\lambda^2}{\lambda^2} d\Omega_n(\lambda) \to \int_0^\infty [1 - K(x, \lambda)] \frac{1+\lambda^2}{\lambda^2} d\Omega(\lambda) \tag{7.14}$$

where the integrand is defined by continuity, using Lemma 7.1, to be $1/4(1/4 + x^2)$ at $\lambda = 0$. Combining (7.9), (7.13), and (7.14) we obtain

$$(7.15) \qquad -\log \varphi(x) = \int_0^\infty [1 - K(x, \lambda)] \frac{1 + \lambda^2}{\lambda^2} d\Omega(\lambda)$$

provided $0 \leqq x < x_0$. But x_0 was the smallest zero of $\varphi(x)$ and thus $-\log \varphi(x) \to \infty$ as $x \uparrow x_0$ while the integral in (7.15) remains bounded. Hence $\varphi(x)$ never vanishes and (7.15) must hold for all x. Finally defining $\psi(x) = -\log \varphi(x)$ we obtain Theorem 7.3.

THEOREM 7.4. *A function ψ has the representation* (7.5) *if and only if it can be written in the form*

$$(7.16) \qquad \psi(x) = \frac{c}{4}\left(\frac{1}{4} + x^2\right) + \int_{0+}^\infty [1 - K(x, \lambda)] dG(\lambda)$$

where $c \geqq 0$ and G is monotone nondecreasing, right continuous for $\lambda > 0$, $G(\infty) = 0$, and satisfying.

$$(7.17) \qquad \int_{0+}^1 \lambda^2 dG(\lambda) < \infty \, .$$

Proof. If ψ has the representation (7.5) define $c = \Omega(0+) \geqq 0$ and $G(\lambda) = -\int_{\lambda+}^\infty t^{-2}(1 + t^2) d\Omega(t)$ for $\lambda > 0$, then using Lemma 7.1 (ii) it is clear that (7.16) holds and that G has the required properties. Conversely if (7.16) hold define $\Omega(0) = 0$ and $\Omega(\lambda) = c + \int_{0+}^\lambda t^2(1 + t^2)^{-1} dG(t)$, then clearly $\Omega \in \mathscr{F}$ and (7.5) holds.

8. Uniqueness and Simple Consequences of (7.5). In order that our theory be reasonably complete the following uniqueness theorem is required.

THEOREM 8.1. *The representation* (7.5), *and hence* (7.16) *also, is unique.*

Proof. It is sufficient to prove that *if Ω is* of bounded variation and

$$(8.1) \qquad \psi(x) = \int_0^\infty [1 - K(x, \lambda)] \frac{1 + \lambda^2}{\lambda^2} d\Omega(\lambda) = 0$$

for all x, then $\psi = 0$. We will use the following formula [5, p. 168]

$$(ch\lambda + ch\mu)^{-1} = \pi\int_0^\infty \frac{xsh\pi x}{ch^2\pi x} K(x, \lambda)K(x, \mu)dx \tag{8.2}$$

which holds for all $\lambda, \mu \geqq 0$. (Robin's $K_x(ch\lambda)$ is our $K(x, \lambda)$.) Since $K(0, \lambda) = 1$, if we multiply (8.1) by $\pi xsh\pi x(ch\pi x)^{-2}K(x, \mu)$ and integrate from 0 to ∞ we obtain, using (8.2),

$$\int_0^\infty \frac{ch\lambda - 1}{ch\lambda + ch\mu} \frac{1 + \lambda^2}{\lambda^2} d\Omega(\lambda) = 0 . \tag{8.3}$$

The interchange of order of integration is justified since ($|K(x, \mu)| \leqq 1$ and $K(0, \mu) = 1$)

$$\begin{aligned}\pi\int_0^\infty \frac{xsh\pi x}{ch^2\pi x} |K(x, \mu)| \int_0^\infty [1 - K(x, \lambda)] \frac{1 + \lambda^2}{\lambda^2} d|\Omega|(\lambda)dx \\ \leqq \int_0^\infty \frac{ch\lambda - 1}{2(1 + ch\lambda)} \frac{1 + \lambda^2}{\lambda^2} d|\Omega|(\lambda) < \infty ,\end{aligned}$$

where $|\Omega|$ stands for the total variation of Ω. But (8.3) may be written

$$\int_1^\infty (t + s)^{-1}d\Omega^*(t) = 0 \tag{8.4}$$

for all $s \geqq 1$, where for $t \geqq 1$

$$\Omega^*(t) = \int_0^{ch^{-1}(t)} \lambda^{-2}(1 + \lambda^2)(ch\lambda - 1)d\Omega(\lambda) . \tag{8.5}$$

Noting that Ω^* is of bounded variation on each *finite* interval $1 \leqq t \leqq T$, we can apply the uniqueness theorem for Stieltjes transform [7, p. 336]. This leads to the conclusion that Ω^*, and hence Ω, is identically zero.

THEOREM 8.2. *Given a ψ of the form* (7.5) *then $\varphi = e^{-\psi}$ is an infinitely divisible element of $\mathscr{F}_0$.*

Proof. Since $K(\cdot, \lambda)$ is in $\mathscr{P}_0$ for each $\lambda \geqq 0$, it follows from Theorem 6.2 that $\exp\{-b[1 - K(\cdot, \lambda)]\}$ is in $\mathscr{P}_0$ for all $b, \lambda \geqq 0$. Thus if we approximate the integral defining $\psi(x)$ by a Riemann sum and use the above fact and Theorem 5.2, we find that $\varphi(x) = \exp[-\psi(x)]$ is in $\mathscr{P}$. Since φ is in $\mathscr{P}$ it can be extended to a function $\varphi(z)$ which is continuous on the strip $-1/2 \leqq y \leqq 1/2$ and analytic on $-1/2 < y < 1/2$.

Using Lemma 7.1 (i) it is immediate that the integral in (7.5) converges for z in the strip $-1/2 \leqq y \leqq 1/2$ to a function which we denote by $\psi(z)$. It is also clear that ψ is continuous on $-1/2 \leqq y \leqq 1/2$ and analytic on $-1/2 < y < 1/2$. It now follows that $\varphi(z) = \exp[-\psi(z)]$ for all z in the strip $-1/2 \leqq y \leqq 1/2$. Since $\psi(-i/2) = 0$ it follows that $\varphi \in \mathscr{P}_0$. Similarly $\exp[-1/n\psi]$ is in $\mathscr{P}_0$ for each $n > 0$ and thus $\varphi = \exp(-\psi)$ is infinitely divisible.

COROLLARY 8.1. *φ is a generalized limit if and only if φ is infinitely divisible.*

Proof. This results from Theorems 7.1, 7.3, and 8.2.

Actually contained in the proofs of Theorem 7.3, 8.1, and 8.2 is the following result which we state explicitly for completeness.

THEOREM 8.3. *Let $\varphi_{nK} \in \mathscr{P}_0$ for $n = 1, 2, \cdots$ and $k = 1, 2, \cdots, k_n$ with $k_n \rightarrow \infty$ and satisfy* (7.1). *Let φ_n be defined by* (7.2), *then a necessary and sufficient condition $\varphi_n \rightarrow \varphi \in \mathscr{P}_0$ is that*

$$\Omega_n(\lambda) = \sum_{k=1}^{k_n} \int_0^\lambda \frac{t^2}{1+t^2} dF_{nk}(t)$$

be Bernoulli convergent to $\Omega \in \mathscr{F}$. In this case $\varphi = \exp(-\psi)$ where ψ is defined by (7.5).

9. Proofs of the Lemmas. We begin with Lemma 7.1 (i). In view of the definition (4.9) of $K(z, \lambda)$ it will suffice to show $\lambda^{-2} | 1 - P_\nu(ch\lambda) | \leqq M_r$ for $0 \leqq \lambda \leqq \lambda_0$ and $|\nu| \leqq r$. Choose λ_0 such that the inequalities.

$$ch\lambda - 1 < 1 \, ; \, \lambda^{-2}(ch\lambda - 1) < 1 \tag{9.1}$$

hold for $0 \leqq \lambda \leqq \lambda_0$. Since [3, p. 122]

$$P_\nu(ch\lambda) = F\left(-\nu, \nu + 1; 1; \frac{1}{2}(1 - ch\lambda)\right) \tag{9.2}$$

provided $|1 - ch\lambda| < 2$, we easily find that for $0 \leqq \lambda \leqq \lambda_0$

$$\lambda^{-2} | 1 - P_\nu(ch\lambda) | \leqq F\left(|\nu|, |\nu| + 1; 1; \frac{1}{2}\right).$$

Taking $M_r = F(r, r + 1; 1; 1/2)$ we obtain the desired conclusion.

Let λ_0 be as above, then from (9.2) and (4.9) we have for $0 \leqq \lambda \leqq \lambda_0$

$$K(x, \lambda) = F\left[\frac{1}{2} - ix, \frac{1}{2} + ix; 1; \frac{1}{2}(1 - ch\lambda)\right].$$

Expanding the hypergeometric function we find that

$$\lambda^{-2}[1 - K(x, \lambda)] + \frac{1}{2}\lambda^{-2}\left[\left(\frac{1}{4} + x^2\right)(1 - ch\lambda)\right] = G(x, \lambda)$$

where $G(x, \lambda) \rightarrow 0$ as $\lambda \rightarrow 0$ for each fixed x. The second conclusion of Lemma 7.1 is now immediate.

Finally all of the conclusions of Lemma 7.2 with the possible exception of $H_T(\lambda) > 0$ for $\lambda > 0$ are easy consequences of Lemma 7.1 and

(4.5). But if $H_T(\lambda_0) = 0$ for $\lambda_0 > 0$ we would have $K(x, \lambda_0) = 1$ for all x in $[0, T]$. In particular $K(0, \lambda_0) = 1$. Using (4.5) and the fact that $K(0, 0) = 1$ this leads to the conclusion that

$$\int_0^\infty \frac{dt}{[ch\lambda_0 + cht]^{1/2}} = \int_0^\infty \frac{dt}{[1 + cht]^{1/2}}$$

which is clearly impossible if $\lambda_0 > 0$. This then completes the proofs of our lemmas.

10. Gaussian and Stable Distributions. It is an immediate consequence of Theorems 8.2 and 7.4 that for each $t > 0$ the function

$$\varphi(x) = \exp\left[-t\left(\frac{1}{4} + x^2\right)\right] \tag{10.1}$$

is an infinitely divisible element of $\mathscr{F}_0$. Following Bochner [1] a φ of the form (10.1) will be called *Gaussian* (*or normal*). Let $U_t(\lambda)$ be the element in $\mathscr{F}_0$ corresponding to $\exp[-t(1/4 + x^2)]$. If one uses the inversion formula of Fock [5, p. 165] and the fact that [5, p. 154]

$$K(x, \lambda) = \frac{\sqrt{2}}{\pi} \coth \pi x \int_\lambda^\infty \frac{\sin xt\, dt}{(cht - ch\lambda)^{1/2}}$$

one finds that

$$U_t(\lambda) = \int_0^\lambda u(t, \mu) sh\mu d\mu$$

where

$$u(t, \lambda) = \frac{e^{-t/4}}{(2t)^{3/2}} \int_\lambda^\infty \frac{se^{-s^2/4t} ds}{(chs - ch\lambda)^{1/2}} . \tag{10.2}$$

The function $u(t, \lambda)$ defined in (10.2) is therefore the density (with respect to $sh\lambda d\lambda$) of the rotationally invariant Gaussian distribution on the hyperbolic plane. It is not difficult to check directly that

$$\int_0^\infty u(t, \lambda) sh\lambda d\lambda = 1 \quad \text{for all} \quad t > 0 ,$$

although it is not necessary for us to do so since we know that $U_t \in \mathscr{F}_0$. Finally it is interesting, but not unexpected, to note that

$$\frac{\partial u}{\partial t} = (sh\lambda)^{-1} \frac{\partial}{\partial \lambda}\left(sh\lambda \frac{\partial u}{\partial \lambda}\right) . \tag{10.3}$$

The differential operator on the right side of (10.3) is the radial part of the Laplace-Beltrami operator in geodesic polar coordinates (λ, θ) for the hyperbolic plane.

Let $0 < \alpha < 1$ and let $g_\alpha(u)$ be the positive continuous function defined on $u \geqq 0$ by the relation

$$(10.4) \qquad e^{-s^\alpha} = \int_0^\infty e^{-su} g_\alpha(u) du$$

for $s \geqq 0$. See § 4.3 of [1]. (The function g_α is the density of a one-sided stable law of index α on the real line.) For $t > 0$ we define

$$(10.5) \qquad u_\alpha(t, \lambda) = \int_0^\infty u(t^{\frac{1}{\alpha}} s, \lambda) g_\alpha(s) ds$$

where $u(t, \lambda)$ is the normal density (10.2). Clearly $u_\alpha(t, \cdot)$ is a probability density with respect to $sh\lambda d\lambda$. Moreover

$$(10.6) \qquad \int_0^\infty K(x, \lambda) u_\alpha(t, \lambda) sh(\lambda) d\lambda = \exp[-t(1/4 + x^2)^\alpha]$$

Thus for each $t > 0$ the function $\exp[-t(1/4 + x^2)^\alpha]$ is an infinitely divisible element of $\mathscr{P}_0$. By analogy with the Euclidean case one might call the densities (10.5) or the transforms (10.6) *stable*. We will investigate the properties of these distributions in a future paper, in particular, we will give a fuller justification of the name stable.

11. Higher Dimensional Hyperbolic Spaces. All that has gone before can be easily extended to a more general class of integral transforms that are related to the higher dimensional hyperbolic spaces. Let μ be a real parameter with $\mu > -1/2$ and define the kernels

$$(11.1) \qquad \begin{aligned} K_\mu(x, \lambda) &= 2^\mu \Gamma(\mu + 1)(sh\lambda)^{-\mu} P^{(\mu)}_{-1/2+ix}(ch\lambda) \\ &= \left(\frac{1 + ch\lambda}{2}\right)^{-\mu} F\left(1/2 - ix, 1/2 + ix; \mu + 1; \frac{1 - ch\lambda}{2}\right), \end{aligned}$$

where $P_\nu^{(\mu)}$ is the usual associated Legendre function [3]. Similarly we define $K_\mu(z, \lambda)$ for complex z by replacing x by z in (11.1). Clearly $K_\mu(z, \lambda)$ is analytic in z and it is not difficult to verify that

$$(11.2) \qquad | K_\mu(z, \lambda) | \leqq 1 \quad \text{if} \quad -\mu - 1/2 \leqq y \leqq 1/2$$

and

$$(11.3) \qquad K_\mu(z, 0) = 1, \qquad K_\mu(-i(\mu + 1/2), \lambda) = 1\,.$$

Moreover it follows from 3.7(8) of [3] that

$$(11.4) \qquad K_\mu(x, \lambda) = \sqrt{\frac{2}{\pi}} \frac{2^\mu \Gamma(\mu + 1)}{\Gamma(\mu + 1/2)(sh\lambda)^{2\mu}} \int_0^\lambda (ch\lambda - cht)^{\mu - 1/2} \cos xt dt\,,$$

and combining this with (11.2) results in

$$-1 \leqq K_\mu(x, \lambda) \leqq 1 \,. \tag{11.5}$$

If $\mu = 0$ then K_μ reduces to the kernel K considered in the previous sections.

Let $\mu = N/2 - 1$ where $N \geqq 2$ is an integer, then the functions $K_\mu(z, \cdot)$ are the spherical functions on N-dimensional hyperbolic space and the functions $K_\mu(x, \cdot)$ are the spherical functions of positive type.

If for arbitrary $\mu > -1/2$ we define the μ-transform of an element $F \in \mathscr{F}$ by

$$\varphi(x) = \int_0^\infty K_\mu(x, \lambda) dF(\lambda) \tag{11.6}$$

then all of the results of the preceding sections can be carried over to μ-transforms with only minor changes. In particular, in (7.16) one must replace $c/4(1/4 + x^2)$ by $c/[4(\mu + 1)][(\mu + 1/2)^2 + x^2]$ and then the Gaussian elements have the form $\exp(-t[(\mu + 1/2)^2 + x^2])$. The proofs require only minor technical changes with the exception of Theorem 6.1.

In order to prove the analog of Theorem 6.1 for μ-transforms one needs the following formula

$$K_\mu(x, \lambda) K_\mu(x, t) = \frac{2^{2\mu} \Gamma(1 + \mu)^2}{\pi \Gamma(2\mu + 1)} \int_0^\pi K_\mu(x, w) \sin^{2\mu}\theta d\theta \tag{11.7}$$

where

$$w = ch\lambda cht + sh\lambda sht \cos\theta \,.$$

Formula (11.7) is a simple consequence of the addition theorem for associated Legendre functions (formula 80 of Peter Henrici, Addition Theorems for General Legendre and Gegenbauer Functions, Journ. of Rat. Mech. and Anal. (4) 1955; note the misprint in this formula, namely $-\mu - n$ should be $-\nu - n$) and the orthogonality relations for Gegenbauer polynomials. Using (11.7) the fact that the product of two μ-transforms is again a μ-transform is proved by an argument similar to the one used in § 6.

Note added in proof. Some results which are similar to a part of ours appeared in a paper by F. I. Karpelevitch, V. N. Tutubalin, and M. G. Šur entitled "Limit theorems for convolutions of distributions on Lobachevsky's plane and space", Theory of Probability and its Applications, **4**, (1959), 432-436. These authors were particularly interested in convergence to the normal distribution.

REFERENCES

1. S. Bochner, *Harmonic Analysis and the Theory of Probability*, Berkeley and Los Angeles, 1955.

2. C. Caratheodory, *Theory of Functions of a Complex Variable, Volume I*, New York, 1954.

3. A. Erdelyi, *Higher Transcendental Functions, Volume I,* Bateman Manuscript Project. New York, 1953
4. R. Godement, *Introduction aux travaux de A. Selberg,* Séminaire Bourbaki, No. 144 (1956-57).
5. L. Robin, *Fonctions Sphériques de Legendre et Fonctions Spheriodales,* Tome III. Paris, 1959.
6. A. Selberg, *Harmonic analysis and discontinuous groups in weakly symmetric Riemannian spaces with applications to Dirichlet series,* J. Indian Math. Soc., **20** (1956), 47-87.
7. D. V. Widder, *The Laplace Transform,* Princeton, 1946.

SEATTLE WASHINGTON

SEQUENCES IN GROUPS WITH DISTINCT PARTIAL PRODUCTS

BASIL GORDON

1. In an investigation concerning a certain type of Latin square, the following problem arose:

Can the elements of a finite group G be arranged in a sequence $a_1, a_2, \cdots, a_n$ so that the partial products $a_1, a_1a_2, \cdots, a_1a_2 \cdots a_n$ are all distinct?

In the present paper a complete solution will be given for the case of Abelian groups, and the application to Latin squares will be indicated. Let us introduce the term *sequenceable group* to denote groups whose elements can be arranged in a sequence with the property described above. The main result is then contained in the following theorem.

THEOREM 1. *A finite Abelian group G is sequenceable if and only if G is the direct product of two groups A and B, where A is cyclic of order 2^k $(k > 0)$, and B is of odd order.*

Proof (i). To see the necessity of the condition, suppose that G is sequenceable, and let $a_1, a_2, \cdots, a_n$ be an ordering of the elements of G with $a_1, a_1a_2, \cdots, a_1a_2 \cdots a_n$ all distinct. The notation $b_i = a_1a_2 \cdots a_i$ will be used throughout the remainder of the paper. It is immediately seen that $a_1 = b_1 = e$, the identity element of G; for if $a_i = e$ for some $i > 1$, then $b_{i-1} = b_i$, contrary to assumption. Hence $b_n \neq e$, i.e., the product of all the elements of G is not the identity. It is well known (cf [2]) that this implies that G has the form $A \times B$ with A cyclic of order $2^k(k > 0)$ and B of odd order.

(ii) To prove sufficiency of the condition, suppose that $G = A \times B$, with A and B as above. We then show that G is sequenceable by constructing an ordering $a_1, a_2, \cdots, a_n$ of its elements with distinct partial products. From the general theory of Abelian groups, it is known that G has a basis of the form $c_0, c_1, \cdots, c_m$, where c_0 is of order 2^k, and where the orders $\delta_1, \delta_2, \cdots, \delta_m$ of $c_1, c_2, \cdots, c_m$ are odd positive integers each of which divides the next, i.e., $\delta_i \mid \delta_{i+1}$ for $0 < i < m$. If j is any positive integer, then there exist unique integers $j_0, j_1, \cdots, j_m$ such that

$$(1) \qquad \begin{aligned} j &\equiv j_0 (\mathrm{mod}\ \delta_1 \delta_2 \cdots \delta_m) \\ j_0 &= j_1 + j_2\delta_1 + j_3\delta_1\delta_2 + \cdots + j_m\delta_1 \cdots \delta_{m-1} \\ &0 \leqq j_1 < \delta_1 \end{aligned}$$

Received January 3, 1961.

$$0 \leqq j_2 < \delta_2$$
$$\vdots$$
$$0 \leqq j_m < \delta_m .$$

The proof of the existence and uniqueness of this expansion will be omitted here; it is entirely analogous to the expansion of an integer in powers of a number base.

We are now in a position to define the desired sequencing of G. It is convenient to define the products $b_1, b_2, \cdots, b_n$ directly, to prove they are all distinct, and then to verify that the corresponding a_i, as calculated from the formula $a_1 = e$, $a_i = b_{i-1}^{-1}b_i$, are all distinct. If i is of the form $2j + 1 (0 \leqq j < n/2)$, let

$$b_{2j+1} = c_0^{-j}c_1^{-j_1}c_2^{-j_2} \cdots c_m^{-j_m} ,$$

where $j_1, j_2, \cdots, j_m$ are the integers defined in (1). On the other hand, if i is of the form $2j + 2 (0 \leqq j < n/2)$, let

$$b_{2j+2} = c_0^{j+1}c_1^{j_1+1}c_2^{j_2+1} \cdots c_m^{j_m+1} .$$

The elements $b_1, b_2, \cdots, b_n$ thus defined are all distinct. For if $b_s = b_t$ with $s = 2u + 1$, $t = 2v + 1$, then

$$\begin{aligned} u &\equiv v (\mathrm{mod}\ 2^k) \\ u_1 &\equiv v_1 (\mathrm{mod}\ \delta_1) \\ &\vdots \\ u_m &\equiv v_m (\mathrm{mod}\ \delta_m) . \end{aligned} \tag{2}$$

From the inequalities in (1) we conclude that $u_1 = v_1, \cdots, u_m = v_m$. Hence $u_0 = v_0$, so that $u \equiv v (\mathrm{mod}\ \delta_1 \cdots \delta_m)$; coupled with the first of equations (2), this gives $u \equiv v\ (\mathrm{mod}\ n)$, which implies $u = v$. Similarly $b_{2u+2} = b_{2v+2}$ implies $u = v$, so that the "even" b's are distinct.

Next suppose

$$b_{2u+1} = b_{2v+2} .$$

Then

$$\begin{aligned} -u &\equiv v + 1\ (\mathrm{mod}\ 2^k) \\ -u_1 &\equiv v_1 + 1\ (\mathrm{mod}\ \delta_1) \\ &\vdots \\ -u_m &\equiv v_m + 1\ (\mathrm{mod}\ \delta_m) \end{aligned}$$

or equivalently,

$$u + v + 1 \equiv 0\ (\mathrm{mod}\ 2^k) \tag{3}$$

$$\begin{aligned} u_1 + v_1 + 1 &\equiv 0 \pmod{\delta_1} \\ &\vdots \\ u_m + v_m + 1 &\equiv 0 \pmod{\delta_m} . \end{aligned}$$

Since $0 < u_1 + v_1 + 1 \leqq 2(\delta_1 - 1) + 1 < 2\delta_1$, we must have $u_1 + v_1 + 1 = \delta_1$. Reasoning similarly for $i = 2, \cdots, m$ we obtain

$$\begin{aligned} u_1 + v_1 + 1 &= \delta_1 \\ u_2 + v_2 + 1 &= \delta_2 \\ &\vdots \\ u_m + v_m + 1 &= \delta_m . \end{aligned}$$

Multiplying the $(i + 1)$'st equation of this system by $\delta_1\delta_2 \cdots \delta_i$ $(1 \leqq i < m)$ and adding, we get $u_0 + v_0 + 1 = \delta_1 \cdots \delta_m$, which implies $u + v + 1 \equiv o(\delta_1 \cdots \delta_m)$. Combining this with the first of equations (3), we find that $u + v + 1 \equiv 0 \pmod{n}$, which, on account of the inequality $0 < u + v + 1 < n$, is impossible. Hence $b_1, b_2, \cdots, b_n$ are all distinct.

Next we calculate $a_1, a_2, \cdots, a_n$. If $i = 2j + 2$ $(0 \leqq j < n/2)$, then

$$a_i = b_{i-1}^{-1} b_i = c_0^{2j+1} c_1^{2j_1+1} \cdots c_m^{2j_m+1} .$$

These are all different by the same argument as above. If $i = 2j + 1$, and $j_1 \neq 0$, then

$$a_i = c_0^{-2j} c_1^{-2j_1} c_2^{-2j_2-1} \cdots c_m^{-2j_m-1} .$$

If $i = 2j + 1$ and $j_1 = 0$, but $j_2 \neq 0$, then $a_i = c_0^{-2j} c_2^{-2j_2} c_3^{-2j_3-1} \cdots c_m^{-2j_m-1}$, while if $j_1 = j_2 = 0$ but $j_3 \neq 0$, then $a_i = c_0^{-2j} c_3^{-2j_3} c_4^{-2j_4-1} \cdots c_m^{-2j_m-1}$, etc. These a_i's are obviously distinct from each other by the same reasoning as before. Because of the exponent of c_0 they are also distinct from the a_i with i even. This completes the proof of the theorem.

As an example of the construction of Theorem 1, consider the group $G = C_2 \times C_3 \times C_3$. We use basis elements c_0, c_1, c_2 of orders 2, 3, 3 respectively. Using the notation (α, β, γ) for the element $c_0^\alpha c_1^\beta c_2^\gamma$, the sequences a_i and b_i are then the following:

a_i	b_i
(0 0 0)	(0 0 0)
(1 1 1)	(1 1 1)
(0 1 2)	(1 2 0)
(1 0 1)	(0 2 1)
(0 2 2)	(0 1 0)
(1 2 1)	(1 0 1)
(0 0 1)	(1 0 2)
(1 1 0)	(0 1 2)

a_i	b_i
(0 1 0)	(0 2 2)
(1 0 0)	(1 2 2)
(0 2 0)	(1 1 2)
(1 2 0)	(0 0 2)
(0 0 2)	(0 0 1)
(1 1 2)	(1 1 0)
(0 1 1)	(1 2 1)
(1 0 2)	(0 2 0)
(0 2 1)	(0 1 1)
(1 2 2)	(1 0 0)

2. Application to Latin squares. Consider the following Latin square:

1	2	3	4
2	4	1	3
3	1	4	2
4	3	2	1

Given any ordered pair $(\alpha\beta)$ with $\alpha \neq \beta$, it occurs as a pair of consecutive entries in some row of this square. In general, an $n \times n$ Latin square (c_{st}) whose elements are the integers $1, \cdots, n$ will be called *horizontally complete* if for every ordered pair (α, β) with $1 \leqq \alpha,\ \beta \leqq n$ and $\alpha \neq \beta$, the equations

$$c_{st} = \alpha \qquad (4)$$
$$c_{s,t+1} = \beta$$

are solvable. Similarly a vertically complete square is one for which

$$c_{st} = \alpha$$
$$c_{s+1,t} = \beta$$

can be solved for any such choice of α, β. A square which is both horizontally and vertically complete is called complete.

Note that in a horizontally complete square, the solution of equations (4) is unique, since the total number of consecutive pairs $a_{st}, a_{s,t+1}$ is equal to the total number of order pairs (α, β) with $\alpha \neq \beta$. Conversely, uniqueness implies existence for the same reason.

Complete Latin squares are useful in the design of experiments in which it is desired to investigate the interaction of nearest neighbors.

THEOREM 2. *Suppose that G is a sequenceable group, and let a_1, $a_2 \cdots, a_n$ be an ordering of its elements such that $b_1, b_2, \cdots, b_n$ are distinct. Then the matrix $(c_{st}) = (b_s^{-1}b_t)$ is a complete Latin square.*

Proof. It is immediately seen that (c_{st}) is a Latin square, since either $b_s^{-1}b_t = b_s^{-1}b_u$ or $b_t^{-1}b_s = b_u^{-1}b_s$ imply $t = u$ by elementary properties of groups. To show that (c_{st}) is horizontally complete, suppose

$$c_{st} = c_{uv}$$
$$c_{s,t+1} = c_{u,v+1}\ .$$

We must show that $s = u$ and $t = v$. From the definition of c_{st},

$$b_s^{-1}b_t = b_u^{-1}b_v \tag{5}$$
$$b_s^{-1}b_{t+1} = b_u^{-1}b_{v+1}\ . \tag{6}$$

Inverting both sides of (5) yields $b_t^{-1}b_s = b_v^{-1}b_u$. Combining this with (6) we get $(b_t^{-1}b_s)(b_s^{-}b_{t+1}) = (b_v^{-1}b_u)(b_u^{-1}b_{v+1})$, or $b_t^{-1}b_{t+1} = b_v^{-1}b_{v+1}$, i.e., $a_{t+1} = a_{v+1}$. This implies $t = v$. Substituting in (5) we obtain $b_s^{-1}b_t = b_u^{-1}b_t$, from which $s = u$ follows immediately. The proof that (c_{st}) is vertically complete is entirely similar and will be omitted.

This method enables one to construct a complete Latin square of order n for any even n (note that B may be trivial in Theorem 1). Whether or not complete, or even horizontally complete, squares exist for odd n is an open question.

3. Extension to non-Abelian groups. The problem of determining which non-Abelian groups G are sequencable is unsolved at the present time. Considerable information about the nature of a sequence $a_1, \cdots, a_n$ with distinct partial products, if one exists, can be obtained by mapping G onto the Abelian group G/C, where C is the commutator subgroup. Using this technique, for example, it can be shown that the non-Abelian group of order 6 and the two non-Abelian groups of order 8 are not sequencable. On the other hand the non-Abelian group of order 10 is sequencable. To see this, denote its elements by $e, a, b, ab, ba, aba, bab, abab, baba, ababa$, where $a^2 = b^2 = (ab)^5 = e$. A suitable ordering is then given by $e, ab, abab, ababa, bab, aba, b, a, baba, ba$, the partial products being $e, ab, baba, a, abab, bab, ba, b, aba, ababa$. In view of Theorem 1 and the results of [2], one might conjecture that G is sequencable if and only if it does not possess a complete mapping. However, the symmetric group S_3 does not possess a complete mapping (cf [1]) and is also not sequenceable. Whether or not the two properties are at least mutually exclusive is still an open question.

References

1. L. J. Paige, *Complete mappings of finite groups*, Pacific J. Math. **1** (1951), 111-116.

2. M. Hall and L. J. Paige, *Complete mappings of finite groups*, Pacific J. Math. **5** (1955), 541-549.

UNIVERSITY OF CALIFORNIA, LOS ANGELES

RELATIVE SELF-ADJOINT OPERATORS IN HILBERT SPACE

MAGNUS R. HESTENES

1. Introduction. Let A be a closed operator from a Hilbert space $\mathfrak{H}$ to a Hilbert space $\mathfrak{H}'$. The main purpose of this present paper is to develop a spectral theory for an operator A of this type. This theory is analogous to the given in the self-adjoint case and reduces to the standard theory when A is self-adjoint. The spectral theory here given is based on generalization of the concept of self-adjointness. Let A^* denote the adjoint of A. An operator T on $\mathfrak{H}$ to $\mathfrak{H}'$ will be said to be an elementary operator if $TT^*T = T$. If T is elementary, the operator TA^*T can be considered to be an adjoint of A relative to T. If $A = TA^*T$, then A will be said to be self-adjoint relative to T. The polar decomposition theorem for A implies the existence of a unique elementary operator R relative to which A is self-adjoint and having the further property that R has the same null space as A and that A^*R is a nonnegative self-adjoint operator in the usual sense. Every elementary operator T relative to which A is self-adjoint is of the form $T = T_0 + R_1 - R_2$, where $R = R_1 + R_2$ and T_0, R_1, R_2 are $*$-orthogonal. Two operators B and C are said to be $*$-orthogonal if $B^*C = 0$ and $BC^* = 0$ on dense sets in $\mathfrak{H}$ and $\mathfrak{H}'$ respectively.

An operator B will be called a section of an operator A if there is an operator C $*$-orthogonal to B such that $A = B + C$. If R is the elementary operator associated with A, there exists a one parameter family A_λ, R_λ $(0 < \lambda < \infty)$ of sections of A, R respectively such that R_λ is the elementary operator belonging to A_λ, $\| A_\lambda \| \leqq \lambda$, A_μ $(\mu < \lambda)$ is a section of A_λ and $A = \int_0^\infty \lambda dR_\lambda$. From this result it is seen that A possesses a spectral decomposition relative to any elementary operator T relative to which A is self-adjoint. These results can be extended to the case in which A is normal relative to T. When $\mathfrak{H}' = \mathfrak{H}$ and T is the identity, these results give the usual spectral theory for self-adjoint operators. Examples are given in §§ 4 and 10 below. In particular spectral resolutions are given for the gradient of a function and its adjoint, the divergence of a vector. The finite dimensional case has been treated in a recent paper by the author[1].

The results given below are elementary in nature and are based

Received October 11, 1960. The preparation of this paper was sponsored in part by the Office of Naval Research and the U. S. Army Research Office. Reproduction in whole or in part is permitted for any purpose of the U. S. Government.

[1] M. R. Hestenes, Relatively hermitian matrices, to be published in the Pacific Journal of Mathematics.

upon the fundamental ideas concerning Hilbert spaces. These ideas can be found in the standard treatises on Hilbert space. The concept of $*$-commutativity is introduced. This concept is used in the development of the spectral theory. It is shown that a reciprocally compact operator has a discrete principal spectrum. The concept of reciprocal compactness is connected with the concept of ellipticity of differential operators, as is indicated in the last section below.

2. Preliminaries. Let $\mathfrak{H}$ and $\mathfrak{H}'$ be two Hilbert Spaces over a scalar field $\mathfrak{C}$. The field $\mathfrak{C}$ will be taken to be either the field of real numbers or the field of complex numbers. The two case can be treated simultaneously by defining the conjugate $\bar{b}$ of b to be b itself in the field of reals. The spaces $\mathfrak{H}$ and $\mathfrak{H}'$ may coincide. The same notations will be used for the inner product in each of the two spaces. Thus, the symbol (x_1, x_2) denotes the *inner product* of x_1 and x_2, whether x_1 and x_2 are in $\mathfrak{H}$ or in $\mathfrak{H}'$. The *norm* of x will be denoted by $\| x \|$. *Strong convergence* of a sequence $\{x_n\}$ to x_0 will be denoted by $x_n \Rightarrow x_0$ and *weak convergence* by $x_n \rightarrow x_0$.

The *closure* of a subclass $\mathscr{B}$ of $\mathfrak{H}$ will be denoted by $\overline{\mathscr{B}}$ and its *orthogonal complement* in $\mathfrak{H}$ by $\mathscr{B}^{\perp}$. Clearly $\mathscr{B}^{\perp}$ is a subspace of $\mathfrak{H}$. By the sum $\mathfrak{A} + \mathscr{B}$ of two linear subclasses $\mathfrak{A}$ and $\mathscr{B}$ will be meant the class of all elements of the form $x + y$ with x in $\mathfrak{A}$ and y in $\mathscr{B}$. It will be called a *direct sum* if $\mathfrak{A}$ and $\mathscr{B}$ have no nonnull elements in common.

A linear tranformation A will be said to be *from* $\mathfrak{H}$ to $\mathfrak{H}'$ if its *domain* $\mathfrak{D}_A$ is in $\mathfrak{H}$ and its *range* $\mathscr{R}_A$ is in $\mathfrak{H}'$. If $\mathscr{D}_A = \mathfrak{H}$ the phrase "on $\mathfrak{H}$ to $\mathfrak{H}'$" will be used to emphasize this fact. The phrase "A in $\mathfrak{H}$" will be used occassionally in case $\mathfrak{H}' = \mathfrak{H}$. A linear transformation B from $\mathfrak{H}$ to $\mathfrak{H}'$ will be called an *extension* of A, written $A \leqq B$ or $B \geqq A$, in case $\mathscr{D}_B \supset \mathscr{D}_A$ and $B = A$ on $\mathscr{D}_A$. If $\overline{\mathscr{D}_A} = \mathfrak{H}$, then A will be said to be *dense* in $\mathfrak{H}$. The transformation A will be said to be *bounded* if it maps bounded subsets of $\mathscr{D}_A$ into bounded sets of $\mathfrak{H}'$. If A is bounded, its *norm* $\| A \|$ is defined to be the least upper bound of $\| Ax \|$ for all x in $\mathscr{D}_A$ having $\| x \| = 1$. If whenever $x_n \in D_A$, $x_n \Rightarrow x_0$, $Ax_n \Rightarrow y_0$ we also have $x_0 \in \mathscr{D}_A$ and $Ax_0 = y_0$, then A will be said to be *closed*. If whenever $x_0, x_n \in \mathscr{D}_A$ and $x \Rightarrow x_0$, $Ax_n \Rightarrow y_0$ we have $Ax_0 = y_0$ then A is said to be *preclosed*. A closed dense linear transformation is bounded if and only if $\mathscr{D}_A = \mathfrak{H}$. The minimal closed extension of A, if it exists, will be called the *closure* of A and will be denoted by $\bar{A}$. If A is preclosed, its closure exists. By the *null class* $\mathfrak{N}_A$ of A will be meant all x in $\mathscr{D}_A$ such that $Ax = 0$. There is a unique extension of A whose domain is $\mathscr{D}_A + \overline{\mathfrak{N}}_A$ and whose null space is $\overline{\mathfrak{N}}_A$. If A is closed then $\mathfrak{N}_A$ is closed.

Consider now a dense linear transformation A from $\mathfrak{H}$ to $\mathfrak{H}'$ and let

$\mathscr{D}_A{}^*$ be the class of all vectors y in $\mathfrak{H}'$ for which there exists a vector A^*y in $\mathfrak{H}$ such that the relation.

$$(Ax, y) = (x, A^*y)$$

holds for all x in $\mathscr{D}_A$. The transformation A^* from $\mathfrak{H}'$ to $\mathfrak{H}$ so defined is a closed linear transformation whose domain is $\mathscr{D}_A{}^*$, whose null class $\mathfrak{N}_A{}^*$ is $\mathscr{R}_A^\perp$ and whose range $\mathscr{R}_A{}^*$ is a subclass of $\mathfrak{N}_A^\perp$.

A linear transformation A from $\mathfrak{H}$ to $\mathfrak{H}$ will be said to be *self-adjoint* if it is dense and if $A^* = A$. A self-adjoint linear transformation A will be said to be *nonnegative*, written $A \geqq 0$, if the inequality $(Ax, x) \geqq 0$ holds for all x in $\mathscr{D}_A$. By a *projection* E in $\mathfrak{H}$ will be meant a self-adjoint operator such that $E^2 = E$.

It will be convenient to use the term "*operator*" to denote a closed dense linear transformation. We shall have occassion to use the following well known result.

THEOREM 2.1. *Let A be an operator from $\mathfrak{H}$ to $\mathfrak{H}'$. Then its adjoint A^* is an operator from $\mathfrak{H}'$ to $\mathfrak{H}$. Moreover, $A^{**} = A$, $\mathfrak{N}_A{}^* = \mathscr{R}_A^\perp$, $\mathfrak{N}_A = \mathscr{R}_A^{\perp *}$. For each vector x_0 in $\mathfrak{H}$ and y_0 in $\mathfrak{H}'$ there is a unique vector x in $\mathscr{D}_A$ and y in $\mathscr{D}_A{}^*$ such that*

$$x_0 = x + A^*y\ , \quad y_0 = Ax - y\ . \tag{2.1}$$

The transformation A^*A is a nonnegative self-adjoint operator in $\mathfrak{H}$ whose null space is $\mathfrak{N}_A$. Similarly AA^* is a nonnegative self-adjoint operator in $\mathfrak{H}'$ whose null space is $\mathfrak{N}_A{}^*$. The operator A is bounded if and only if A^* is bounded. In this event $\| A \| = \| A^* \|$.

3. The reciprocal and *-reciprocal of a closed operator. Consider a linear transformation A from $\mathfrak{H}$ to $\mathfrak{H}'$ whose domain $\mathfrak{D}_A$ is expressible as a direct sum $\mathscr{D}_A = \mathscr{C}_A + \mathfrak{N}_A$, where $\mathfrak{N}_A$ is the null space of A and $\mathscr{C}_A$ is orthogonal to $\mathfrak{H}_A$. The class $\mathscr{C}_A$ will be called the *carrier* of A. If $\mathfrak{N}_A$ is closed, then $\mathscr{D}_A$ has such a representation. Consequently, the carrier of a closed linear transformation is well defined.

The transformation A establishes a one-to-one correspondence between its carrier and its range. The inverse transformation on $\mathscr{R}_A$ onto $\mathscr{C}_A$, when extended linearly so as to have $\mathscr{R}_A^\perp$ as its null space and $\mathscr{R}_A + \mathscr{R}_A^\perp$ as its domain, defines a linear transformation A^{-1} which will be called the *reciprocal*[2] of A. The carrier of A^{-1} is the range of A and the range of A^{-1} is the carrier of A. It is clear that A^{-1} is dense in $\mathfrak{H}'$ and that $\mathfrak{N}_{A-1}$ is closed.

The reciprocal of A^{-1} for an arbitrary linear transformation A will

[2] See E. H. Moore General analysis I, Memoirs, American Philosophical Society (1935). See also, J. von Neumann, *On regular rings*, Proc. Nat. Acad. Sci, **22**, (1936), 707-715.

be defined to be the reciprocal of the extension A_0 of A whose domain is $\mathscr{D}_A + \overline{\mathfrak{N}}_A$ and whose null space is $\overline{\mathfrak{H}}_A$. The carrier of A_0 will be defined to be the carrier of A. The reciprocal of A^{-1} is accordingly the extension of A whose domain is $\mathscr{D}_A + \overline{\mathfrak{N}}_A + \mathscr{D}_A^{\perp}$ and whose null space is $\overline{\mathfrak{N}}_A + \mathscr{D}_A^{\perp}$. Hence A is the reciprocal of A^{-1} if and only if A is dense in $\mathfrak{H}$ and its null space is closed. If $\mathfrak{N}_A$ is closed, then A is closed if and only if A^{-1} is closed. If A possesses an inverse, then A^{-1} is the inverse of A.

Theorem 3.1. *The adjoint of the reciprocal of an operator A is the reciprocal of its adjoint, that is, $(A^{-1})^* = (A^*)^{-1}$. The operators A and A^{-1*} have the same null spaces.*

Clearly $\mathfrak{N}_{A^{-1*}} = \mathscr{R}_{A^{-1}}^{\perp} = \mathfrak{N}_A$. Let x be a vector in $\mathscr{C}_{A^{-1*}}$. Then $(A^{-1}y_0, x) = (y_0, A^{-1*}x)$ for every y_0 in $\mathscr{C}_{A^{-1}} = \mathscr{R}_A$. Hence $(x_0, x) = (Ax_0, A^{-1*}x)$ if $x_0 \in \mathscr{C}_A$ and hence if $x_0 \in \mathscr{D}_A$. It follows that $A^{-1*}x$ is in $\mathscr{D}_A$ and that $x = A^*A^{-1}x$. Consequently $\mathscr{C}_{A^{-1*}} \subset \mathscr{R}_{A^*} = \mathscr{C}_{A^{*-1}}$. Conversely if $x \in \mathscr{C}_{A^{*-1}}$, then $(x_0, x) = (Ax_0, A^{*-1}x)$ holds for all x_0 in $\mathscr{C}_A$ or equivalently $(A^{-1}y_0, x) = (y_0, A^{*-1}x)$ holds for all y_0 in $\mathscr{C}_{A^{-1*}}$. It follows that x is in $\mathscr{C}_{A^{-1*}}$ and that $A^{-1*}x = A^{*-1}x$. It follows that A^{-1*}, A^{*-1} coincide on their carriers, as well as their null spaces and hence are identical.

The element A^{*-1} plays an important role in the results given below and will be called the *$*$-reciprocal* of A.

As an immediate consequence of the last theorem we have

Theorem 3.2. *Let A be an operator from $\mathfrak{H}$ to $\mathfrak{H}'$. Then A^{-1}, A^*, $A^{*-1} = A^{-1*}$ are operators. The products A^*A, $A^{-1}A^{*-1}$ are nonnegative self-adjoint operators, are reciprocals of each other and have the same null space as A. Similarly, the products AA^*, $A^{*-1}A^{-1}$ are nonnegative self-adjoint operators, are reciprocals of each other and have the same null space as A^*.*

A linear transformation A will be said to be *reciprocally bounded* if its reciprocal is bounded, or, equivalently if there is a positive number $m > 0$ such that $\|Ax\| \geqq m\|x\|$ on the carrier of A. The following theorem is self-evident.

Theorem 3.3. *Let A be an operator from $\mathfrak{H}$ to $\mathfrak{H}'$. Then A is reciprocally bounded if and only if its range is closed. Hence A is reciprocally bounded if and only if the equation $Ax = y$ has a solution x in $\mathscr{D}_A$ whenever y is orthogonal to every solution z of $A^*z = 0$. The operator A is reciprocally bounded if and only if A^* is reciprocally bounded. Finally, A is reciprocally bounded if and only if A^*A (or AA^*) is reciprocally bounded.*

The concept of reciprocal boundedness is the basis for a large class of existence theorems for ordinary and partial differential equations. In view of the last conclusion in the theorem existence theorems for non-self-adjoint problems follows from those for self-adjoint problems.

THEOREM 3.4. *Let A be an operator from $\mathfrak{H}$ to $\mathfrak{H}'$. If α and β are positive numbers, then*

$$\alpha A + \beta A^{*-1} , \quad \alpha A^* + \beta A^{-1}$$

are reciprocally bounded operators and are adjoints of each other.

In order to prove the theorem it is sufficient to consider a transformation of the form $B = \lambda A + (1/\lambda)A^{*-1}$, where λ is a positive number. Let y_0 be a vector in $\overline{\mathscr{R}}_A$. By Theorem 2:1, with A replaced by λA, there is a unique vector x in $\mathscr{D}_A$ and y in $\mathscr{D}_{A^*}$ such that

$$0 = x + \lambda A^* y , \quad y_0 = \lambda A x - y .$$

The vector y is therefore in $\overline{\mathscr{R}}_A = \mathfrak{N}_{A^*}^{\perp}$ and in the carrier of A^*. Consequently, $y = (1/\lambda)A^{*-1}x$ and

$$y_0 = \left(\lambda A + \frac{1}{\lambda} A^{*-1}\right)x = Bx .$$

The range of B is therefore closed. It follows that B is reciprocally bounded and closed. Similarly $C = \lambda A^* + (1/\lambda)A^{-1}$ is reciprocally bounded and closed. Clearly $C = B^*$. This proves the theorem.

COROLLARY 1. *If A is self-adjoint operator, and α, β are positive numbers, then $\alpha A + \beta A^{-1}$ is a reciprocally bounded self-adjoint operator, Moreover the reciprocal of A^2 is $A^{-2} = (A^{-1})^2$.*

THEOREM 3.5. *Let $C = BA$, $D = A^{-1}B^{-1}$, where A is an operator from $\mathfrak{H}$ to $\mathfrak{H}'$ and B is an operator from $\mathfrak{H}'$ to a Hilbert space $\mathfrak{H}''$. Suppose that $\mathfrak{N}_{A^*} = \mathfrak{N}_B$. Then $\mathfrak{N}_C = \mathfrak{N}_A$, $\mathfrak{N}_D = \mathfrak{N}_{B^*}$. If D is dense, then $D = C^{-1}$. If either A or B^{-1} is bounded then C and D are closed.*

Suppose that $x_n \in \mathscr{D}_C$, $x_n \Rightarrow x_0$, $Cx_n \Rightarrow z_0$. Set $y_n = Ax_n$, $z_n = By_n = Cx_n$. If A is bounded, then $y_n \Rightarrow Ax_0$. Since $By_n = Cx_n \Rightarrow z_0$ it follows that $z_0 = BAx_0 = Cx_0$. Consequently C is closed. Observe that this conclusion is valid even if $\mathfrak{N}_{A^*} \neq \mathfrak{N}_B$. Since $y_n \in \mathscr{C}_B$ we have $y_n = Ax_n = B^{-1}z_n$. If B^{-1} is bounded, then $y_n = Ax_n \Rightarrow B^{-1}z_0$. Hence $B^{-1}z_0 = Ax_0$, that is $z_0 = BAx_0 = Cx_0$. Consequently C is closed in this event also. The remaining statements in the theorem are readily verified.

COROLLARY. *If A is bounded and reciprocally bounded and $\mathfrak{N}_{A^*} \supset \mathfrak{N}_B$, then the products C and D described in Theorem 3.5 are operators and are reciprocals of each other.*

This follows readily from Theorem 3.5 because we can replace A by $F'A$ where F' is the projection in $\mathfrak{H}'$ whose null class is $\mathfrak{N}_B$. We then have $\mathfrak{N}_{A^*} = \mathfrak{N}_B$.

4. Examples. The results here given were motivated in part by certain applications to differential equations. It will be convenient to explain in part two of these applications at this time.

EXAMPLE 1. Let $\mathfrak{H}$ be the class of all real valued Lebesgue square integrable functions x in the interval $0 \leqq t \leqq \pi$. This class with

$$(x, y) = \int_0^\pi x(t)y(t)dt$$

as its inner product and the real numbers as scalars from a Hilbert space. Let $\mathfrak{A}$ be the class of all absolutely continuous functions $x(t)$ $(0 \leqq t \cong \pi)$ whose derivatives $\dot{x}$ are in $\mathfrak{H}$. Let A be the differential operator d/dt having as its domain the class $\mathscr{D}_A$ of all functions in $\mathfrak{A}$ having $x(0) = x(\pi) = 0$. The carrier of A is $\mathscr{D}_A$ itself. Then range $\mathscr{R}_A$ consists of all functions y in $\mathfrak{H}$ satisfying the condition

$$\int_0^\pi y(t)dt = 0 \,. \tag{4:1}$$

Since $\mathscr{R}_A$ is closed it follows that A is reciprocally bounded. The reciprocal of A is

$$A^{-1}y = \int_0^t y(s)ds - \frac{t}{\pi}\int_0^\pi y(s)ds \,.$$

The adjoint A^* of A is the operator $-d/dt$ with $\mathscr{D}_{A^*} = \mathfrak{A}$ as its domain and $\mathfrak{A} \cap \mathscr{R}_A$ as its carrier. Since A is reciprocally bounded so also A^*. Moreover $\mathscr{R}_{A^*} = \mathfrak{H}$. The reciprocal of A^* is

$$A^{*-1}x = -\int_0^t x(s)ds + \frac{1}{\pi}\int_0^\pi\int_0^r x(s)dsdr$$

by virtue of the relation (4.1). Let $\mathscr{B}$ be all functions in $\mathfrak{A}$ whose derivatives are also in $\mathfrak{A}$. The operator A^*A is the operator $-d^2/dt^2$ having as its domain all functions in $\mathscr{B}$ such that $x(0) = x(\pi) = 0$. The range of A^*A is $\mathfrak{H}$. The operator AA^* is the operator $-d^2/dt^2$ having as its domain all functions x in $\mathscr{B}$ whose derivative $\dot{x}$ satisfies the conditions $\dot{x}(0) = \dot{x}(\pi) = 0$. The range of AA^* coincides with that of A.

The operator AA^* is also reciprocally bounded.

A preview of the theory to be presented below can be given for this example by recalling certain known facts. Let

$$x_n(t) = \sqrt{\frac{2}{\pi}}\sin nt\ , \quad y_0(t) = \sqrt{\frac{1}{\pi}}\ , \ y_n(t) = \sqrt{\frac{2}{\pi}}\cos nt \quad (n = 1, 2, 3, \cdots)\ .$$

The function x_n form a complete orthonormal system in $\mathfrak{H}$. A function x in $\mathfrak{H}$ is accordingly given by the fourier sine series.

$$x = \sum_{n=1}^{\infty} a_n x_n\ , \quad a_n = (x, x_n)$$

where convergence is taken to be convergence in the mean of order 2. Similarly a function y in $\mathfrak{H}$ is expressible in the form

$$y = b_0 y_0 + \sum_{n=1}^{\infty} b_n y_n\ , \qquad b_j = (y, y_j)\ (j = 1, 2, \cdots)\ .$$

If x and y are in the appropriate domains we have

$$\begin{aligned} Ax &= \sum_{n=1}^{\infty} n a_n y_n\ , \quad A^{-1}y = \sum_{1=n}^{\infty} \frac{1}{n} b_n x_n\ , \\ A^{*-1}x &= \sum_{n=1}^{\infty} \frac{1}{n} a_n y_n\ , \quad A^*y = \sum_{n=1}^{\infty} n b_n x_n\ , \end{aligned} \tag{4.2}$$

as one readily verifies. These formulas can be put in another form by defining the operators R and R_i $(i = 1, 2, 3, \cdots)$ by the formulas

$$Rx = \sum_{n=1}^{\infty} a_n y_n\ , \quad R_i x = a_i y_i \qquad (i = 1, 2, 3, \cdots)\ .$$

Observe that

$$R^*y = \sum_{n=1}^{\infty} b_n x_n\ , \quad R_i^* y = b_i x_i \qquad (i = 1, 2, 3, \cdots)\ .$$

The operator R maps $\mathfrak{H}$ isometrically onto $\mathscr{R}_A$. Its adjoint R^* maps $\mathscr{R}_A$ isometrically onto $\mathfrak{H}$ and annihilates $\mathscr{R}_A^{\perp}$. We have the relations

$$\begin{aligned} &R = \sum_{n=1}^{\infty} R_n\ , \quad R^*R_n = R_n^*R_n\ , \quad RR_n^* = R_nR_n^* \\ &R_i^*R_j = 0\ , \quad R_iR_j^* = 0 \quad (i \neq j)\ . \end{aligned} \tag{4.3}$$

Moreover, by (4.2) we have

$$\begin{aligned} A &= \sum_{n=1}^{\infty} nR_n\ , \quad A^{-1} = \sum_{n=1}^{\infty} \frac{1}{n} R_n^* \\ A^{*-1} &= \sum_{n=1}^{\infty} \frac{1}{n} R_n\ , \quad A^* = \sum_{n=1}^{\infty} nR_n^*\ . \end{aligned} \tag{4.4}$$

These formulas constitute a spectral resolution of A, A^{-1}, A^{*-1}, A^*. It is our purpose to show that every operator A can be resolved in terms of elementary operators having the properties similar to those given in (4.3).

The example just given can be modified so as to include all complex valued functions in $\mathfrak{H}$ and so that $A = i(d/dt)$. Then A is a symmetric operator but is not self-adjoint. The theory for this case is not significantly different from that just described.

EXAMPLE 2. Let $\mathfrak{H}$ be the class of all real valued Lebesgue square integrable functions $x(s, t)$ on the square $0 \leqq s \leqq \pi$, $0 \leqq s \leqq \pi$. Then $\mathfrak{H}$ together with the inner product $(x, y) = \int_0^\pi\int_0^\pi x(s, t)y(s, t)dsdt$ defines a Hilbert space with the real numbers as its scalar field. Let $\mathfrak{A}$ be the class of all functions x in $\mathfrak{H}$ such that

(i) $x(s, t)$ is absolutely continuous in s on $0 \leqq s \leqq \pi$ for almost all t on $0 \leqq t \leqq \pi$ and is absolutely continuous in t on $0 \leqq t \leqq \pi$ for almost all s on $0 \leqq s \leqq \pi$;

(ii) The partial derivatives x_s, x_t, (which exist almost everywhere) are in $\mathfrak{H}$. Let $\mathfrak{H}'$ be the Hilbert space defined by the cartesian product $\mathfrak{H} \times \mathfrak{H}$. Observe that the gradient of x, written grad x, is defined on $\mathfrak{A}$ and maps $\mathfrak{A}$ into $\mathfrak{H}'$.

We shall be concerned with the operator $Ax = \text{grad } x$ whose domain $\mathscr{D}_A$ consists of all functions x in $\mathfrak{A}$ which vanish on the boundary, in the sense that $x(0, t) = x(\pi, t) = 0$ for almost all t on $0 \leqq t \leqq \pi$ and $x(s, 0) = (x\,\pi) = 0$ for almost all s on $0 \leqq s \leqq \pi$. It can be shown that the mapping A so defined is a closed dense operator A from $\mathfrak{H}$ to $\mathfrak{H}'$. In fact it is the closure of the transformation grad x restricted to functions of class C' that vanish on the boundary of the given square. Its adjoint A^* is defined by $A^*y = -\text{div } y$, where div y is the closure of the usual divergence operator defined on the class of all vectors y in $\mathfrak{H}'$ of class C'. The ranges of A and A^* are closed. Consequently A and A^* are reciprocally bounded. The operators A^{-1} and A^{*-1} are bounded and can be given an integral representation but we shall not pause to do so here.

The functions

$$x_{mn}(s, t) = \frac{2}{\pi}\sin ms \sin nt \qquad (m, n = 1, 2, 3, \cdots)$$

form a complete orthonormal system in $\mathfrak{H}$. Consequently every vector x in $\mathfrak{H}$ can be expressible in the form

$$x = \sum_{m,n=1}^{\infty} a_{mn}x_{mn} ,$$

where convergence is taken in the mean of order 2. The vectory y_{mn} in $\mathfrak{H}'$ whose components are

$$\frac{2m}{\pi\sqrt{m^2+n^2}}\cos ms \sin nt\ , \quad \frac{2n}{\pi\sqrt{m^2+n^2}}\sin ms \cos nt$$

form an orthonormal system in $\mathfrak{H}'$ that is incomplete. However, it is complete in $\mathscr{R}_A$. Consequently every vector y in $\mathfrak{H}'$ is expressible in the form

$$y = y_0 + \sum_{m,n=1}^{\infty} b_{mn} y_{mn}$$

where $y_0 \in \mathscr{R}_A^{\perp}$, that is, $A^* y_0 = 0$. If x and y are in the appropriate domains we have.

$$Ax = \sum \lambda_{mn} a_{mn} y_{mn}\ , \quad A^{-1}y = \sum \frac{b_{mn}}{\lambda_{mn}} x_{mn}$$

$$A^{*-1}x = \sum \frac{a_{mn}}{\lambda_{mn}} y_{mn}\ , \quad A^* y = \sum \lambda_{mn} b_{mn} x_{mn}$$

where $\lambda_{mn} = (m^2+n^2)^{1/2}$ and m, n summed over the positive integers. Defining R and R_{mn} by the formulas.

$$Rx = \sum a_{mn} y_{mn}\ , \quad R_{mn}x = a_{mn} y_{mn}$$

it is found that R and R_{mn} satisfies relation analogous to (4.3) and that R maps $\mathfrak{H}$ isometrically onto $\mathscr{R}_A$. Moreover,

$$A = \sum \lambda_{mn} R_{mn}\ , \quad A^{-1} = \sum \lambda_{mn}^{-1} R_{mn}^*$$
$$A^{*-1} = \sum \lambda_{mn}^{-1} R_{mn}\ , \quad A^* = \sum \lambda_{mn} R_{mn}^*\ .$$

These formulas are analogous to (4:4) and with minor modifications illustrate the spectral theory given below for an arbitrary closed operator whose reciprocal is compact.

5. Some properties of nonnegative self-adjoint operators. It is the purpose of this section to establish certain properties of nonnegative self-adjoint operators. The first of these is given in the following

THEOREM 5.1. *Let A be a nonnegative self-adjoint operator from $\mathfrak{H}$ to $\mathfrak{H}$ and let E be the projection*

(5.1) $$E = \overline{A^{-1}A} = \overline{AA^{-1}}\ .$$

There exists a unique pair of nonnegative self-adjoint operators C and D such that

(5.2) $$C + D = E\ , \quad A = CD^{-1} = D^{-1}C\ , \quad A^{-1} = C^{-1}D = DC^{-1}\ .$$

The operators C and D are bounded and are given by the formulas

$$C^{-1} = A^{-1} + E\,, \quad D^{-1} = A + E\,. \tag{5.3}$$

They have the same null space as A. Moreover

$$CD = DC\,, \quad C^{-1}D^{-1} = D^{-1}C^{-1} = C^{-1} + D^{-1}\,. \tag{5.4}$$

In order to prove this result let C and D be defined by the formula (5.3). Then C and D are bounded. In fact $\|C\| \leqq 1$, $\|D\| \leqq 1$. The set $\mathscr{D} = \mathscr{D}_A \cap \mathscr{D}_{A^{-1}}$ is the domain of each of the transformations $C^{-1}D^{-1}$, $D^{-1}C^{-1}$, $C^{-1} + D^{-1}$. In view of (5:1) we have

$$C^{-1}D^{-1} = A^{-1}A + A + A^{-1} + E = C^{-1} + D^{-1} = D^{-1}C^{-1}\,.$$

These operators are accordingly reciprocally bounded operators and are the reciprocals of CD and DC, by Theorem 3.5. Hence (5.4) holds. In addition

$$\begin{aligned} C^{-1}C &= D^{-1}D = E = D^{-1}C^{-1}CD = (C^{-1} + D^{-1})CD = D + C\,, \\ C^{-1} &= C^{-1}(C + D) = E + C^{-1}D = (C + D)C^{-1} = E + DC^{-1}\,, \\ D^{-1} &= D^{-1}(C + D) = D^{-1}C + E = (C + D)D^{-1} = CD^{-1} + E\,. \end{aligned}$$

Comparing this result with (5.3) it is seen that (5.2) holds. On the other hand equations (5.2) imply that A, C, D have the same null space and it follows from the computation just made that (5.3) holds. This proves the theorem.

THEOREM 5.2. *Let A be a nonnegative self-adjoint operator from $\mathfrak{H}$ to $\mathfrak{H}$. There is a unique nonnegative self-adjoint operator P from $\mathfrak{H}$ to $\mathfrak{H}$ such that $P^2 = A$. The operator P will be called the square root of A and will be denoted alternatively by $A^{1/2}$. The square root of A^{-1} is P^{-1}.*

If A is bounded, this result can be established by elementary means[3]. In this event every bounded self-adjoint operator that commutes with A also commutes with $A^{1/2}$. The truth of the theorem for the unbounded case can be obtained from the spectral theorem. Assuming the truth of the theorem for the bounded case one can establish its truth for the unbounded case without the direct use of the spectral theorem. As a first step in the proof we shall prove the following

LEMMA 5.1. *Let P and A be two self-adjoint operators from $\mathfrak{H}$ to $\mathfrak{H}$ such that $P^2 = A$. Then A is nonnegative and $(P^{-1})^2 = A^{-1}$.*

Clearly A is nonnegative. In order to show that $C = (P^{-1})^2$ is the

[3] See, for example, F. Reisz and B. Nagy, Lecons d'Analyse Fonctionelle, p. 262.

reciprocal of A observe that on $\mathscr{D}_A$ we have $CA = (P^{-1}P)^2 \subseteqq E$, where $E = \overline{A^{-1}A}$. It follows that $C \supseteqq A^{-1}$. Since C and A^{-1} are self-adjoint we have $A^{-1} = (A^{-1})^* \supseteqq C^* = C$. Hence $A^{-1} = C$, as was to be proved.

COROLLARY. *A reciprocally bounded nonnegative self-adjoint operator possesses a unique square root.*

We are now in position to complete the proof of Theorem 5.2. To this end let C and D be related to A as described in Theorem 5.1. Since C and D are bounded and commute, their square roots M and N satisfy the relations

$$M^2 + N^2 = E\,, \quad MN = NM\,, \quad M^{-1}N^{-1} = N^{-1}M^{-1}\,.$$

Moreover $E = M^{-1}M = N^{-1}N$ and

$$N^{-1} = N^{-1}(M^2 + N^2) = N^{-1}M^2 + N = (M^2 + N^2)N^{-1} = M^2N^{-1} + N\,.$$

Hence $N^{-1}M^2 = M^2N^{-1}$ and

$$\begin{aligned} MN^{-1} &= EMN^{-1} = N^{-1}M^2N^{-1} = M^{-1}N^{-1}M^2 = N^{-1}M^{-1}M^2 \\ &= N^{-1}EM = N^{-1}M \end{aligned}$$

Similarly $M^{-1}N = NM^{-1}$. In addition

$$(N^{-1}M)^2 = N^{-1}MN^{-1}M = N^{-2}M^2 = D^{-1}C = A\,, \quad (M^{-1}N)^2 = A^{-1}\,.$$

Setting $y = Nx$ with x in the carrier of N and using the fact that $MN = NM \geqq 0$ we find that

$$(MN^{-1}y, y) = (Mx, Nx) = (MNx, x) \geqq 0$$

for all y in the carrier of MN^{-1}. Hence MN^{-1} is a nonnegative self-adjoint operator whose square is A. It remains to show that if P is a nonnegative self-adjoint operator whose square is A, then $P = MN^{-1}$. To do so observe that

$$(P + P^{-1})^2 = P^2 + P^{-2} + 2E = A + A^{-1} + 2E = C^{-1}D^{-1} = (M^{-1}N^{-1})^2\,.$$

Since reciprocally bounded operators have unique square roots it follows that

$$P + P^{-1} = M^{-1}N^{-1}\,.$$

Moreover

$$\begin{aligned} &PM^{-2} = PC^{-1} = PA^{-1} + P \subseteqq P + P^{-1} = M^{-1}N^{-1} = N^{-1}M^{-1} \\ &P = PM^{-2}M^2 \subseteqq N^{-1}M^{-1}M^2 = N^{-1}M\,. \end{aligned}$$

Since P and $N^{-1}M$ are self-adjoint, they are equal. This completes the proof of Theorem 5.2.

6. Elementary operators and the polar form. By an *elementary operator* R from $\mathfrak{H}$ to $\mathfrak{H}'$ will be meant one that is its own $*$-reciprocal, or equivalently one whose adjoint is its reciprocal. It is characterized by the relation

$$RR^*R = R \,. \tag{6.1}$$

An elementary operator maps its carrier isometrically onto its range. If $R \neq 0$ then $\|R\| = 1$. It is easily seen that an operator R is elementary if and only if $E = R^*R$ is a projection in $\mathfrak{H}$. Similarly R is elementary if and only if $E' = RR^*$ is a projection in $\mathfrak{H}'$. If $\mathfrak{H} = \mathfrak{H}'$, then an elementary operator R is normal if and only if $E = E'$, that is, if and only if R and R^* have the same null spaces. A projection is a nonnegative self-adjoint elementary operator. An elementary operator R is self-adjoint if and only if it is expressible as the difference $R = E_+ - E_-$ of two projections E_+ and E_- that are orthogonal. For if R is self-adjoint, then

$$E_+ = \tfrac{1}{2}(E + R) \,, \quad E_- = \tfrac{1}{2}(E - R)$$

satisfy the relations

$$E_+^2 = E_+ = E_+^* \,, \quad E_-^2 = E_- = E_-^* \,, \quad E_+E_- = E_-E_+ = 0$$

and hence are projections. Moreover

$$R = E_+ - E_- \,, \quad E = E_+ = E_- = R^2 \,.$$

Conversely, if R is expressible in this form it is a self-adjoint elementary operator, as one readily verifies.

It should be observed in passing that if R is an elementary operator from $\mathfrak{H}$ to $\mathfrak{H}'$ and F is a projection in $\mathfrak{H}$ that commutes with the projection $E = R^*R$, then $S = RF$ is also elementary. This follows because $S^*S = FR^*RF = FEF = FE$ is projection. Similarly if F' is a projection in $\mathfrak{H}'$ that commutes with RR^*, then $F'R$ is elementary.

Let R be an elementary operator from $\mathfrak{H}$ to $\mathfrak{H}'$. An operator A from $\mathfrak{H}$ to $\mathfrak{H}'$ will be said to be *self-adjoint relative to* R in case

$$A = RA^*R \,. \tag{6.2}$$

If $\mathfrak{H} = \mathfrak{H}'$ and R is the identity, this concept reduces to the usual definition of self-adjointness. We have the following

THEOREM 6.1. *Let A be an operator from $\mathfrak{H}$ to $\mathfrak{H}'$ that is self-adjoint relative to an elementary operator R. Then A^{*-1} is self-adjoint relative to R. Similarly A^*, A^{-1} are self-adjoint relative to*

R^*. *Moreover* $\mathfrak{N}_R \subset \mathfrak{N}_A$ *and* $\mathfrak{N}_{R^*} \subset \mathfrak{N}_{A^*}$. *The operators* A *and* R *satisfy the further relations*

$$A = RR^*A = AR^*R\,, \quad RA^*A = AA^*R \tag{6.3a}$$

$$A^*R = R^*A\,, \quad AR^* = RA^* \tag{6.3b}$$

$$(A^*R)^2 = A^*A\,, \quad (AR^*)^2 = AA^*\,. \tag{6.3c}$$

It is clear from (6.2) that $\mathfrak{N}_R \subset \mathfrak{N}_A$, $\mathfrak{N}_{R^*} \subset \mathfrak{N}_{A^*}$ and $R^*AR^* = A^*$. Moreover

$$RR^*A = RR^*RA^*R = RA^*R = A = AR^*R$$
$$RA^*A = RA^*RA^*R = AA^*R\,.$$

Hence (6.3a) holds. The relation (6.3b) and (6.3c) follow from the computations

$$A^*R = R^*AR^*R = R^*A\,, \quad RA^* = RR^*AR^* = AR^*$$
$$(A^*R)^2 = A^*RR^*A = A^*A\,, \quad (AR^*)^2 = AR^*AR^* = AA^*\,.$$

In view of the corollary to Theorem 3.5 it is seen that $RA^{-1}R$ is the reciprocal of $A^* = R^*AR^*$, that is, $A^{*-1} = RA^{-1}R$. This proves the theorem.

It is easily seen from the formula (6.2) that $\mathfrak{N}_R = \mathfrak{N}_A$ if and only if $\mathfrak{N}_{R^*} = \mathfrak{N}_{A^*}$. In addition we have the following

COROLLARY. *An operator* A *is self-adjoint relative to an elementary operator and only if*

$$A = RR^*A = AR^*R\,, \quad R^*A = A^*R\,, \quad RA^* = AR^*\,. \tag{6.4}$$

The existence of elementary operator R relative to which A is self-adjoint is established in the following

THEOREM 6.2. *Given an operator* A *from* $\mathfrak{H}$ *to* $\mathfrak{H}'$ *there is a unique elementary operator* R *such that* A *is self-adjoint relative to* R, $\mathfrak{N}_A = \mathfrak{N}_R$ *and* A^*R *is nonnegative. The operators* $A^{-1}R$, AR^* *and* $A^{*-1}R^*$ *are also nonnegative and* $\mathfrak{N}_{A^*} = \mathfrak{N}_{R^*}$.

In order to prove this result P be the square root of A^*A. Then P is nonnegative and $\mathfrak{N}_P = \mathfrak{N}_A$. We shall show that the operator $R = (P^{-1}A^*)^*$ has the properties described theorem. Observe first that

$$R \supseteqq AP^{-1}\,, \quad R^* \supseteqq P^{-1}A^*\,, \tag{6.5}$$

and hence that

$$E = R^*R \supseteqq P^{-1}A^*AP^{-1} = P^{-1}P^2P^{-1} = (P^{-1}P)(PP^{-1})\,.$$

This is possible only in case E is a projection. Hence R is elementary and $\mathfrak{N}_R = \mathfrak{N}_P = \mathfrak{N}_A$, $\mathfrak{N}_{A^*} = \mathfrak{N}_{R^*}$. By Theorem 3.5, the operators R^*A and A^*R are closed. Moreover by (6.4).

$$E = R^*R \supseteqq R^*AP^{-1}\ , \quad E = R^*R \supseteqq P^{-1}A^*R\ .$$

It follows that $P^*R = R^*A \geqq 0$. Consequently

$$RP = RA^*R = RR^*A = A\ ,$$

the last inequality holding since $\mathfrak{N}_{A^*} = \mathfrak{N}_{R^*}$ and RR^* is a projection on $\mathfrak{H}'$. Since $A^{-1}R = P^{-1}$, $AR^* = RPR^*$, $A^{*-1}R^* = RA^{-1}RR^* = RP^{-1}R^*$, these products are nonnegative. The uniqueness of R follows from (6.3c) and the uniqueness of the square root of A^*A.

The elementary operator R described in Theorem 6.2 will be called the *elementry operator belonging to or associated with* A. The projections $E = R^*R$, $E' = RR^*$ are such that $E'A = AE = A$ and will be called the *projection associated with* A. It should be observed that if we set $P = A^*R$, $Q = AR^*$, then, $A = RA^*R = RP = QR$. This formula is commonly called the *polar decomposition* of A. It was first established for an unbounded operator in Hilbert space by J. von Neumann.[4]

COROLLARY. *If R is the elementary operator associated with A, then R is the elementary operator associated with A^{*-1} and R^* is the elementary operator associated with A^* and A^{-1}.*

THEOREM 6.3. *Let A be an operator from $\mathfrak{H}$ to $\mathfrak{H}$ and let R be the associated elementary operator. Then A is normal if and only if R^* commutes with A. If A is normal so also is R. The operator A is self-adjoint if and only if R is self-adjoint and commutes with A. Finally A is self-adjoint and nonnegative if and only if R is a projection.*

Since A^*A and AA^* are equal if and only if their square roots R^*A and AR^* are equal, it follows that A is normal if and only if A commutes with R^*. If $A = A^*$ then $R^*A = AR^* = RA$, by (6.3b). Hence $R = R^*$ and R commutes with A. Conversely if R commutes with A and $R = R^*$, then A is normal and $A^*R = RA = AR$. Hence $A^* = A$. If R is a projection, $AR = RA^*R^2 = RA^*R = A = RA$. Hence A is self-adjoint and nonnegative. The converse is immediate and the theorem is established.

COROLLARY 1. *If A is a self-adjoint operator from $\mathfrak{H}$ to $\mathfrak{H}$, it is expressible as the difference $A = A_+ - A_-$ of orthogonal nonnegative*

[4] Neumann, J. v. *Über Adjungierte Funktionaloperatoren* Annals of Math., **33** (1932), 294-310.

self-adjoint operators.

This follows because its associated elementary operator R is self-adjoint and hence is the difference $R = E_+ - E_-$ of two orthogonal projections. Since R and $E = E_+ + E_-$ commute with A so also does E_+ and E_-. Using this fact it is seen that $A_+ = AE_+$, $A_- = AE_-$ have properties described in the corollary.

COROLLARY 2. *If A is self-adjoint relative to an elementary operator T so also is its associated elementary operator R, that is $R = TR^*T$.*

7. **$*$-orthogonality and sections.** Two operators A and B will be said to be *$*$-orthogonal* if their carriers are orthogonal and their ranges are orthogonal. This is equivalent to the statemnt that $A^*B = 0$ (or $B^*A = 0$) on a dense set in $\mathfrak{H}$ and $AB^* = 0$ (or $BA^* = 0$) on a dense set in $\mathfrak{H}'$. It is clear that A is $*$-orthogonal to B if and only if A^{*-1} is $*$-orthogonal to B. If one of the pairs A, B; A^*, B^*; A^{-1}, B^{-1}; A^{*-1}, B^{*-1} form a $*$-orthogonal pair, then the remaining pairs form $*$-orthogonal pairs. Finally two operators A and B are $*$-orthogonal if and only if their associated elementary operators R and S are $*$-orthogonal. The following result is readily verified.

THEOREM 7.1. *Let B and C be $*$-orthogonal operators from $\mathfrak{H}$ to $\mathfrak{H}'$. Then $A = B + C$ is an operator and $A^{-1} = B^{-1} + C^{-1}$, $A^* = B^* + C^*$, $A^{*-1} = B^{*-1} + C^{*-1}$. Moreover A is elementary if and only if B and C are elementary. If S and T are respectively the elementary operators associated with B and C, then $R = S + T$ is the elementary operator associated with $A = B + C$.*

An operator B will be called a *section of an operator A*, if there is an operator C $*$-orthogonal to B such that $A = B + C$. If B is a section of A, its associated elementary operator S is a section of the associated elementary operator R of A. As a first result characterizing sections of A we have the following.

THEOREM 7.2. *Let $E = \overline{A^{-1}A}$, $E' = \overline{AA^{-1}}$ be the projections associated with A. Let F, F' be projections in $\mathfrak{H}$ and $\mathfrak{H}'$ respectively. Suppose that $AF \supseteqq F'A$. Then $EF = FE$ and $F'F' = F'E'$. Moreover AF is a section of A and its adjoint is A^*F'.*

Since the domain of $F'A$ is $\mathscr{D}_A$ it follows from the relation $AF \supseteqq F'A$ that $B = AF$ is dense. Since B is closed, it is an operator. Since $AE = A$ it follows that $AFE \supseteqq F'A$. Hence $AFE - AF = A(EFE -$

$EF) = 0$ on $\mathscr{D}_A$. This possible only in case $EFE = EF$ and hence only in case $EF = FE$. Similarly, since $A^*F' \supseteqq FA^*$, it follows that $F'E' = E'F'$. Moreover $B^* = A^*F'$. The projections associated with B are accordingly $G = EF$ and $G' = E'F'$. The operator $C = A(E - G)$ has $E - G$ and $E' - G'$ as its associated projections. It follows that C and B are $*$-orthogonal. Moreover $A = B + C$ and the theorem is established.

THEOREM 7.3. *An operator B is a section of A if and only if $A^*B = B^*B$ and $AB^* = BB^*$.*

If $A = B + C$, where B is *-orthogonal to C, then $A^*B = (B^* + C^*)B = B^*B$ and $AB^* = B^*B$. Conversely suppose that $A^*B = B^*B$ and $AB^* = BB^*$. Let $F = \overline{B^{-1}B}$, $F' = \overline{BB^{-1}}$. Then

$$B^* = A^*BB^{-1} \subseteqq A^*F' , \quad B = AB^*B^{*-1} \subseteqq AF .$$

It follows that $F'A \subseteqq B \subseteqq AF$ and hence that $B = AF$. In view of Theorem 7.2 the operator B is a section of A, as was to be proved.

THEOREM 7.4. *Let R be an elementary operator and let $E = R^*R$. Let F be a projection in $\mathfrak{H}$. Then $S = RF$ is a section of R if and only if $EF = FE$. Similarly if F' is projection to $\mathfrak{H}'$, then $F'R$ is a section of R if and only if $E'F' = F'E'$, where $E' = RR^*$.*

If $S = RF$ is a section of R, then

$$S^*S = R^*S = R^*RF = EF ,$$

is a projection in $\mathfrak{H}$. Hence $EF = FE$. Conversely if $EF = FE$ then

$$R^*S = R^*RF = EF = FEF = FR^*RF = S^*S ,$$
$$SS^* = RFR^* = RS^* .$$

Consequently, S is a section of R, by Theorem 7.3. The last statement in the theorem follows similarly.

8. $*$-commutativity. A bounded operator B from $\mathfrak{H}$ to $\mathfrak{H}'$ will be said to *$*$-commute* with an operator A from $\mathfrak{H}$ to $\mathfrak{H}'$ if

$$A^*B \supseteqq B^*A , \quad AB^* \supseteqq BA^* . \tag{8.1}$$

It should be observed that products A^*B and AB^* appearing in (8.1) are closed and dense and hence are operators. In the present section we shall derive some elementary properties of $*$-commutative operators of this type. Throughout this section the operator B is restricted to be bounded, while A is arbitary. The associated projections will be denoted by

$$E = \overline{A^{-1}A}\,, \quad E' = \overline{AA^{-1}}\,, \quad F = \overline{B^{-1}B}\,, \quad F' = \overline{BB^{-1}}\,, \tag{8.2}$$

as a first result we have

LEMMA 8.1. *Suppose that B $*$-commutes with A. The product A^*B is self-adjoint and is the closure of B^*A. Similarly, the product AB^* is self-adjoint and is the closure of BA^*.*

Suppose that B $*$-commutes with A and that A^*B is not the closure of B^*A. Then there is vector $x_0 \neq 0$ in the domain of A^*B such that

$$(x_0, x) + (A^*Bx_0, B^*Ax) = 0$$

for all x in $\mathscr{D}_A$. Since $(B^*A)^* = A^*B$ it follows that

$$(x_0, x) + (A^*BA^*Bx_0, x) = 0 \tag{8.3}$$

for all x in $\mathscr{D}_A$, and hence for all x in $\mathfrak{H}$. Choosing $x = B^*Bx_0$ and making use of (8.1) we find that

$$(x_0, B^*Bx_0) + (A^*AB^*Bx_0, B^*Bx_0) = 0.$$

Since B^*Bx_0 is in $\mathscr{D}_A$ we have

$$\| Bx_0 \|^2 + \| AB^*Bx_0 \|^2 = 0$$

and hence $Bx_0 = 0$. Using (8.3) we find that $x_0 = 0$, contrary to our choice of x_0. The closure of B^*A is accordingly A^*B. The last statement in the lemma follows by symmetry.

LEMMA 8.2. *Suppose that $\mathfrak{N}_A = \mathfrak{N}_B$ and $\mathfrak{N}_{A^*} = \mathfrak{N}_{B^*}$. If the first of the relations*

$$A^*B \supseteqq BA\,, \quad AB^* \supseteqq BA^*\,, \tag{8.4a}$$

$$A^{-1}B \supseteqq B^*A^{*-1}\,, \quad A^{*-1}B^* \supseteqq BA^{-1}\,, \tag{8.4b}$$

$$B^{-1}A^{*-1} \supseteqq A^{-1}B^{*-1}\,, \quad B^{*-1}A^{-1} \supseteqq A^{*-1}B^{-1}\,, \tag{8.4c}$$

$$B^{-1}A \supseteqq A^*B^{*-1}\,, \quad B^{*-1}A^* \supseteqq AB^{-1}\,, \tag{8.4d}$$

holds, so the others hold also. If (8.4) *holds, the products appearing on the right are operators.*

The last statement in the theorem follows from Theorem 3.5. Suppose now that (8.4a) holds. Then $A^*B = B^*A$ on $\mathscr{D}_A$. Consequently, on $\mathscr{D}_{A-1}$ we have

$$A^{*-1}A^*BA^{-1} = A^{*-1}B^*AA^{-1} = A^{*-1}B^*\,.$$

Hence the second relation in (8.4b) holds. The first relation follows similarly. The right and left members of (8.4c) are the reciprocals of the corresponding right and left numbers of (8.4a). Hence (8.4c) holds. Similarly (8.4d) holds.

LEMMA 8.3. *Suppose that B *-commutes A. Then A and B are expressible uniquely as sums of sections*

$$A = A_0 + A_1\,, \quad B = B_0 + B_1 \tag{8.5}$$

such that $(\alpha) A_0$ *is *-orthogonal to B and B_0 is *-orthogonal to A:* (β) *B_1 *-commutes with A_1 and $\mathfrak{N}_{A_1} = \mathfrak{N}_{B_1}$, $\mathfrak{N}_{A_1^*} = \mathfrak{N}_{B_1^*}$. Moreover*

$$A_1 = B^{*-1}A^*B\,, \quad A_1^* = B^{-1}AB^*\,. \tag{8.6}$$

Conversely, if A and B are expressible in the form (8.5) *such that* (α) *and* (β) *hold, then B *-commutes with A.*

Suppose first that B *-commutes with A. Using (8.1) and (8.2) it is seen that

$$EB^*A = B^*A\,, \quad E'BA^* = BA^*\,, \quad FA^*B = A^*B\,, \quad F'AB^* = AB^*\,.$$

Hence

$$EB^*E' = B^*E' : \quad E'BE = BE\,, \quad FA^*F' = A^*F' \quad F'AF = AF\,.$$

Consequently $BE = E'B$, $AF \supseteqq F'A$. In view of Theorem 7.2 it follows that $A_1 = AF$, $B_1 = BE$ are respectively sections of A and B, each having EF and $E'F'$ as their associated projections. We have accordingly $\mathfrak{N}_{A_1} = \mathfrak{N}_{B_1}$, $\mathfrak{N}_{A^*} = \mathfrak{N}_{B_1^*} = {}_B\mathfrak{N}_1^*$. Choose A_0 and B_0 so that (8.5) holds. The operator A_0 has $E - EF$ and $E' - E'F'$ as its associated projections and is accordingly *-orthogonal to B, B_0, B_1 and A_1. Similarly B_0 is *-orthogonal to A, A_0, A_1 and B_1. Using (8.1) again we see that

$$B_1^*A_1 = EB^*AF \subseteqq EA^*BF = (A_0^* + A_1^*)(B_0 + B_1) = A_1^*B_1\,.$$

Likewise $B_1A_1^* \subseteqq A_1B_1^*$. This proves the first conclusion of the lemma. The last statement is immediate.

It remains to obtain the formulas (8.6). To this end observe that

$$B^{*-1}A^*B = B_1^{*-1}A_1^*B_1 \supseteqq B_1^{*-1}B_1^*A_1 = A_1\,.$$

In view of the result we may suppose that $A = A_1$. Assume that $A \neq B^{*-1}A^*B$. Since $B^{*-1}AB^*$ and A are closed, there is a vector $x_0 \neq 0$ such that

$$(x_0, x) + (B^{*-1}A^*Bx_0, Ax) = 0 \tag{8.7}$$

for all x in $\mathscr{D}_A$. Consequently, by (8.4d),

$$(x_0, x) = -(A^*B^{*-1}A^*Bx_0\, x) = -(B^{-1}AA^*Bx_0, x)$$

for all x in $\mathscr{D}_A$ and hence for x in $\mathfrak{H}$. Choosing $x = B^*Bx_0$ we find that

$$\| Bx_0 \|^2 = -(AA^*Bx_0\, Bx_0) = -\| A^*Bx_0 \|^2 .$$

This relation together with (8.7) can hold only in case $x_0 = 0$. It follows that the first formula in (8.6) holds. The second is obtained by symmetry and the lemma is established.

COROLLARY 1. *Suppose that B $*$-commutes with A. The associated projections* (8.2) *satisfy the relation $EF = FE$, $E'F' = F'E'$. Moreover $\mathfrak{N}_A = \mathfrak{N}_B$ if and only if $\mathfrak{N}_A^* = \mathfrak{N}_B^*$.*

As a further result we have

COROLLARY 2. *If an elementary operator T $*$-commutes with A, then TA^*T is a section of A.*

In view of Lemma 8.2 and 8.3 we have

COROLLARY 3. *Suppose that B $*$-commutes with A. Then B $*$-commutes with A^{*-1} and with $\alpha A + \beta A^{*-1}$, where α and β are positive numbers.*

The restriction that α, β are positive is made only to insure that $\alpha A + \beta A^{*-1}$ be closed.

LEMMA 8.4. *Let T be an elementary operator such that $TA^*T = A$ and suppose that B $*$-commutes with T. Then B $*$-commutes with A if and only if $AT^*B \supseteqq BT^*A$.*

If B $*$-commutes with A then

$$AT^*B = TA^*TT^*B = TA^*B \supseteqq TB^*A = BT^*A .$$

Conversely, if $AT^*B \supseteqq BT^*A$, then

$$A^*B = T^*AT^*B \supseteqq T^*BT^*A = B^*TT^*A = B^*A$$
$$AB^* = TA^*TB^* \supseteqq TB^*TA^* = BT^*TA^* = BA^*$$

as was to be proved.

9. Decomposition of an operator.

As a first result we have

THEOREM 9.1. *Let R be the elementary operator associated with* *an operator A from $\mathfrak{H}$ to $\mathfrak{H}'$. Let T be second elementary operator that $*$-commutes with A. Then T $*$-commutes with R and the operators A, R, T are expressible uniquely as sums and difference*

$$(9.1)\quad A = A_0 + A_+ + A_- , \quad R = R_0 + R_+ + R_- , \quad T = T_0 + R_+ - R_-$$

of mutually $*$-orthogonal operators such that R_0, R_+, R_- are the elementary operators associated respectively with A_0, A_+, A_- and T_0 is $*$-orthogonal to A. Moreover T is $*$-orthogonal to A_0 and R_0 and $*$-commutes with A_+, A_-, R_+ and R_-. Conversely if A, R, T are expressible in the form (9.1) *then T $*$-commutes with A and R.*

Suppose that T $*$-commutes with A. Then, by Lemma 8:3, they are expressible in the forms $A = A_0 + A_1$, $T = T_0 + T_1$, where A_0 is $*$-orthogonal to T and A_1, T_0 is $*$-orthogonal to A and T, $\mathfrak{N}_{T_1} = \mathfrak{N}_{A_1}$, $\mathfrak{N}_{T_1^*} = \mathfrak{N}_{A_1^*}$ and T_1 $*$-commutes with A_1 Moreover, by Theorem 7.1 $R = R_0 + R_1$, where R_0 is the elementary operator belonging to A_0 and R_1 is the elementary operator belonging to A_1. In view of this result we can restrict ourselves to the case in with $A_0 = 0$, $T_0 = 0$, $R_0 = 0$. Then $\mathfrak{N}_A = \mathfrak{N}_T = \mathfrak{N}_R$ and $\mathfrak{N}_{A^*} = \mathfrak{N}_{T^*} = \mathfrak{N}_{R^*}$. Since A^*T is self-adjoint, its associated elementary operator S is self-adjoint and hence is expressible as the difference $S = E_+ - E_-$ of two orthogonal projections E_+, E_- whose sum is $E = R^*R$. The operator A^*TS is nonnegative and self-adjoint. It follows from Theorem 6.1 that $\mathrm{R} = TS$ and $T = RS$. Setting $R_+ = RE_+$, $R_- = RE_-$ we see that

$$R = RE = R_+ + R_- , \quad \mathrm{T} = RS = R_+ - R_- .$$

Since $AR^* = AR_+^* + AR_-^*$ and $AT^* = AR_+^* - AR_-$ are self-adjoint, so also are AR_+^* and AR_-^*. Moreover $AR_+^* \geqq 0$ and $AR_-^* \geqq 0$ since they are orthogonal and $AR^* \geqq 0$. The elementary operators R_+ and R_- are therefore the elementary operators associated respectively with $A_+ = AE_+$ and $A_- = AE_-$. Since R_+ and R_- are $*$-orthogonal it follows that A_+ and A_- are $*$-orthogonal. Consequently A, R, T are expressible in the form (9.1). The remaining statements in the theorem are easily established.

COROLLARY. *Two elementary operators R and T on $\mathfrak{H}$ to $\mathfrak{H}'$ $*$-commute if and only if there exist mutually $*$-orthogonal elementary operators R_0, R_+, R_-, T_0 such that $R = R_0 + R_+ + R_-$, $T = T_0 + R_+ - R_-$. Moreover, this representation is unique.*

THEOREM 9.2. *Let R be the elementary operator associated with an operator A from $\mathfrak{H}$ to $\mathfrak{H}'$. Given a positive number λ there are unique decompositions.*

(9.2) $$A = A_{\lambda+} + A_{\lambda 0} + A_{\lambda-} , \quad R = R_{\lambda+} + R_{\lambda 0} + R_{\lambda-}$$

of A and R into sections such that $R_{\lambda+}$, $R_{\lambda 0}$, $R_{\lambda-}$, $R_{\lambda+} - R_{\lambda-}$ are respectively the elementary operators associated with $A_{\lambda+}$, $A_{\lambda 0}$, $A_{\lambda-}$, $A - \lambda R$. Moreover,

(9.3a) $$A_{\lambda+} = R_{\lambda+}A^*R_{\lambda+} , \quad A_{\lambda 0} = \lambda R_{\lambda 0} , \quad A_{\lambda-} = R_{\lambda-}A^*R_{\lambda-} .$$

The relations

(9.3b) $$\| A_{\lambda+}x \| > \lambda \| R_{\lambda+}x \| , \quad (A_{\lambda+}x, R_{\lambda+}x) > \lambda \| R_{\lambda+}x \|^2 ,$$

hold for all x in $\mathscr{D}_A$ such that $R_{\lambda+}x \neq 0$ and the relations

(9.3c) $$\| A_{\lambda-}x \| < \lambda \| R_{\lambda-}x \| , \quad (A_{\lambda-}x, R_{\lambda-}x) < \lambda \| R_{\lambda-}x \|^2$$

hold for all x in $\mathfrak{H}$ such that $R_{\lambda-}x \neq 0$. If $\lambda < \mu$, then $A_{\lambda-} + A_{\lambda 0}$ is a section of $A_{\mu-}$, and $A_{\mu 0} + A_{\mu+}$ is a section of $A_{\lambda+}$. Similarly $R_{\lambda-} + R_{\lambda 0}$ is a section of $R_{\mu-}$ and $R_{\mu 0} + R_{\mu+}$ is a section of $R_{\lambda+}$.

In order to prove this result let $C = A - \lambda R$, where λ is a fixed positive number. Let T be the elementary operator associated with C. Since R $*$-commutes with A and R, it follows that R $*$-commutes with C. By virtue of Theorem 9:1 R also $*$-commutes with T. Similarly T $*$-commutes R and C and hence also with $A = C + \lambda R$. Applying Theorem 9.1 to A, R, T and to C, T, R it is seen that they aree xpressible uniquely as sums

$$A = A_0 + A_+ + A_- , \quad R = R_0 + R_+ + R_-$$
$$C = C_0 + C_+ - C_- , \quad T = T_0 + R_+ - R_-$$

of mutually $*$-orthogonal operators such that R_+ is the elementary operator associated with A_+ and C_+; R_- is the elementary operator associated with A_- and C_-; R_0 is the elementary operator associated with A_0. Since $\mathfrak{N}_C \supset \mathfrak{N}_A$ it follows that $C_0 = T_0 = 0$. From the relation $C = A - \lambda R$ we obtain the relations

$$A_0 = \lambda R_0 , \quad A_+ = C_+ + \lambda R_+ , \quad A_- = -C_- + \lambda R_- .$$

Moreover, if we set $E_+ = R_+^*R_+$, $E_- = R_-^*R_-$

$$R_+^* A_+ = R_+^*C_+ + \lambda E_+ , \quad R_-^*A_- = -R_-^*C_- + \lambda E_- .$$

It follows that the second relations in (9.3b) and (9.3c) hold. If x is in $\mathscr{D}_A$, then

$$\| A_+x \|^2 - \lambda^2 \| R_+x \|^2 = \| C_+x \|^2 + 2\lambda(R_+^* C_+x, x) \geqq 0 .$$

Hence the first relation in (9.3c) holds. Since $P = R_-^*A_- \geqq 0$ and $Q = R_-^*C_- \geqq 0$ satisfy the relation $P + Q = \lambda E_-$, they are bounded and commute. Hence $PQ = A_-^*C_- = C_-^*A_- \geqq 0$. Using the relations

$$\| A_-x \|^2 + 2(A_-^*C_-x, x) + \| C_-x \|^2 = \lambda^2 \| R_-x \|^2$$

it is seen that the first relation in (9.3c) holds.

In order to prove the last statement with $\mu > \lambda$ apply the results described in the first part with $A_{\lambda+}$, $R_{\lambda+}$, μ playing role of A, R, λ. One then obtains the partitions

$$A_{\lambda+} = A_{\mu+} + A_{\mu 0} + A_{\lambda\mu} , \quad R_{\lambda+} - R_{\mu+} + R_{\mu 0} + R_{\lambda\mu} . \tag{9.4a}$$

Setting

$$A_{\mu-} = A_{\lambda\mu} + A_{\lambda 0} + A_{\lambda-} , \quad R_{\mu-} = R_{\lambda\mu} + R_{\lambda 0} + R_{\lambda-} . \tag{9.4b}$$

We have

$$A = A_{\mu+} + A_{\mu 0} + A_{\mu-} , \quad R = R_{\mu+} + R_{\mu 0} + R_{\mu-}$$

with $R_{\mu+} - R_{\mu-}$ as the elementary operator $A - \mu R$. The last statement of theorem follows from the relations (9.4). This completes the proof of Theorem 9.2.

COROLLARY 1. *Suppose that A is bounded and set $M = \| A \|$. Let m be the largest number such that $\| Ax \| \geqq m \| Rx \|$. If $\lambda \geqq M$, then $R_{\lambda+} = 0$. If $m > 0$ and $0 < \lambda \leqq m$, then $R_{\lambda-} = 0$. If $m < \lambda < M$, then $\| A - \lambda R \| \leqq \max\,[M - \lambda, \lambda - m]$.*

COROLLARY 2. *The operator $R_\lambda = R_{\lambda 0} + R_{\lambda-}$ $(0 < \lambda < \infty)$ is the elementary operator belonging to $A_\lambda = A_{\lambda 0} + A_{\lambda-} = R_\lambda A^* R_\lambda$. Moreover*

(1) $\lim_{\lambda=\infty} R_\lambda = R , \quad \lim_{\lambda=0} R_\lambda = 0 , \quad \lim_{\lambda=\infty} A_\lambda = A , \quad \lim_{\lambda=0} A_\lambda = 0 .$

(2) *If $\lambda < \mu$, then R_λ is a section of R_μ, A_λ is a section of A_μ, and*

$$\lambda \| R_\lambda x \| \leqq \| A_\mu x - A_\lambda x \| \leqq \mu \| R_\mu x \| .$$

(3) $\lim_{\mu=\lambda+0} R_\mu = R_\lambda , \quad \lim_{\mu=\lambda+0} A_\mu = A_\lambda .$

Let $A_\lambda (0 < \lambda < \infty)$ be the one parameter family of sections of A described in the last corollary. By the *principal spectrum* Λ of A will be meant the set of all numbers λ_0 on $0 \leqq \lambda < \infty$ such that A_λ is constant on no neighborhood of λ_0. The principal spectrum of A^* is also Λ. The spectrum of A^{-1} and A^{*-1} is the closure of the reciprocals $1/\lambda$ of the points $\lambda \neq 0$ in Λ.

If R_λ is the elementary operator of A_λ described in the last corollary, we have the representations

$$A = \int_0^\infty \lambda dR_\lambda \, , \quad A^* = \int_0^\infty \lambda dR_\lambda^* \, ,$$
$$A^{*-1} = \int_0^\infty \lambda^{-1} dR_\lambda \, , \quad A^{-1} = \int_0^\infty \lambda^{-1} dR_\lambda^* \, ,$$

where the integrals are defined in the usual manner. It should be observed that $E_\lambda = R_\lambda^* R_\lambda$ and $E'_\lambda = R_\lambda R_\lambda^*$ are resolutions of $E = R^*R$ and $E' = RR^*$, respectively. Since $R_\lambda = RE_\lambda = E'_\lambda R$ we have, from the polar form of A,

$$A = R\int_0^\infty \lambda dE_\lambda = \left(\int_0^\infty \lambda dE'_\lambda\right) R \, .$$

It follows that the results given above can be derived from the self-adjoint case, if one so desires.

An extension of the results given above is found in the following

THEOREM 9.3. *Let A be an operator and T be an elementary operator such that $A = TA^*T$. Given a real number λ there exist a unique decomposition*

$$A = A_{\lambda+} + A_{\lambda 0} + A_{\lambda-} \, , \quad T = T_{\lambda+} + T_{\lambda 0} + T_{\lambda-} \quad (-\infty < \lambda < \infty) \tag{9.6}$$

of A and T into sections such that

$$A_{\lambda+} = T_{\lambda+} A^* T_{\lambda+} \, , \quad A_{\lambda 0} = \lambda T_{\lambda 0} \, , \quad A_{\lambda-} = T_{\lambda-} A^* T_{\lambda-} \tag{9.7a}$$

$$(A_{\lambda+}x, T_\lambda x) > \lambda \, \| T_{\lambda+}x \|^2 \text{ for all } x \text{ in } \mathscr{D}_A \text{ having } T_{\lambda+}x \neq 0 \, , \tag{9.7b}$$

$$(A_{\lambda-}x, T_\lambda x) < \lambda \, \| T_{\lambda-}x \|^2 \text{ for all } x \text{ in } \mathscr{D}_A \text{ having } T_{\lambda-}x \neq 0 \, . \tag{9.7c}$$

If $\mu > \lambda$, then $A_{\lambda 0} + A_{\nu-}$ is a section of $A_{\mu-}$, $T_{\lambda 0} + T_{\lambda-}$ is a section of $T_{\mu-}$, $A_{\mu+} + A_{\mu 0}$ is a section of $A_{\lambda+}$ and $T_\mu + T_{\mu 0}$ is a section of $T_{\lambda+}$.

In order to prove this result observe first that by Theorem 9.1 the operators A, R, T have unique decompositions

$$A = A_1 + A_2 \, , \quad R = R_1 + R_2 \, , \quad T = T_0 + R_1 - R_2$$

where R_1, R_2 are the elementary operators associated with A_1, A_2 respectively and T_0 is $*$-orthogonal to A. The terms A_0 and R_0 described in Theorem 9.1 are zero since $A = TA^*T$. If λ is positive let

$$A_1 = A_{1\lambda+} + A_{1\lambda 0} + A_{1\lambda-} \, , \quad R_1 = R_{1\lambda+} + R_{1\lambda 0} + R_{1\lambda-} \, ,$$

be the decompositions of A_1 and R_1 described in Theorem 9.2. Then

$$A_{\lambda+} = A_{1\lambda+} \, , \quad A_{\lambda 0} = A_{1\lambda 0} \, , \quad A_{\lambda-} = A_{1\lambda-} + A_2$$
$$T_{\lambda+} = R_{1\lambda+} \, , \quad T_{\lambda 0} = R_{1\lambda 0} \, , \quad T_{\lambda-} = R_{1\lambda-} - R_2 + T_0$$

have the properties described in Theorem 9.3. If $\lambda = 0$ set

$$A_{\lambda+} = A_1 , \quad A_{\lambda 0} = 0 , \quad A_{\lambda-} = A_2$$
$$T_{\lambda+} = R_1 , \quad T_{\lambda 0} = T_0 , \quad T_{\lambda-} = -R_2 .$$

If $\lambda = -\mu < 0$ let

$$A_2 = A_{2\mu+} + A_{2\mu 0} + A_{2\mu-} , \quad R_2 = R_{2\mu+} + R_{2\mu 0} + R_{2\mu-}$$

be the decomposition of A_2, R_2 described in Theorem 9.2. Then

$$A_{\lambda+} = A_1 + A_{2\mu-} , \quad A_{\lambda 0} = A_{2\mu 0} , \quad A_{\lambda-} = A_{2\mu+} \qquad (\mu = -\lambda)$$
$$T_{\lambda+} = R_1 - R_{2\mu-} , \quad T_{\lambda 0} = -R_{2\mu 0} , \quad T_{\lambda-} = -R_{2\mu+}$$

have the properties described in Theorem 9.3. The uniqueness of the decomposition follows from (9.7) and the connections between T and R.

COROLLARY. *The operators* $T_\lambda = T_{\lambda 0} + T_{\lambda-}$, $A_\lambda = A_{\lambda 0} + A_{\lambda-} = T_\lambda = A^* T_\lambda$ *have the following properties*:

(1) $\lim_{\lambda=+\infty} T_\lambda = T , \quad \lim_{\lambda=-\infty} T_\lambda = 0 , \quad \lim_{\lambda=+\infty} A_\lambda = A , \quad \lim_{\lambda=-\infty} A_\lambda = 0 .$

(2) *If* $\lambda < \mu$, T_λ *is a section of* T_μ, A_λ *is a section of* A_μ

(3) $\lim_{\mu=\lambda+0} T_\mu = T_\lambda , \quad \lim_{\mu=\lambda+0} A_\mu = A_\lambda .$

(4) $(T_\lambda x, A_\lambda x) \leqq \lambda \| T_\lambda x \|^2$ *for all* x *in* $\mathscr{D}_A$.

In view of the results obtained in the last corollary we shall define the *spectrum* Λ *of* A *relative to* T to be the set of all real numbers λ_0 such that the operators A_λ described in the last corollary is constant on no neighborhood of λ_0. The spectrum of A^* relative to T^* is also Λ. Similarly the spectrum of A^{*-1} relative to T and A^{-1} relative to T^* is the closure of the reciprocal $1/\lambda$ of the points $\lambda \neq 0$ in Λ. Moreover A and A^* are representable

$$A = \int_{-\infty}^{\infty} \lambda d T_\lambda , \quad A^* = \int_{-\infty}^{\infty} \lambda d T_\lambda^* .$$

If $\mathfrak{N}_A = \mathfrak{N}_T$, then

$$A^{*-1} = \int_{-\infty}^{\infty} \lambda^{-1} d T_\lambda , \quad A^{-1} = \int_{-\infty}^{\infty} \lambda^{-1} d T_\lambda .$$

When $\mathfrak{H} = \mathfrak{H}'$ and T is the identity one obtains the usual spectral resolution for self-adjoint operators.

10. Spectrum of the gradient operator. Let $\mathfrak{H}$ be the class of all complex valued Lebesgue square integrable functions $x(t) = x(t_1, \cdots, t_m)$

of points $t = (t_1, \cdots, t_m)$ in an m-dimensional Euclidean space. It is convenient to normalize a function in $\mathfrak{H}$ to be equal to the limit of its integral mean whenever these limits exist and setting $x(t) = 0$ elsewhere. The class so normalized forms a Hilbert space over the field of complex numbers with

$$(x_1, x_2) = \int_{-\infty}^{\infty} x_1(t)\overline{x_2(t)}dt$$

as the inner product, where $\bar{x}(t)$ denotes the conjugate of $x(t)$. As is well known the Fourier transform

$$\hat{x}(s) = c\int_{-\infty}^{\infty} e^{-ist}x(t)dt \ , \quad st = s_1t_1 + \cdots + s_mt_m \ , \tag{10.1}$$

where $c = (2/\pi)^{m/2}$, defines an isometry on $\mathfrak{H}$ onto $\mathfrak{H}$ and hence is an elementary operator, whose inverse is given by

$$x(t) = c\int_{-\infty}^{\infty} e^{ist}\hat{x}(s)ds \ . \tag{10.2}$$

Let $\mathscr{D}$ be the class of all functions x in $\mathfrak{H}$ that are linearly absolutely continuous[5] and whose partial derivatives are in $\mathfrak{H}$. A function x in $\mathscr{D}$ is characterized by the condition that $s_1\hat{x}(s), \cdots, s_m\hat{x}(s)$ are square integrable, where $\hat{x}(s)$ is the Fourier transform of x. In fact one has

$$-i\frac{\partial x(t)}{\partial t_\alpha} = c\int_{-\infty}^{\infty} e^{ist}s_\alpha x(s)ds \ .$$

The gradient operator A defined by $-i(\partial/\partial t_1), \cdots, -i(\partial/\partial t_m)$ is a closed operator from $\mathfrak{H}$ to the cartesian product $\mathfrak{H}'$ of $\mathfrak{H}$ by itself m times. The domain A is $\mathscr{D}$. It is not difficult to see that A is the closure of the restriction of A to the class of functions of class C^∞ with compact support.

Let $y(t) = [y_1(t), \cdots, y_m(t)]$ be a function in $\mathfrak{H}'$. If $y(t)$ is of class C^∞ and has compact support, then the divergence

$$\operatorname{div} y = i\frac{\partial y_1}{\partial t_1} + \cdots + i\frac{\partial y_m}{\partial t_m}$$

is in $\mathfrak{H}$. This operator from $\mathfrak{H}'$ to $\mathfrak{H}$ is preclosed and its closure is the adjoint A^* of A. If $\hat{y}(s)$ is the Fourier transform of $y(t)$ then y is in $\mathscr{D}_{A^*}$ if and only if the sum $s_\alpha\hat{y}_\alpha(s)$ is square integrable. Moreover

$$A^*y = c\int_{-\infty}^{\infty} e^{ist}s_\alpha\hat{y}_\alpha(s)ds \ , \quad (\alpha \text{ summed}) \ .$$

[5] Calkin J. W. "Functions of several variables and absolute continuity I," *Duke Math J.* Vol 6 (1940) pp. 170-186. See also Morrey Jr. C. B. "Functions of several variables and absolute continuity II, Duke Math J. **6** (1940), 187-215.

The elementary operator R associated with A is given by the formulas.

$$(Rx)_\alpha = c\int_{-\infty}^{\infty} e^{ist}\frac{s_\alpha \hat{x}(s)}{|\, s\,|}ds \qquad (\alpha = 1, \cdots m),$$

$$R^* y = c\int_{-\infty}^{\infty} e^{ist}\frac{s_\alpha \hat{y}_\alpha(s)}{|\, s\,|}ds$$

where $|\, s\,|$ is the distance from s to the origin. The carrier of R^* is the set of all functions y in $\mathfrak{H}'$ whose Fourier transforms $\hat{y}_\alpha(s)$ are of the form $s_\alpha \hat{x}(s)/|\, s\,|$ such that $\hat{x}(s)$ is in $\mathfrak{H}$. Similarly the carrier of A^* consists of all functions y in $\mathfrak{H}'$ whose Fourier transform is of the form $\hat{y}_\alpha(s) = s_\alpha \hat{x}(s)$., such that $|\, s\,|^2 \hat{x}(s)$ is in $\mathfrak{H}$. It is easily seen that

$$(A^{*-1}x)_\alpha = c\int_{-\infty}^{\infty} e^{ist}\frac{s_\alpha \hat{x}(s)}{|\, s\,|^2}ds$$

$$A^{-1}y = c\int_{-\infty}^{\infty} e^{ist}\frac{s_\alpha \hat{y}_\alpha(s)}{|\, s\,|^2}ds\ .$$

The operator A^*A is, of course, the Laplacian.

The operators A_λ, R_λ described in Corollary 2 to Theorem 9.2 are defined by the formulas

$$(A_\lambda x)_\alpha = c\int_{-\infty}^{\infty} e^{ist}\varphi_\lambda(s) s_\alpha \hat{x}(s) ds$$

$$(R_\lambda x)_\alpha = c\int_{-\infty}^{\infty} e^{ist}\varphi_\lambda(s)\frac{s_\alpha \hat{x}(s)}{|\, s\,|}ds$$

where $\varphi_\lambda(s)$ is the characteristic function of the sphere $|\, s\,| \leqq \lambda$. The principal spectrum of A is accordingly point set $0 \leqq \lambda < \infty$.

11. Principal values and principal vectors. In the present section we shall be interested in certain special points of the principal spectrum of A which we shall call principal values of A. Before defining this concept it will be convenient to introduce the concept of the rank of an operator. By the *rank* of an operator A will be meant the dimension of its carrier, or equivalently the dimension of its range. It is clear that the ranks of A, A^*, A^{-1}, A^{*-1}, A^*A, AA^* are the same. If the rank of A is finite, then A is bounded and reciprocally bounded.

A number λ on $0 < \lambda < \infty$ will be said to be a *principal value* of an operator A if the rank of the section $A_{\lambda 0} = \lambda R_{\lambda 0}$ of A described in Theorem 9.2 is not zero. The rank of $A_{\lambda 0}$ will be called the *order of* λ as a principal value of A and $A_{\lambda 0}$ will be called the corrersponding *principal section* of A. The non-null vectors in the carrier of $A_{\lambda 0}$ will be called the *principal vectors* of A corresponding to λ. The non-null vectors in range of A_{λ_0} will be called the *principal reciprocal vectors* of A corresponding to λ. The latter are the principal vectors of A^{-1} corresponding to $1/\lambda$. The order of $1/\lambda$ as principal value of A^{-1} is equal to the order of λ as a principal value of A. A number λ is a principal value of A if and only if it is a principal value of A^* and its order as a principal value of A is equal to its order as a principal value of A^*. A positive number λ is a principal value of A if and only if λ^2 is a principal value of A^*A. Again the order of corresponding principal values are the same. The principal values of A^*A are the nonzero eigenvalues of A^*A. The eigenvectors A^*A corresponding to nonzero eigenvalues are the principal vectors of A. Similarly the eigenvectors of AA^* corresponding to nonzero eigenvalues are the principal reciprocal vectors of A. Principal values of A belong to the principal spectrum of A. Isolated points of the principal spectrum of A are principal values of A.

A principal value λ of A can be characterized in another way. A value λ is a principal value of A if and only if there is a non-null vector x in its carrier such that $Ax = \lambda Rx$, where R is its associated elementary operator of A. The vector $y = Rx$ is a principal reciprocal vector of A and satisfies the relation $A^*y = \lambda R^*y$. Consequently,

$$Ax = \lambda y\ , \quad A^*y = \lambda x\ . \tag{11.1}$$

Conversely if λ is a positive number such that there exist a vector $x \neq 0$ on $\mathscr{D}_A$ and a vector $y \neq 0$ in $\mathscr{D}_A$ such that (11.1) holds, then λ is a principal value of A, x is a principal vector and y is a principal reciprocal vector. From these remarks, it follows that the principal values of A are the positive eigenvalues of the self-adjoint operator

$$\begin{pmatrix} 0 & A^* \\ A & 0 \end{pmatrix}$$

from the cartesian product $\mathfrak{H} \times \mathfrak{H}'$ to $\mathfrak{H} \times \mathfrak{H}'$. It is clear that the foregoing results could have been obtained from the study of this self-adjoint operator. However, the author prefers the more direct approach here given.

THEOREM 11.1. *Suppose the principal spectrum of A apart from $\lambda = 0$ consists of a set of isolated points $\lambda_1, \lambda_2, \cdots$. Then λ_i is a principal value of A and has associated with it a unique elementary operator R_i as described in Theorem 9.2. The elementary operators $R_1, R_2, \cdots$, are mutually $*$-orthogonal and*

$$A = \Sigma\lambda_i R_i\ , \quad A^* = \Sigma\lambda_i R_i^*\ , \quad A^{*-1} = \Sigma\frac{1}{\lambda_i}R_i\ , \quad A^{-1} = \Sigma\frac{1}{\lambda_i}R_i^*\ .$$

12. Further results on $*$-commutativity. Throught this section we shall be concerned with a closed operator A and bounded operator B from $\mathfrak{H}$ to $\mathfrak{H}'$. As a first results we have the following converse of a statement in Lemma 8:4.

LEMMA 12.1. *If B $*$-commutes with $\alpha A + \beta A^{*-1}$ for every pair of positive numbers α and β, then B $*$-commutes with A.*

Suppose that B $*$-commutes with $C = \alpha_1 A + \beta_1 A^{*-1}$ and $D = \alpha_2 A + \beta_2 A^{*-1}$ where α_1, β_1, α_2, β_2 are positive numbers such that $\alpha_1\beta_2 - \beta_2\alpha_1 = 1$. These operators have the common domain $\mathscr{D} = \mathscr{D}_A \cap \mathscr{D}_{A^{*-1}}$. The operator $\beta_2 C - \alpha_1 D$ is the restriction of A to $\mathscr{D}$. Since $C^*B = B^*C$ and $D^*B = B^*D$ on $\mathscr{D}$ it follows that $A^*B = B^*A$ on $\mathscr{D}$. In order to show that $A^*B = B^*A$ on $\mathscr{D}_A$ consider a vector x in $\mathscr{D}_A$. Let $x_n = E_n x - E_{1/n}x$, where $E_\lambda = R_\lambda^* R_\lambda$ and R_λ is the section of R described in Corollary 2 to Theorem 9.2. The vector x_n is in $\mathscr{D}$ and $x_n \Rightarrow x$, $Ax_n \Rightarrow Ax$. Consequently $A^*Bx_n = B^*Ax_n \Rightarrow B^*Ax$. Since A^*B is closed we have $A^*Bx = B^*Ax$. We have accordingly $A^*B \supseteqq B^*A$. Similarly $AB^* \supseteqq B^*A$ and the lemma is proved.

COROLLARY. *If $\mathfrak{H}' = \mathfrak{H}$ and $\alpha A + \beta A^{*-1}$ is self-adjoint for all pairs of positive numbers α and β, then A is also self-adjoint.*

This result is obtained from lemma by selecting $B = I$, the identity.

LEMMA 12.2. *Suppose that $\mathfrak{H}' = \mathfrak{H}$ and that B $*$-commutes with A. If one of A or B is self-adjoint and positive, the other is self-adjoint.*

Consider first the case in which A is bounded and $A = A^* > 0$. Since $AB = B^*A$ and $AB^* = BA$, the difference $C = B - B^*$ satisfies the relation $AC = -CA$. Hence $A^nC = (-1)^n CA^n$ and $e^{tA}C = Ce^{-tA}$. The inverse e^{tA} is e^{-tA}. We have accordingly

$$C = e^{-tA}Ce^{-tA} .$$

Since $A > 0$ it follows from the spectral theorem for A that $\lim_{t=+\infty} e^{-tA}x = 0$ for each x in $\mathfrak{H}$. Consequently $C = 0$, that is, $B = B^*$ for the case here considered.

If A is reciprocally bounded, then A^{*-1} is bounded and B $*$-commutes with A^{*-1}. If A is self-adjoint and positive, so also is $A^{*-1} = A^{-1}$ and B $*$-commutes with A^{-1}. Hence $B = B^*$ by virture of the result just obtained. If $B = B^* > 0$ then A^{*-1} is self-adjoint and hence A is also self-adjoint.

In the general case if $A = A^* > 0$, then $A + A^{-1}$ is self-adjoint, reciprocally bounded and positive. Since B $*$-commutes with $A + A^{-1}$, it follows that $B = B^*$. If on the other hand $B = B^* > 0$, then $C = \alpha A + \beta A^{*-1}$ is self-adjoint whenever α and β are positive. This follows because C is reciprocally bounded and $*$-commutes with B. By virtue of the last corollary the operator A is self-adjoint and the lemma is established.

LEMMA 12.3. *Let R and S be the elementary operators associated with A and B respectively. If B $*$-commutes with A, then B $*$-commutes with R, and S $*$-commutes with A.*

In order to prove this result we may suppose, by Lemma 8.3 that $\mathfrak{N}_A = \mathfrak{N}_B$ and $\mathfrak{N}_{A^*} = \mathfrak{N}_{B^*}$. In fact, we may assume that $\mathfrak{N}_A = \mathfrak{N}_B = 0$, $\mathfrak{N}_{A^*} = \mathfrak{N}_{B^*} = 0$. Under these assumptions $P = A^*R = R^*A$ is positive and self-adjoint. Setting $Q = R^*B$ we obtain

$$PQ = A^*RR^*B = A^*B \supseteqq B^*A = B^*RR^*A = Q^*P$$
$$P^*Q = R^*AB^*R \supseteqq R^*BA^*R = QP\,.$$

Hence Q $*$-commutes with P and $Q = R^*B = B^*R$ by the last lemma. Similarly $RB^* = BR^*$. Consequently B $*$-commutes with R.

In order to prove that S $*$-commutes with A it is sufficient, by Lemma 12.1, to show that S $*$-commutes with $C = \alpha A + \beta A^{*-1}$, where α and β are positive numbers. The operator C is reciprocally bounded and $*$-commutes with B. The operator C^{*-1} is bounded and $*$-commutes with B. Hence S $*$-commutes C^{*-1} and also with C. This completes the proof of the lemma.

LEMMA 12.4. *If an elementary operator T $*$-commutes with A, then T $*$-commutes with a section A_1 of A if and only if it $*$-commutes with the elementary operator R_1 associated with A_1.*

Let A_1 be a section of A and let A_0 be the section of A such that $A = A_0 + A_1$. Let R_0, R_1 be the elementary operator associcated with A_0, A_1 respectively. Suppose that T $*$-commutes with R_1. Since T $*$-commutes with $R = R_0 + R_1$, it follows that T $*$-commutes with R_0. Consequently $T_0 = R_0T^*R_0$ and $T_1 = R_1T^*R_1$ are $*$-orthogonal sections of T. The section $T_2 = T - T_0 - T_1$ is $*$-orthogonal to A. Consequently the operators A^*T and AT^* are expressible as sums

$$A^*T = A_0^*T_0 + A_1^*T_1\,, \quad AT^* = A_0T_0^* + A_1T_1^*$$

of orthogonal operators. Hence $A_1^*T_1$ and $A_1T_1^*$ are self-adjoint and T_1 $*$-commutes with A_1. Since $A_1^*T = A_1^*T_1$ and $A_1T^* = A_1T_1^*$, it follows that T $*$-commutes with A_1 as was to be proved.

13. Representations of operators as products. The present section will be devoted to an extension of Lemma 5.2 and some of its consequences.

THEOREM 13.1. *Let A be an operator from $\mathfrak{H}$ to $\mathfrak{H}'$ and let R be its associated elementary operator. There is a unique pair of operators C and D from $\mathfrak{H}$ to $\mathfrak{H}'$ such that*

$$C + D = R\ , \quad A = D^{*-1}R^*C = CR^*D^{*-1}\ . \tag{13.1}$$

The operators C and D are determined by the formulas

$$C^{-1} = A^{-1} + R^*\ , \quad D^{-1} = A^* + R^* \tag{13.2}$$

and have R as their associated elementary operator. The operators C and D are bounded and $*$-commute. In addition

$$C^{-1}RD^{-1} = D^{-1}RC^{-1} = C^{-1} + D^{-1} \tag{13.3a}$$

$$\begin{aligned} A^{-1} &= C^{-1}RD^* = D^*RC^{-1}\ , \quad A^* = C^*RD^{-1} = D^{-1}RC^*\ , \\ A^{*-1} &= C^{*-1}R^*D = DR^*C^{*-1}\ . \end{aligned} \tag{13.3b}$$

This result is an easy consequence of Lemma 5.2. The operator $A_1 = R^*A$ is self-adjoint and nonnegative. Let C_1 and D_1 be the bounded nonnegative self-adjoint operators related to A_1 as described in Lemma 5.2. The operators $C = RC_1$, $D = RD_1$ have the properties described in the theorem, as one readily verifies. An alternate proof can be made by defining C and D by (13.2) and making computations analongous to those made in the proof of Lemma 5.2. Finally, a proof can be made by the use of the integral representation (9.5) of A. In this case C and are defined by the formulas.

$$C = \int_0^\infty \frac{\lambda}{1+\lambda} dR_\lambda\ , \quad D = \int_0^\infty \frac{1}{1+\lambda} dR_\lambda\ .$$

THEOREM 13.2. *Let C be the operator related to A as described in Theorem 13.1. A bounded operator B $*$-commutes with A if and only if B $*$-commutes with C.*

If B $*$-commutes with A, then B $*$-commutes with R and $A^{*-1} = c^{*-1} - R$. Consequently B $*$-commutes with C^{*-1} and hence with C. Conversely if B $*$-commutes with, C, then B $*$-commutes with R, $C^{*-1} = A^{*-1} + R$, A^{*-1} and A. This proves the theorem. It is clear that the results described in the theorem hold equally well with C replaced by $D = R - C$.

We state without proof the following

THEOREM 13.3. *Let C be the operator related to A as described in Theorem* 13.1 *and let*

$$C_t = \left(\frac{1-t}{2}\right)R + tC\ , \quad D_t = R - C_t\,(-1 \leqq t \leqq 1)\ .$$

The one-parameter family of operators

$$A_t = D_t^{-1}R^*C_t = C_tR^*D_t^{-1} \quad (-1 \leqq t \leqq 1)$$

contains A for $t = 1$, A^{*-1} for $t = -1$, R for $t = 0$ and is such that A_t $(-1 < t < 1)$ is bounded and reciprocally bounded.

As a further result we have

THEOREM 13.4. *Let C and C_1 be the bounded operators related respectively to two operators A and A_1 as described in Theorem* 13.1. *Then A_1 is a section of A if and only if C_1 is a section of C.*

Let R_1 and R be the elementary operator associated with A_1 and A hence also with C_1 and C. If A_1 is a section of A, then

$$R_1C^{-1}R_1 = R_1A^{-1}R_1 + R_1R^*R_1 = A_1^{-1} + R_1 = C_1^{-1}\ .$$

Since R_1 $*$-commutes with C it follows that C_1 is a section of C. The converse is readily verified.

The result given in Theorem 13.2 enables us to extend the definition of $*$-commutativity to two unbounded operators A_1 and A_2. To this end let C_1 and C_2 be the bounded operators related respectively to A_1 and A_2 as described in Theorem 13.1. The operators A_1 and A_2 will be said to *$*$-commute* if the operators C_1 and C_2 $*$-commute. This definition is consistent with the one given heretofore for the case in which one of the operators is bounded. The result described in Lemma 12.3 is valid without the assumption that B is bounded.

14. Further decomposition of operators. In this section we assume that A and B are arbitrary operators from $\mathfrak{H}$ to $\mathfrak{H}'$. As an extension of Theorem 9.1 we have

THEOREM 14.1. *Let R and S be the elementary operators associated with A and B respectively. If B $*$-commutes with A, then A, R, B S are expressible uniquely as sums and differences*

$$(14.1)\qquad \begin{aligned} A &= A_0 + A_+ + A_-\ , \quad R = R_0 + R_+ + R_- \\ B &= B_0 + B_+ - B_-\ , \quad S = S_0 + R_+ - R_- \end{aligned}$$

of mutually $*$-orthogonal operators such that (α) R_0, R_+, R_- *are the elementary operators beloging to* A_0, A_+, A_- *respectively;* (β) S_0, R_+, R_- *are the elementary operators belonging to* B_0, B_+, B_- *respectively;* (γ) A_0 *and* R_0 *are $*$-orthogonal to* B_0 *and* S_0; (δ) B_+ *$*$-commutes with* A_+ *and* B_- *$*$-commutes with* A_-. *Conversely if* A, R, B, S *are so expressible then* B *and* S *$*$-commute with* A *and* R.

In view of the results given in the last section we may assume that A and B are bounded. Suppose that B $*$-commutes with A, then B $*$-commutes with R, and S $*$-commutes with A, by Lemma 12.3. By virtue of Theorem 9.1 applied to A, R, S, it is seen that A, R, S have the decomposition (14.1) such that condition (α) holds and S_0 is $*$-orthogonal to R_0. Applying Theorem 9.1 to the operators B, S, R it is seen that B, S, R have the decomposition (14.1) such that condition (β) holds and S_0 is $*$-orthogonal to R_0. Since the decomposition of R and S are unique, the decomposition (14.1) holds such that (α), (β), and (γ) hold. Since

$$A^*B = A_+^*B_+ - A_-^*B_- \,, \quad AB^* = A_+B_+^* - A_-B_-^*$$

are self-adjoint it follows that each of the operators on the right are also self-adjoint. Consequently B_+ $*$-commutes with A_+ and B_- $*$-commutes with A_-. The converse is immediate and the lemma is established.

COROLLARY. *If B $*$-commutes with A there is elementary operator T such that* $A = TA^*T$ *and* $B = TB^*T$.

The operator $T = S_0 + R$ has this property.

THEOREM 14.2. *Suppose that A and B $*$-commute and are self-adjoint relative to an elementary operator T. Then the operators* A_0, A_+, A_- B_0, B_+, B_- *described in Theorem* 14.1 *are also self-adjoint relative to T.*

Since the elementary operators R and S belonging to A and B are self-adjoint relative to T it follows that

$$R_+ = \tfrac{1}{2}(SR^*S + RS^*R) \,, \quad R_- = \tfrac{1}{2}(SR^*S - RS^*R) \,,$$

are self-adjoint relative to T. The same is true for R_0 and S_0. The theorem follows readily with the help of Lemma 12.4.

The result just given can be extended as described in the following

THEOEM 14.3. *Suppose that B $*$-commutes with A and T is an elementary operator such that* $A = TA^*T$ *and* $B = TB^*T$. *Then* T, *A and B can be decomposed uniquely in the form*

$$
\begin{aligned}
T = T_0 &= T_1 + T_2 + T_3 + T_4 + T_5 + T_6 + T_7 + T_8 \\
A &= A_1 - A_2 + A_3 + A_4 + A_5 - A_6 \qquad (14.2)\\
B &= B_3 - B_4 - B_5 + B_6 + B_7 - B_8
\end{aligned}
$$

into mutually $*$-orthogonal operators such that (α) *T_j is the elementary operator associated with A_j ($j = 1, 2, \cdots, 6$)* (β) *T_k is the elementary operator associated with B_k ($k = 3, 4, \cdots, 8$)* (γ). *The operators T_i, A_j, B_k $*$-commute.*

In order to prove this result let A, B, R, S have the decomposition (14.1). By virtue of the last theorem the operator T $*$-commutes with each of the operators given in (14.1). Applying Theorem 14.1 to A_0 and T we see that A_0 can be expressed as the difference $A_0 = A_1 - A_2$ of $*$-orthogonal operators A_1 and A_2 whose associated elementary operators T_1 and T_2 are sections of T. Similarly $B_0 = B_7 - B_8$, where B_7 and B_8 are $*$-orthogonal operators whose associated elementary operators T_7 and T_8 are section of T. Applying Theorem 14.1 to A_+, T; B_+, T; A_-, T and B_-, T we obtain differences $A_+ = A_3 - A_4$, $B_+ = B_3 - B_4$, $A_- = A_5 - A_6$, $B_- = B_5 - B_6$ of $*$-orthogonal operators such that A_i, B_i have the same associated elementary operator T_i, a section of T.

From these relations one obtains the decomposition (14.2), the section T_0 of T being $*$-orthogonal to A and B. In view of Theorem 14.1, the operator T_i, A_j, B_k $*$-commute. This proves the theorem.

THEOREM 14.4. *Let $A_{\lambda+}, A_{\lambda 0}, A_{\lambda-}$ ($0 < \lambda < \infty$) be the sections of A described in Theorem 9.2 and let $B_{\mu+}, B_{\mu 0}, B_{\mu-}$ ($0 < \mu < \infty$) be the corresponding sections of B. Suppose that B $*$-commutes A. Then the operators $B, B_{\mu+}, B_{\mu 0}, B_{\mu-}, B_\mu = B_{\mu 0} + B_{\mu-}$ $*$-commute with each of the operators $A, A_{\lambda+}, A_{\lambda 0}, A_{\lambda-}, A_\lambda = A_{\lambda 0} + A_{\lambda-}$.*

In order to prove this result recall, by Theorem 9.2, that $T = R_{\lambda+} - R_{\lambda-}$ is the elementary operator associated with $C = A - \lambda R$. Since B $*$-commutes with A it $*$-commutes with R, C, T, RC^*T and hence also with

$$C_+ = C + RC^*T = 2(A_{\lambda+} - \lambda R_{\lambda+}) \, , \quad C_- = C - RC^*T = 2(A_{\lambda-} - \lambda R_{\lambda-}) \, .$$

The elementary operators of C_+ and C_- are $R_{\lambda+}$ and $R_{\lambda-}$ respectively. It follows that B $*$-commutes with $A_{\lambda+}, A_{\lambda-}$ and hence also with $A_{\lambda 0}$. Similarly $B_{\mu+}, B_{\mu 0}, B_{\mu-}$ $*$-commutes with A. The operators therefore $*$-commute, as described in the theorem.

THEOREM 14.5. *Suppose that B $*$-commutes with A and that T is* *an elementary operator such that $A = TA^*T$, $B = TB^*T$. Let A_λ, T_λ $(-\infty < \lambda < \infty)$ be the sections A and T described in the corollary to Theorem 9.3. Let $\tilde{B}_\mu$, $\tilde{T}_\mu$ $(-\infty < \mu < \infty)$ be the sections of B and T obtained by having B playing the role of A in this corollary. Then A, A_λ, T_λ, B, $\tilde{B}_\mu$, $\tilde{T}_\mu$, $T_{\lambda\mu} = T_\lambda T_* \tilde{T}_\mu$ $*$-commute with each other. Moreover*

$$A = \int_{-\infty}^{\infty} \lambda d T_\lambda = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \lambda d T_{\lambda\mu} , \quad B = \int_{-\infty}^{\infty} \mu d \tilde{T}_\mu = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mu d T_{\lambda\mu} .$$

If the scalars are the complex numbers, then

$$A + iB = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (\lambda + i\mu) d T_{\lambda\mu} .$$

15. Bounded normal operators relative to T. In the present section it will be assumed that the scalars are the complex numbers. Let T be an elementary operator from $\mathfrak{H}$ to $\mathfrak{H}'$ and let $\mathfrak{A}(T)$ be the class of all bounded operators A such that the relation

$$AT^*T = TT^*A = A \tag{15.1}$$

holds. Let $\mathscr{B}(T)$ be the class of all operators A in $\mathfrak{A}(T)$ that $*$-commute with T. These are the operators A in $\mathfrak{A}(T)$ that satisfy the relation $A = TA^*T$, that is, the operators in $\mathscr{B}(T)$ that are self-adjoint relative to T. Every operator A in $\mathfrak{A}(T)$ is expressible uniquely in the form $A = A_1 + iA_2$ where A_1 and A_2 are in $\mathfrak{B}(T)$. The operators A_1 and A_2 are given by the formulas

$$A_1 = \tfrac{1}{2}(A + TA^*T) , \quad A_2 = \tfrac{1}{2i}(A - TA^*T) . \tag{15.2}$$

It should be observed that, by virtue of Lemma 8.4, two operators A and B in $\mathscr{B}(T)$ $*$-commute if and only if $AT^*B = BT^*A$.

Let $\mathscr{C}(T)$ be the class of all operators A in $\mathfrak{A}(T)$ such that $TA^*A = AA^*T$. An operator A in $\mathscr{C}(T)$ will be said to be *normal with respect to T*. It is clear that an operator that is self-adjoint relative to T is also normal relative to T. An operator A in $\mathfrak{A}(T)$ is in $\mathscr{C}(T)$ if and only if the operators A_1 and A_2 defined by (15.2) $*$-commute. In order to prove this fact observe that

$$B = A_1T^*A_1 + A_2T^*A_2 , \quad C = A_1T^*A_2 - A_2T^*A_1$$

are in $\mathscr{B}(T)$ and

$$TA^*A = B + iC , \quad AA^*T = B - iC .$$

Consequently $TA^*A = AA^*T$ if and only if $C = 0$, that is, if and only if A_1 $*$-commutes with A_2. If A is in $\mathscr{C}(T)$, there is by virtue of Theorem 14.5, a section T_α corresponding to each complex number α

such that

$$A = \int \alpha d T_\alpha$$

where the integral is taken over the complex plane.

Given an operator A in $\mathscr{C}(T)$ let $\mathscr{C}(A, T)$ be the class of all operators B in $\mathscr{C}(T)$ such that $TA^*B = BA^*T$ and $AT^*B = BT^*A$. If B is in $\mathscr{C}(A, T)$, then $TB^*A = AB^*T$ also. Moreover, TB^*T is in $\mathscr{C}(A, T)$. Let $\mathfrak{M}(A, T)$ be all operators B in $\mathscr{C}(T)$ such that $\mathscr{C}(A, T) \subset \mathscr{C}(B, T)$. If B and C are in $\mathfrak{M}(A, T)$, so also are $\alpha B + \beta C$ and BT^*C, where α and β are complex mumbers. Moreover

$$\| BT^*C \| \leqq \|B\| \, \| C \| .$$

It follows that if we define BT^*C to be the product of B and C, the class $\mathfrak{M}(A, T)$ is a Banach algebra with the operator T as a unit element and TB^*T as an involution. The subclass $\mathscr{L}(A, T)$ of all operators B in $\mathfrak{M}(A, T)$ such that $B = TB^*T$ form a Banach algebra over the reals.

16. Compact and reciprocally compact operators. An operator A from $\mathfrak{H}$ to $\mathfrak{H}'$ will be said to be *compact* if given a bounded sequence $\{x_n\}$ in $\mathscr{D}_A$, the sequence $\{Ax_n\}$ has a strongly convergent subsequence. An operator A will be said to be *reciprocally compact* if its reciprocal is compact. Since compact operators are bounded, it follows that reciprocally compact operators are reciprocally bounded. It should be observed that an operator A is compact if an only if given a weakly convergent sequence $\{x_n\}$ in $\mathscr{D}_A$, the sequence $\{Ax_n\}$ converges strongly.

THEOREM 16.1 *An operator A is of finite rank if and only if it is compact and reciprocally bounded. An operator A is of finite rank if and only if it is bounded and reciprocally compact An operator A is of finite rank if and only if it is compact and reciprocally compact.*

Suppose that A is compact and reciprocally bounded. Then $\mathscr{C}_A$ and $\mathscr{R}_A$ are closed. Let $\{x_n\}$ be a sequence in $\mathscr{C}_A$ converging weakly to a point x_0. Since A is compact $y_n = Ax_n$ converges strongly to $y_0 = Ax_0$. it follows that $x = A^{-1}y_n$ converges strongly to $x_0 = A^{-1}y$. Consequently weak convergence on $\mathscr{C}_A$ implies strong convergence. It follows that $\mathscr{C}_A$ is of finite dimension. Hence A is of finite rank. Conversely if A is of finite rank, then A is compact and reciprocally bounded. The remaining statements follow readily.

THEOREM 16.2 *Let A be the sum $A = B + C$ of two $*$-orthogonal* *operators B and C from $\mathfrak{H}$ to $\mathfrak{H}'$. Then A is compact, reciprocally compact, bounded, or reciprocally bounded if and only if B and C have the same property. If C is of finite rank, then A is compact, reciprocally compact, bounded, or reciprocally bounded if and only if B has the same property.*

The first conclusion is immediate from the definitions of the terms involved. The second follows from the first. In view of the second statement sections of finite rank can be disregarded in determining the properties of compactness, reciprocal compactness, boundedness and reciprocal boundedness.

THEOREM 16.3. *An operator A is compact if and only if its reciprocally bounded sections are of finite rank. Similarly, an operator A is reciprocally compact if and only if its bounded sections are of finite rank.*

The second statement follows from the first. If A is compact, its sections are compact and hence its reciprocally bounded sections are of finite rank, by Theorem 16.1. Suppose now that A is an operator whose reciprocally bounded sections are finite rank. Then as was seen in § 9, given a number $\lambda > 0$, the operator A can be written as the sum $A = A_{\lambda+} + A_\lambda$ of two $*$-orthogonal operators such that $A_{\lambda+}$ is reciprocally bounded and A_λ is of norm at most λ. In view of our hypothese $A_{\lambda+}$ is of finite rank and hence is compact. It follows that A is bounded and that $\mathscr{D}_A = \mathfrak{H}$. Let $\{x_n\}$ be a sequence in $\mathscr{D}_A$ converging weakly to zero. Then,

$$|| Ax_n || \leqq || A_{\lambda+}x_n || + || A_\lambda x_n || \leqq || A_{\lambda+}x_n || + \lambda || x_n || .$$

Since $A_{\lambda+}$ is compact we have $\lim_{n=\infty} || A_{\lambda+}x_n || = 0$. Consequently $\lim_{n=\infty} \sup || Ax_n || \leqq \lambda M$ where M is a bound for the sequence $|| x_n ||$. Since λ is arbitrary it follows that $Ax_n \Rightarrow 0$ and hence that A is compact, as was to be proved.

THEOREM 16.4. *An operator A is compact if and only if its spectrum (apart from $\lambda = 0$) consists of a bounded set of isolated principal values of finite order. It is reciprocally compact if and only if its spectrum consists of isolated principal values of finite order bounded away from zero.*

Again, the second statement follows from the first. In order to prove the first statement we use the decomposition $A = A_{\lambda+} + A_\lambda$ of A into the $*$-orthogonal sections described in § 9, where λ is an arbitrary positive number. The points of the spectrum of A that exceed λ comprise the spectrum of $A_{\lambda+}$. The remaining points of the spectrum of A comprise the spectrum of A_λ. If A is compact, then $A_{\lambda+}$ is of finite rank. Consequently the points of the spectrum of A that exceed λ consist of a finite number of principal values of $A_{\lambda+}$, each being of finite order. Since λ is arbitrary it follows that the spectrum of A consists of a bounded set of isolated principal values of finite order. Conversely if the spectrum of A consists of a bounded set of isolated principal values of finite order, then $A_{\lambda+}$ is of finite rank for every value of λ. Consequently A is compact, as was to be proved.

The following corollary is immediate.

COROLLARY. *If one of the operators A, A^*, A^*A, AA^* is compact, then the others are compact. Similarly, if one of the operators A, A^*, A^*A, AA^* is reciprocally compact so also are the others.*

17. Operators of finite character. By the *nullity* of an operator will be meant the dimension of its null space. An operator A will be said to be of *finite character* if it is of finite nullity and if its bounded sections have finite rank, or equivalently by, if it is of finite nullity and is reciprocally compact. Operators of this type play an important role in the calculus of variations and in existence theorems for elliptic partial differential equations. In fact the condition of ellipticity is equivalent to the condition that an operator be of finite character relative to a suitably chosen norm, provided the domain of the independent variable is bounded. The operators described in § 4 are of finite character.

THEOREM 17.1. *An operator A is of finite character if and only if given a bounded sequence $\{x_n\}$ in $\mathscr{D}_A$ such that $\{Ax_n\}$ is also bounded, then $\{x_n\}$ has a strongly convergent subsequence. An operator A is of finite character if and only if it is of finite nullity and given a sequence $\{x_n\}$ in the carrier $\mathscr{C}_A$ of A such that $\{Ax_n\}$ is bounded, then $\{x_n\}$ has a strongly convergent subsequence.*

Suppose that A is of finite character. Then the nullity of A is finite, and A^{-1} is compact. Let $\{x_n\}$ be a sequence in $\mathscr{C}_A$ such that $\{Ax_n\}$ is bounded. Setting $y_n = Ax_n$ we have $x_n = A^{-1}y_n$. Since $\{y_n\}$ is in the carrier of A^{-1} and A^{-1} is compact it follows that $\{x_n\}$ has a strongly convergent subsequence. Suppose next that $\{x_n\}$ is a bounded sequence in $\mathscr{D}_A$ such that $\{Ax_n\}$ is bouneded. Then x is expressible in the form $x_n = x_{n0} + x_{n1}$, where $x_{n0} \in \mathfrak{N}_A$ and $x_{n1} \in \mathscr{C}_A$. Since $\mathfrak{N}_A$ is of finite dimension and $Ax_n = Ax_{n1}$, the boundedness conditions imposed imply that $\{x_n\}$ has a strongly convergent subsequence. The criteria given in the theorem are accordingly necessary conditions for A to be of character.

Suppose conversely that every bounded sequence $\{x_n\}$ in $\mathscr{D}_A$ for which $\{Ax_n\}$ is bounded has a strongly convergent subsequence. Then the nullity of A is finite, since otherwise there would exist a orthonormal sequence $\{x_n\}$ in $\mathfrak{N}_A$. Such a sequence would have $Ax_n = 0$ and would possess no strongly convergent subsequence. The reciprocal A^{-1} is bounded. If this were not so we could select a sequence $\{x_n\}$ in $\mathscr{C}_A$ such that $\| x_n \| = 1$ and $\| Ax_n \| \leqq 1/n$. In view of the last inequality the sequence could be chosen so as to converge strongly to a vector x_0. Since A is closed it would follow that x_0 would be in $\mathscr{C}_A$, $\| x_0 \| = 1$ and $Ax_0 = 0$. This is impossible. Hence A^{-1} is bounded. Consider next a bounded sequence $\{y_n\}$ in $\mathscr{C}_{A^{-1}}$. Set $x_n = A^{-1}y_n$. Since A^{-1} is bounded the sequence $\{x_n\}$ is also bounded and hence, by our criterion, has a strongly convergent subsequence. The operator A^{-1} is therefore compact. Hence A is of finite character, as was to be proved.

COROLLARY 1. *Let A be an operator from $\mathfrak{H}$ to $\mathfrak{H}'$ and let B be the operator that maps a point x in $\mathscr{D}_A$ into the pair $\{x, Ax\}$ in $\mathfrak{H} \times \mathfrak{H}'$. The nullity of B is zero. Moreover A is of finite character if and only if B is of finite character.*

COROLLARY 2. *If B and C are $*$-orthogonal operators and C is finite rank, then $A = B + C$ is of finite character if and only if B is of finite character.*

Let T be an elementary operator such that $A = TA^*T$ and let R be the elementary operator associated with A. By Theorem 9.1, T is expressible uniquely in the form $\mathrm{T} = T_0 + R_+ - R_-$ where T_0, R_+, R_- are $*$-orthogonal and $\mathrm{R} = R_+ + R_-$. The operator T will be said to be *of finite index relative* to A in case one of the operators R_+ and R_- is of finite rank. The minimum of the ranks of R_+ and R_- will be called the *index* of T. Clearly the index of T is the minimum of the ranks of the sections $A_+ = R_+A^*R_+$ and $A_- = R_-A^*R_-$ of A. In the self-adjoint case with $T = I$, the identity, this index is the smaller of the ranks of the orthogonal nonnegative operators A_1, A_2 such that $A = A_1 - A_2$. In this event this index is frequently called *the index* of A or of the quadratic form (Ax, x).

THEOREM 17.2. *Let T be an elementary operator such that $TA^*T = A$. Every bounded sequence $\{x_n\}$ such that $\{(Ax_n, Tx_n)\}$ is bounded has a strongly convergent subsequence if and only if A is of finite character and T is of finite index relative to A.*

This criterion, stated in a somewhat different form, is the basis for a large class of existence theorems for weak solutions of partial differential equations.

Since $\| T^*Ax \| = \| Ax \|$ it follows from Theorem 17.1 that A is finite character if and only if T^*A is of finite character. Moreover T^*A is self-adjoint. It follows that it is sufficient to consider the case $A = A^*$ and $T = I$. Let $\{x_n\}$ be a bounded sequence such that $\{Ax_n\}$ is bounded. From the inequality

$$| (Ax, x) | \leqq \| x \| \| Ax \|$$

it follows that $\{(Ax_n, x_n)\}$ is bounded. Consequently if the criterion described in the theorem holds, then $\{x_n\}$ has a strongly convergent subsequence. By virtue of Theorem 17.1 the operator A is of finite character. It remains to show that if A is expressed as the difference $A = B - C$ of two orthogonal nonnegative self-adjoint operators, then either B or C is of finite rank. If this were not the case one could select an orthogonal sequence $\{y_n\}$ in $\mathscr{C}_B$ and $\{z_n\}$ in $\mathscr{C}_C$ such that $(By_m, y_n) = (Cz_m z_n) = \delta_{mn}$. The vectors $x_n = \alpha_n(y_n + z_n)$ then satisfy the relation

$$(Ax_m, x_n) = \alpha_m\alpha_n[(By_m, y_n) - (Cy_m, y_n)] = 0 \ \ (m, n = 1, 2, 3 \cdots) .$$

Choosing α_n such that $\| x_n \| = 1$, we obtain an orthogonal sequence $\{x_n\}$ such that $(Ax_n, x_n) = 0$. This sequence cannot have a strongly convergent subsequence. Consequently either B or C is of finite rank, as was to be proved.

Conversely suppose that B or C is of finite rank and $A = B - C$ is of finite character. For definiteness suppose that C is of finite rank. Then B is of finite character. Consider now a bounded sequence of vectors $\{x_n\}$ in $\mathscr{C}_A$, such that $\{(Ax_n, x_n)\}$ is bounded. Select y_n in $\mathscr{C}_B$ and z_n in $\mathscr{C}_C$ such that $x_n = y_n + z_n$. Then

$$(Ax_n, x_n) = (By_n, y_n) - (Cz_n, z) .$$

It follows that $\{(By_n, y_n)\}$ is bounded. Consequently, $\{y_n\}$ has a convergent subsequence. Since $\{z_n\}$ is restricted to a finite dimensional subspace of $\mathscr{D}_A$, it follows that $\{x_n\}$ has a strongly convergent subsequence. This completes the proof of the theorem.

THEOREM 17.3. *Let A be an operator from $\mathfrak{H}$ to $\mathfrak{H}'$ of finite character and let B be an operator from $\mathfrak{H}$ to a Hilbert space $\mathfrak{H}''$. If $\mathscr{D}_B \subset \mathscr{D}_A$, then B is of finite character.*

Since $\mathscr{D}_B \subset \mathscr{D}_A$ there is a constant α such that if x is in $\mathscr{D}_B$ then

$$\| Ax \| \leqq \alpha \{\| Bx \| + \| x \|\} .$$

If $\{x_n\}$ is a sequence in $\mathscr{D}_B$ such that $\| x_n \|$, $\| Bx_n \|$ are bounded, then $\| Ax_n \|$ is bounded also. It follows from Theorem 17.1 that $\{x_n\}$ converges strongly in subsequence. Consequently B is of finite character, by virtue of Theorem 17.1.

A linear transformation K from $\mathfrak{H}$ to $\mathfrak{H}''$ will be said to be *compact relative* to A if $\mathscr{D}_K \supset \mathscr{D}_A$ and if for every bounded sequence $\{x_n\}$ in $\mathscr{D}_A$ such that $\{Ax_n\}$ is bounded, the sequence $\{Kx_n\}$ has a strongly convergent subsequence.

THEOREM 17.4. *Let A be an operator from $\mathfrak{H}$ to $\mathfrak{H}'$ of finite character. Let K be an operator from $\mathfrak{H}$ to $\mathfrak{H}''$ such that $\mathscr{D}_K \supset \mathscr{D}_A$. Then K is compact relative to A if and only if given a positive number α there is a number β such that the inequality*

$$\| Kx \| \leqq \alpha \| Ax \| + \beta \| x \| \tag{17.1}$$

holds on $\mathscr{D}_A$.

Suppose that K is compact relative to A. Suppose further there is an $\alpha > 0$ such that (17.1) holds on $\mathscr{D}_A$ for no constant β. We can select a non-null sequence $\{x_n\}$ such that

$$\| Kx_n \| \geqq \alpha \| Ax_n \| + n \| x_n \| .$$

We can suppose that $\| Kx_n \| = 1$. Then $\| Ax_n \|$ is bounded and $x_n \Rightarrow 0$. Since K is compact relative to A it follows that $Kx_n \Rightarrow 0$, in subsequence, contrary to the fact that $\| Kx_n \| = 1$.

Suppose that (17.1) holds as stated. Let $\{x_n\}$ be a bounded sequence such that $\{Ax_n\}$ is bounded. A subsequence, rename it $\{x_n\}$, converges strongly to a vector x_0. The point x_0 is in $\mathscr{D}_A$ since A is closed. Given $\alpha > 0$ choose β so that (17.1) holds. Then

$$\| K(x_n - x_0) \| \leqq \alpha \| A(x_n - x_0) \| + \beta \| x_n - x_0 \|$$

and

$$\limsup_{n=\infty} \| Kx_n - Kx_0 \| \leqq \alpha \limsup \| A(x_n - x_0) \| .$$

Since α is arbitrary it follows that $\{Kx_n\}$ converges strongly to Kx_0. The operator K is therefore compact relative to A, as was to be proved.

THEOREM 17.5. *Let A and K be operators from $\mathfrak{H}$ to $\mathfrak{H}'$ such that K is compact relative to A. The operator A is of finite character if and only if $B = A + K$ is an operator of finite character.*

In order to see that B is closed when A is of finite character let $\{x_n\}$ be a sequence such that $x_n \Rightarrow x_0$, $Bx_n \Rightarrow y_n$. In view of (17.1) with $\alpha < 1$

$$|| Kx_n || \leqq \alpha || Ax_n || + \beta || x_n || \leqq \alpha || Bx_n || + \alpha || Kx_n || + \beta || x_n || .$$

We see that $\{Kx_n\}$ is bounded. Consequently $\{Ax_n\}$ is bounded also. It follows that $\{Kx_n\}$ converges to Kx_0 and that $Ax_n \Rightarrow y_0 - Kx_0$. Since A is closed $y_0 - Kx_0 = Ax_0$, that is, $y_0 = Bx_0$. The operator B is accordingly closed. Since B and A has the same domain, B is of finite character. Conversely if B is an operator of finite character, so also is A since $\mathscr{D}_B = \mathscr{D}_A$.

In a similar manner we obtain

THEOREM 17.6. *Let A be an operator from $\mathfrak{H}$ to $\mathfrak{H}'$ and let K be an operator from $\mathfrak{H}$ to $\mathfrak{H}''$ that is compact relative to A. Let B be the operator that maps a point x in $\mathscr{D}_A$ into the point $\{Ax, Kx\}$ in $\mathfrak{H}'$ x $\mathfrak{H}''$. Then A is of finite character if and only if B is an operator of finite character.*

THEOREM 17.7. *Let A be an operator from $\mathfrak{H}$ to $\mathfrak{H}$ and suppose that every bounded sequence $\{x_n\}$ in $\mathscr{D}_A$ for which $\{(Ax_n, x_n)\}$ is bounded has a strongly convergent subsequence. Then A is of finite character. Moreover, a linear subclass $\mathscr{B}$ of $\mathscr{D}_A$ on which $(Ax, x) = 0$ is of finite dimension.*

The proof of this result is like that of Theorem 17.2 and is equivalent to the result given in Theorem 17.2 is $A = A^*$. In this theorem the role of (Ax, x) may be replaced by $(Ax, x) + (x, Ax)$.

18. **Elliptic partial differential equations.** The purpose of the present section is to indicate the connections between the results described in the preceding pages with the theory of elliptic partial differential equations. To this end let Ω be a bounded region in an m-dimensional Euclidean space of points $t = (t_1, \cdots, t_m)$. The boundary of Ω will be assumed to be nonsingular and to be of class C^∞. The results given below are valid under much weaker assumptions but we shall not consider them at this time.

The symbol α will be used to designate an m-tuple $\alpha = (\alpha_1, \cdots, \alpha_m)$ of nonnegative integers. Let $|\alpha| = \alpha_1 + \cdots + \alpha_m$. The symbol D_α. will be used to denote the differential operator

$$D_\alpha = (-i)^{|\alpha|} \frac{\partial^{|\alpha|}}{\partial t_1^{\alpha_1} \cdots \partial t_m^{\alpha_m}} .$$

Let $\mathfrak{H}_{nk}$ be the class of all Lebesgue square integrable complex valued functions $x_\alpha^j(t)$ $(t \in \Omega; j = 1, \cdots, n; |\alpha| = k)$, normalized so as to be equal to the limit of their integral means whenever this limit exists and to be zero elsewhere. The class $\mathfrak{H}_{nk}$ with

$$(x, y)_k = \int_\Omega x_\alpha^j(t)\overline{y_\alpha^j(t)}dt \ (j \leqq n, |\alpha| = k) ,$$

(j and α summed) as its inner product forms a Hilbert space over the field of complex number. The symbol x_k will be used to denote an element in $\mathfrak{H}_{nk}$. The cartesian product of $\mathfrak{H}_{n0}, \mathfrak{H}_{n1}, \cdots, \mathfrak{H}_{nk}$ will be denoted by $\mathfrak{H}_n^k$. Its elements x are of the form $x = (x_0, x_1, \cdots, x_k)$. An element x in $\mathfrak{H}_n^k$ such that x_0: $x^j(t)$ is of class C^k and such that x_r is the set of derivatives $x_\alpha^j = D_\alpha x^j \ |\alpha| = r$ of order r will be denoted by $\mathscr{C}_n^k$. The closure of $\mathscr{C}_n^k$ will be denoted by $\mathscr{D}_n^k$. In view of our normalization of the functions in $\mathfrak{H}_{nr}$, it can be shown the formula $x_\alpha^j(t) = D_\alpha x^j(t)$ $|\alpha| = r \leqq k$ holds almost everywhere on Ω, where $x^j(t)$ are the functions defining x_0 in $(x_0, x_1, \cdots, x_k)$. The projection of $\mathscr{D}_k^n$ in $\mathfrak{H}_{nk}$ will be denoted by $\mathscr{D}_{nk}$. The class $\mathscr{D}_{nk}$ is a closed subspace of $\mathfrak{H}_{nk}$.

Since an element $(x_0, x_1, \cdots, x_k)$ in $\mathscr{D}_n^k$ is uniquely determined by its inital element x_0, a function G_k on $\mathfrak{H}_n^0 = \mathfrak{H}_{n0}$ to $\mathfrak{H}_n^k$ is defined. The range of the function is $\mathscr{D}_n^k$. Its domain $\mathscr{D}_{G_k}$ is the projection of $\mathscr{D}_n^k$ on $\mathfrak{H}_{n0}$. The functions G_k $(k = 1, 2, 3, \cdots)$ have the following properties:

(1) The function G_k is a closed and dense linear transformation from $\mathfrak{H} = \mathfrak{H}_{n0}$ to $\mathfrak{H}_n^k$.

(2) The operator G_k $(k > 0)$ is of finite character and zero nullity.

(3) The operator G_j $(j < k)$ is compact relative to G_k.

These results follow from well known connections between partial derivatives and can be found in papers on this subject.

Let C be a bounded operator from $\mathscr{D}_n^k$ $(k > 0)$ to a Hilbert space $\mathfrak{H}_{q0}$. Given a restriction B_k of G_k that is closed and dense in $\mathfrak{H} = \mathfrak{H}_{n0}$, the product $A_k = CB_k$ defines a dense linear transformation. Such an operator will be said to be *elliptic* in case it is closed. This definition of ellipticity is an extension of the one usually given. An elliptic operator of this type is necessarily of finite character by Theorem 17.3 since B_k has this property. It is clear that A_k is elliptic if and only if there is a constant $h > 0$ such that

$$\| B_k x \| \leqq h [\| A_k x \| + \| x \|] \tag{18.1}$$

for all x in $\mathscr{D}_{B_k}$. It should be observed that if A_k $(k \geqq 1)$ is elliptic, then the equation $A_k x = y$ has a solution x for all y orthogonal to the solutions z of $A^*_k z = 0$. The existence of strong solutions is thereby established.

In order to illustrate these ideas consider the case in which the operator C is defined by a formula of the form

$$c_\alpha^{\delta j}(t)x_\alpha^j(t) \quad (\delta = 1, \cdots, q; j = 1, \cdots, n: |\alpha| \leqq k) \tag{18.2}$$

where j and α are summed and the coefficients are continuous on the closure of Ω. Select $B_k = G_k$. Then $A_k = CB_k$ is elliptic, that is, an inequality of the form (17.1) holds if and only if the following two conditions are met:

(1) Given a point t in Ω there is no non-null set of real numbers $\xi = (\xi_1, \cdots, \xi_m)$ and no non-null set of complex numbers $\zeta = (\zeta^1, \cdots, \zeta^n)$ such that the relations

$$C_\alpha^{\sigma j}\xi^\alpha\zeta^j = 0 \quad (\sigma = 1, \cdots, q, |\alpha| = k) \tag{18.3}$$

holds, where $\xi^\alpha = \xi_1^{\alpha_1}\xi_2^{\alpha_2} \cdots \xi_m^{\alpha_m}$.

(2) Given a point t on the boundary of Ω the relations (18.3) cannot be satisfied by non-null complex numbers $\zeta = (\zeta^1, \cdots, \zeta^n)$ and by non-null numbers $\xi = (\xi_1, \cdots, \xi_m)$ whose normal component is complex and whose tangential components are real.

If the first of these conditions is met then $A_k = CB_k$ is elliptic, where B_k is the restriction of G_k defined by the closure of the subclass of $\mathscr{D}_n^k$ whose elements are continuous and have $x_\alpha^j(t) = 0$ ($|\alpha| < k$) on the boundary of Ω.

These and related results can be found in the recent papers[6] on partial differential equations by Aronszajn Browder, Friedrichs, Gaarding, Hormander, Morrey, Nirenberg, Schechter and the author.

For a list of references see Hestenes, Magnus R. Quadratic Variational-theory and linear elliptic partial differential equations to be published soon in the Transactions of the American Mathematical Society.

ON A THEOREM OF FEJÉR

FU CHENG HSIANG

1. Let

$$T: (\tau_{n\nu}) \qquad (n = 0, 1, 2, \cdots; \nu = 0, 1, 2, \cdots)$$

be an infinite Toeplitz matrix satisfying the conditions

$$\text{(i)} \qquad \lim \tau_{n\nu} = 0$$

for every fixed ν,

$$\text{(ii)} \qquad \lim \sum_{\nu=0}^{\infty} \tau_{n\nu} = 1$$

and

$$\text{(iii)} \qquad \sum_{\nu=0}^{\infty} |\tau_{n\nu}| \leqq K ,$$

K being an absolute constant independent of n.

Given a sequence (S_n) if

$$\lim \sum_{\nu=0}^{\infty} \tau_{n\nu} S_\nu = S ,$$

then we say that the sequence (S_n) or the series with partial sums S_n is summable (T) to the sum S.

2. Suppose that $f(x)$ is integrable in the Lebesgue sense and periodic with period 2π. Let

$$f(x) \sim \frac{1}{2} a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) .$$

Let

$$\sum_{n=1}^{\infty} n(b_n \cos nx - a_n \sin nx) = \sum B_n(x)$$

be the derived series of the Fourier series of $f(x)$. Fixing x, we write

$$\psi_x(t) = f(x + t) - f(x - t) .$$

Fejér [1] has proved the following

Received July 20, 1960.

THEOREM A. *If $f(x)$ is of bounded variation in* $(0, 2\pi)$, *then* $\{B_n(x)\}$ *is summable* (C, r) *to the jump* $l(x) = \{f(x+0) - f(x-0)\}/\pi$ *for every* $r > 0$ *at each point* x.

Recently, Siddiqi [3] extended Fejér's result and established the following

THEOREM B. *Let* $\Lambda: (\lambda_{n\nu})$ *be a triangular Toeplitz matrix*, i.e., $\lambda_{n\nu} = 0$ *for* $\nu > n$. *If it satisfies, in addition, the condition*

$$\text{(iv)} \qquad \sum_{\nu=0}^{n} | \Delta\lambda_{n\nu} | = 0(1)$$

as $n \to \infty$, *then* $\{B_n(x)\}$ *is summable* (Λ) *to* $l(x)$.

It is known [2] that a series which is summable by the harmonic means is also summable (C, r) for every $r > 0$ but not conversely. We take, for the (C, r) means, $\lambda_{n\nu} = A_{n-\nu}^{r-1}/A_n^r$,

$$A_n^r = \Gamma(n + r + 1)/\Gamma(n + 1)\Gamma(r + 1) ,$$

and for the harmonic means, $\lambda_{n\nu} = 1/(n - \nu + 1)$. Both satisfy (iv). Thus, we infer that Siddiqi's theorem contains Fejér's as a special case.

In this note, we develop Siddiqi's theorem into the following general form for the summability (T) of $\{B_n(x)\}$ at a given point.

THEOREM. *If* $\psi_x(t)$ *is of bounded variation in the neighborhood of* $t = 0$ *and absolutely continuous in* (η, π) *for any* $0 < \eta < \pi$, *then* $\{B_n(x)\}$ *is summable* (T) *to the jump* $l(x)$ *at* x.

3. Let us consider

$$\begin{aligned}
\sigma_n(x) &= \sum_{\nu=0}^{\infty} \tau_{n\nu} B_\nu(x) \\
&= \frac{1}{\pi} \sum_{\nu=0}^{\infty} \tau_{n\nu} \int_0^\pi \psi_x(t) \nu \sin \nu t dt \\
&= l(x) \sum_{\nu=0}^{\infty} \tau_{n\nu} + \frac{1}{\pi} \sum_{\nu=0}^{\infty} \tau_{n\nu} \int_0^\pi \cos \nu t d\psi_x(t) \\
&= l(x) + 0(1) + \frac{1}{\pi} \sum_{\nu=0}^{\infty} \tau_{n\nu} I_\nu .
\end{aligned}$$

We are going to prove that $\sum \tau_{n\nu} I_\nu = 0(1)$ as $n \to \infty$. Since $\psi_x(t)$ is of bounded variation in the neighborhood of $t = 0$, for a given $\varepsilon > 0$, we can choose $\delta > 0$ such that

$$\int_0^\delta | d\psi_x(t) | < \varepsilon .$$

Write

$$I_\nu = \left(\int_0^\delta + \int_\delta^\pi\right) \cos \nu t d\psi_x(t)$$
$$= I'_\nu + II''_\nu ,$$

say. Then

$$\left|\sum_{\nu=0}^{\infty} \tau_{n\nu} I'_\nu\right| \leqq \sum_{\nu=0}^{\infty} | \tau_{n\nu} | \int_0^\delta | d\psi_x(t) |$$
$$< \varepsilon \sum_{\nu=0}^{\infty} | \tau_{n\nu} |$$
$$\leqq K\varepsilon .$$

Remembering that $\psi_x(t)$ is absolutely continuous in (δ, π), we have

$$\int_\delta^\pi \cos \nu t d\psi_x(t) = \int_\delta^\pi \cos \nu t \psi'_x(t) dt .$$

For the given $\varepsilon > 0$, we can find ν_0 such that

$$\left|\int_\delta^\pi \cos \nu t \psi'_x(t) dt\right| < \varepsilon$$

for $\nu < \nu_0$ by Riemann-Lebesgue's theorem. Fixing ν_0, we can take a positive integer n_0 making $| \tau_{n\nu} | < \varepsilon/(\nu_0 + 1)$ $0 \leqq \nu \leqq \nu_0$, $n < n_0$. If we write

$$\sum_{\nu=0}^{\infty} \tau_{n\nu} I''_\nu = \left(\sum_{\nu=0}^{\nu_0} + \sum_{\nu_0+1}^{\infty}\right) \tau_{n\nu} \int_\delta^\pi \cos \nu t \psi'_x(t) dt$$
$$= I_1 + I_2 ,$$

say, then

$$| I_1 | \leqq \sum_{\nu=0}^{\nu_0} | \tau_{n\nu} | \int_\delta^\pi | \psi'_x(t) | dt$$
$$\leqq M \sum_{\nu=0}^{\nu_0} | \tau_{n\nu} |$$
$$< M(\nu_0 + 1)/(\nu_0 + 1)$$
$$= M\varepsilon .$$

for $n > n_0$, where

$$M = \int_0^\pi | \psi'_x(t) | dt .$$

$$| I_2 | = \left| \sum_{\nu=\nu_0+1}^{\infty} \tau_{n\nu} \int_\delta^\pi \cos \nu t \psi'_x(t) dt \right|$$
$$< \varepsilon \sum_{\nu=\nu_0+1}^{\infty} | \tau_{n\nu} |$$

$$\leqq \varepsilon \sum_{\nu=0}^{\infty} |\tau_{n\nu}|$$
$$\leqq K\varepsilon$$

by (iii). From the above analysis, it follows that

$$\left| \sum_{\nu=0}^{\infty} \tau_{n\nu} I_\nu \right| < (M + 2K)\varepsilon$$

for $n > n_0$. Since ε is an arbitrary quantity, we obtain $\sum \tau_{n\nu} I_\nu = 0(1)$ as $n \to \infty$. This proves the theorem.

REFERENCES

1. L. Fejér, *Über die Bestimmung des Sprunges der Funktion aus ihrer Fourierreihe,* J. reine angew Math., **142** (1913), 165-168.

2. M. Riesz, *Sur l'equivalence de certaine methodes de sommation,* Proc. London Math. Soc. (2), **22** (1924), 412-419.

3. J. A. Siddiqi, *On a theorem of Fejér,* Math. Zeitsch., **61** (1954-5), 79-81.

NATIONAL TAIWAN UNIVERSITY
TAIPEI, FORMOSA

ON N-HIGH SUBGROUPS OF ABELIAN GROUPS

JOHN M. IRWIN AND ELBERT A. WALKER

In a recent paper [2] the concept of high subgroups of Abelian groups was discussed. The purpose of this paper is to give further results concerning these high subgroups. All groups considered in this paper are Abelian, and our notation is essentially that of L. Fuchs in [1]. Let N be a subgroup of a group G. A subgroup H of G maximal with respect to disjointness from N will be called an *N-high* subgroup of G, or N-high in G. When $N = G^1$ (the subgroup of elements of infinite height in G), H will be called *high* in G.

After considering N-high subgroups in direct sums, we give a characterization (Theorem 3) of N-high subgroups of G in terms of a divisible hull of G. Next we show (Theorem 5) that if G is torsion, $N \subseteqq G^1$, and H is N-high in G, then H is pure and (Lemma 7) the primary components of any two N-high subgroups have the same Ulm invariants (see [3]). These results generalize the results in [2]. The concept of Σ-groups is introduced, and it is shown that any two high subgroups of torsion Σ-groups are isomorphic. Further, torsion Σ-groups are characterized in terms of their basic subgroups. Theorem 3 of [2] is generalized to show that high subgroups of arbitrary Abelian groups are pure. This leads to the solution of a more general version of Problem 4 of L. Fuchs in [1]. Finally, the question of whether any two high subgroups of a torsion group are isomorphic is considered, and a theorem in this direction is proved.

Preliminaries.

LEMMA 1. *Let M and N be subgroups of a primary group G such that M is neat in G and $M[p] \oplus N[p] = G[p]$. Then M is N-high in G.*

Proof. Suppose M is not N-high in G. Then there exists an N-high subgroup S of G properly containing M. Let $0 \neq s + M$ be in $(S/M)\ [p]$. By the neatness of S in G ([1], pg. 92) we may suppose that $s \in S[p]$. But this contradicts $M[p] \oplus N[p] = G[p]$, and so M is N-high in G.

As a consequence of Lemma 1, we obtain a standard

COROLLARY. ([3], pg. 24). *Let G be a primary group, and H a pure subgroup containing $G[p]$. Then $H = G$.*

Received August 10, 1960.

Proof. Purity implies neatness. Now put $N = 0$ in Lemma 1.

A useful generalization of Lemma 1 to torsion groups is

LEMMA 2. *Let M and N be subgroups of a torsion group G such that M is neat in G and $M[p] \oplus N[p] = G[p]$ for each relevant prime p. Then M is N-high in G.*

Proof. Use the proof of Lemma 1 with the observation that since S properly contains M, $(S/M)[p] \neq 0$ for some relevant prime p.

Concerning N-high subgroups in a direct sum, we have

THEOREM 1. *Let $G = \Sigma G_\alpha$ be an arbitrary direct sum of torsion groups, where H_α and N_α are subgroups of G_α, and where H_α is N_α-high in G_α for each α. Let $N = \Sigma N_\alpha$. Then $H = \Sigma H_\alpha$ is N-high in G.*

Proof. First notice that H is neat in G. To see this, use the neatness of H_α in G_α for each α (see [2] Lemma 10). Next observe that $G[p] = \Sigma G_\alpha[p] = \Sigma H_\alpha[p] \oplus \Sigma N_\alpha[p] = H[p] \oplus N[p]$ for each relevant prime p. Now apply Lemma 2 to get H to be N-high in G.

An interesting result concerning high subgroups (which are our main interest) in a direct sum is a corollary of Theorem 1.

THEOREM 2. *Let $G = \Sigma G_\alpha$ be an arbitrary direct sum of torsion groups where H_α is a high subgroup of G_α for each α. Then $H = \Sigma H_\alpha$ is high in G.*

Proof. By [2], Lemma 9 we have $G^1 = \Sigma(G_\alpha)^1$. Now use Theorem 1 and the definition of high subgroup.

Divisible hulls and high subgroups. Now we shall discuss the notion of a divisible hull for a group G, and the connection of such a hull with high subgroups. A group E minimal among those divisible groups containing G as a subgroup will be called a *divisible hull* of G. We need a few lemmas. The following lemma is almost obvious, and its proof is omitted.

LEMMA 3. *Let E be a divisible hull of a torsion group G. Let $E_1 = \Sigma E_p$ and $G = \Sigma G_p$. Then $E_1 = E$, and E_p is the unique divisible hull of G_p in E for each relevant prime p.*

LEMMA 4. *Let D be a divisible hull of a mixed group G, and E be a divisible hull of the torsion subgroup T of G in D. Then $D = E \oplus F$ where E is torsion and F is torsion free divisible.*

Proof. Since E is divisible, it is a direct summand of D. Thus $D = E \oplus F$ for some subgroup F of D. That E is torsion follows from Lemma 3. Next we show that the torsion subgroup T_F of F is zero. To see this, consider $T_F \cap G \subset T_F \cap T \subset E \cap F = 0$ to get $T_F \cap G = 0$. Then by Kulikov's Lemma, ([1], pg. 66) applied to D, $T_F = 0$ and F is torsion free. That F is divisible is clear, and the lemma is proved.

The following theorem gives a characterization in terms of divisible hulls of N-high subgroups of a torsion group G.

THEOREM 3. *Let N be any subgroup of a torsion group G, and E be a divisible hull of G with D a divisible hull of N in E. Then the set of N-high subgroups of G is the set of intersections of G with complementary summands of D in E.*

Proof. Let $H = A \cap G$, where $A \oplus D = E$. Now by [2], Lemma 1, and [1] pg. 67, we have for each relevant prime p,

$$G[p] = E[p] = A[p] \oplus D[p] = (A \cap G)[p] \oplus N[p] = H[p] \oplus N[p] .$$

By [1] pg. 92, H is neat in G, and finally by Lemma 2, H is N-high in G. Now for the converse, suppose H is N-high in G, so that $H \cap N = 0$. Now $H \cap D = 0$. To see this, notice that $(H \cap D) \cap N = H \cap N = 0$, and by Kulikov's lemma, $H \cap D = 0$. Since D is an absolute direct summand (see [1]), there exists A containing H with $A \oplus D = E$. But $H \subset A \cap G$, and since $(A \cap G) \cap N = 0$, by the maximality of H with respect to $H \cap N = 0$, we have $H = A \cap G$.

The reader will note that in particular Theorem 3 yields a characterization of high subgroups in torsion groups.

In general, a group G may have many high subgroups. It is even possible that $H \cap K = 0$ for two high subgroups H and K of G. The following theorem indicates the extent of the non-uniqueness of N-high subgroups.

THEOREM 4. *Let G be a primary group, let N be a subgroup such that $|N[p]| = |G|$ and such that $[G[p]: N[p] = |G|$. Then there exist $2^{|G|}$ distinct N-high subgroups of G. Furthermore, there exists an independent set $\{H_\alpha\}_{\alpha \in R}$ of N-high subgroups of G such that $|R| = |G|$.*

Proof. Let H be an N-high subgroup of G. By [2], $G[p] = H[p] \oplus N[p]$. Clearly $|H[p]| = |N[p]| = |G|$. Let $H[p] = \sum_{\alpha \in S} \langle x_\alpha \rangle$ and $N[p] = \sum_{\beta \in T} \langle y_\beta \rangle$. Then $|S| = |T| = |G|$. There exists $2^{|G|}$ one-to-one mappings of S onto T. Let f be such a mapping, and let

$$P_f = \sum_{\alpha \in S} \langle x_\alpha + y_{f(\alpha)} \rangle .$$

If g is any one-to-one mapping of S onto T such that $f \neq g$, then it is easy to see that $P_f \neq P_g$. Let H_f be an N-high subgroup of G containing P_f. Then $H_f[p] = P_f$, and since $P_f \neq P_g$, it follows that $H_f \neq H_g$. Hence there exist $2^{|G|}$ N-high subgroups of G. Let $T = \bigcup_{\beta \in R} T_\beta$, where $|T_\beta| = |T|$, $|R| = |T|$, and $T_\beta \cap T_\delta = \phi$ if $\beta \neq \delta$. That is, partition T into $|T|$ subsets each of cardinal $|T|$. Let f_β be a one-to-one mapping of S onto T_β, and let H_β be an N-high subgroup containing

$$P_\beta = \sum_{\alpha \in S} \langle x_\alpha + y_{f_\beta(\alpha)} \rangle .$$

It is straightforward to verify that $\{P_\beta\}_{\beta \in R}$ is an independent set of subgroups of G, that $H_\beta[p] = P_\beta$, and hence that $\{H_\beta\}_{\beta \in R}$ is an independent set of subgroups such that $|R| = |G|$. This concludes the proof.

It is easy to find examples of reduced primary groups G such that $|G^1[p]| = [G[p]: G^1[p]] = |G|$.

Purity of N-high subgroups of torsion groups. We now state and prove a generalization to N-high subgroups of torsion groups of [2] Theorem 3, namely that high subgroups of torsion groups are pure.

THEOREM 5. *Let N be a subgroup of a torsion group G with $N \subset G^1$, and let H be an N-high subgroup of G. Then H is pure in G.*

Proof. That it suffices to consider the primary case here follows from the fact H_p is N_p-high in G_p (see [2] Lemma 10 and [2] Lemma 11). So let G be primary. Now by [2] Lemma 1, we have $G[p] = H[p] \oplus N[p]$. Since $G^1[p] \supset N[p]$, then $G^1[p] = (H \cap G^1)[p] \oplus N[p]$. Now let H_1 be an $(H \cap G^1)[p]$-high subgroup of H. Since N-high subgroups are neat (see [1] c, d pg. 92) and neatness is transitive, we have that H_1 is neat in G. By [2] Lemma 1, $H[p] = H_1[p] \oplus (H \cap G^1)[p]$, so that $G[p] = H_1[p] \oplus G^1[p]$. An application of Lemma 1 yields H_1 high in G. Finally, by [2], Lemma 8, H_1 contains B basic in G, so that by [2] Lemma 2, H is pure in G as stated.

Before stating some corollaries, we would like to pose the following question: characterize all subgroups T of an Abelian group G such that T-high subgroups of G are pure. Suitable examples are easy to find which show that just any subgroup T will not do.

A couple of corollaries of Theorem 5 are

COROLLARY 1. *Let N_1 and N_2 be subgroups of a torsion group G with $N_1 \subseteqq N_2 \subseteqq G^1$. Then every N_1-high subgroup of G contains an N_2-high subgroup, and in particular every N_1-high subgroup K contains a subgroup H high in G.*

Proof. The proof is similar to the proof of Theorem 5.

COROLLARY 2. *Let N be a subgroup of G^1 in a torsion group G, and let S be an infinite subgroup of G with $S \cap N = 0$. Then there exists a subgroup K pure in G with $|K| = |S|$ and $K \cap N = 0$.*

Proof. Substitute N for G^1 in the proof of [2], Theorem 2.

This last corollary is a generalization of the solution in [2] of Fuchs' Problem 4.

Σ-groups. The following ideas arose from an investigation of the question of whether or not all high subgroups of a given group are isomorphic. A natural query in this direction is: If one of the high subgroups of a given group G is a direct sum of cyclic groups, are all of them direct sums of cyclic groups? The answer for torsion groups is yes. It is this observation that gives rise to so called Σ-groups. Before discussing this notion further, we need a few lemmas.

LEMMA 5. *Let N be a subgroup of a torsion group G with H and K both N-high subgroups of G. Then $((H \oplus N)/N)[p] = ((K \oplus N)/N)[p]$ for each relevant prime p.*

Proof. For $h \in H$ we have that $o(h + N) = p$ if and only if $o(h) = p$. Suppose $h \in H[p] \backslash (K \cap H)$. Then there exists $k \in K$, $n \in N$ with $h - k = n$, whence $o(k) = p$. Thus $h + N = (k + N) \in ((K \oplus N)/N)[p]$; and since p was arbitrary, we have by symmetry that

$$((H \oplus N)/N)[p] = ((K \oplus N)/N)[p]$$

as stated.

LEMMA 6. *Let N be a subgroup of a torsion group G with $N \subseteqq G^1$. Let H be an N-high subgroup of G. Then $((H \oplus N)/N)$ is pure in G/N.*

Proof. Suppose $m(g + N) = h + N$ for some $h \in H$, $g \in G$, m a nonzero integer. Then $mg - n = h$ for some $n \in N$, and since $n \in G^1$ and H is pure (Theorem 5), we have $h = mh_1$ for some $h_1 \in H$. Thus $h + N = m(h_1 + N)$ and the lemma is proved.

COROLLARY. *Let N be a subgroup of a reduced torsion group G with $N \subseteqq G^1 \neq 0$, and let H be an N-high subgroup of G. Then H is not closed.*

Proof. This follows easily from a theorem of Kulikov and Papp ([1] pg. 117).

LEMMA 7. *Let H and K be any two N-high subgroups of a primary group G with N a subgroup of G^1. Then for all positive integers n*

(a) *p^nH is N-high in p^nG,*

(b) *p^nH is pure in p^nG,*

(c) $(p^n((H \oplus N)/N))[p] = (p^n((K \oplus N)/N))[p]$,

(d) *H, K, and G have the same* nth *Ulm invariants* (see [3]).

Proof. (a) Use the proof of Theorem 5 (e) in [2] and the fact that $N \subseteqq p^nG$ for all n.

(b) Use (a), $N \subseteqq (p^nG)^1$, and Theorem 5.

(c) First notice that $p^n((H \oplus N)/N) = (p^nH \oplus N)/N$. Now (c) follows immediately from Lemma 5 applied to the right sides of this equation and the corresponding one for K.

(d) The proof is similar to that of Theorem 6 in [2].

The following theorem is a slight generalization of the fact that any two high subgroups of a countable group G are isomorphic. (See [2], Theorem 5 (u).)

THEOREM 6. *Let N be a subgroup of a countable torsion group G with $N \subseteqq G^1$, and G^1 elementary. Then any two N-high subgroups H and K of G are isomorphic.*

Proof. Write $G = \Sigma G_p$, $H = \Sigma H_p$, $K = \Sigma K_p$. Now H_p and K_p are both N_p high in G_p, and $N_p \subseteqq G_p^1$ for each relevant prime p. Let $\bar{H}_p = (H_p \oplus N_p)/N_p$ and $\bar{K}_p = (K_p \oplus N_p)/N_p$. Using Lemmas 5 and 6 and the fact that G^1 is elementary, we get immediately that $(p^\omega \bar{H}_p)[p] = (p^\omega \bar{K}_p)[p]$, and that $(p^{\omega+1}\bar{H}_p)[p] = (p^{\omega+1}\bar{K}_p)[p] = 0$. Thus for $\alpha \geqq \omega$, the αth Ulm invariants of $\bar{H}_p$ and $\bar{K}_p$ are the same. Since $\bar{H}_p \cong H_p$ and $\bar{K}_p \cong K_p$, Lemma 7 (d) implies that H_p and K_p have the same Ulm invariants. Since G^1 is elementary, H and K are reduced, and Ulm's theorem yields $H \cong K$.

REMARK. In Theorem 6, if we take N to be a subgroup of G^1 such that $N[p] \neq G^1[p]$, then neither H nor K will be a direct sum of cyclic groups as is easily seen.

A *Σ-group* is any group G all of whose high subgroups are direct sums of cyclic groups. This means that in a torsion Σ-group every high subgroup is basic. This implies further that in Σ-groups, every high subgroup is an endomorphic image. Examples of Σ-groups are very to easy to find. For instance, direct sums of countable groups turn out to be Σ-groups. Also, any group G such that G/G^1 is a direct sum of cyclic groups is a Σ-group. (See the proof of Theorem 7.) For a non-Σ-group, see [2], Theorem 5(t).

THEOREM 7. *Let H and K be high subgroups of a torsion group G. Then if H is a direct sum of cyclic groups, so is K. Moreover $H \cong K$.*

Proof. Let $\tilde{S}$ be the image of S under the natural homomorphism of G onto G/G^1. Now $\tilde{H} \cong H$, $\tilde{K} \cong K$, and by Lemma 5 we have $\tilde{K}[p] = \tilde{H}[p]$ for each relevant prime p. By [3], Theorem 12, we have that $\tilde{H}[p]$ is the union of a sequence $\tilde{P}_n$ of subgroups of bounded height in $\tilde{H}$. The purity of $\tilde{H}$ (Theorem 5) tells us that $\tilde{P}_n$ has bounded height in $\tilde{K}$ for each n. Hence by [3], Theorem 12, each $\tilde{K}_p$ is a direct sum of cyclic groups, so that $\tilde{K}$ is a direct sum of cyclic groups. Thus K, which is isomorphic to $\tilde{K}$, is a direct sum of cyclic groups. Since H and K are both basic in G, we have $H \cong K$. Thus we have shown that in a torsion group, if one high subgroup is a direct sum of cyclic groups, they all are, and they are all isomorphic.

From Theorem 7 we see that if in a torsion group G there exist two non-isomorphic high subgroups, then no high subgroup is a direct sum of cyclic groups. The next theorem shows that torsion Σ-groups are closed under direct sums.

THEOREM 8. *For torsion groups it is true that a direct sum of Σ-groups is a Σ-group.*

Proof. By Theorem 2, a direct sum of highs is high. But such a direct sum is basic. An application of Theorem 7 completes the proof.

COROLLARY. *A direct sum of countable torsion group is a Σ-group.*

Proof. It suffices by Theorem 8 to verify that a countable torsion group is a Σ-group, and this is very easy.

REMARK. Examples exist of torsion groups G such that G/G^1 is a direct sum of cyclic groups, but such that G is not a direct sum of countable groups. Therefore, we see that the class of torsion Σ-groups properly contains the class of all torsion groups that are the direct sum of countable groups.

The next theorem gives an interesting characterization of torsion Σ-groups.

THEOREM 9. *A torsion group G is a Σ-group if and only if G*

contains a maximal basic subgroup.

Proof. If G is a Σ-group, then any high subgroup will be a maximal basic subgroup of G. Now suppose G contains a maximal basic subgroup B. Let H be a high subgroup containing B, and suppose $B \neq H$. By [1] pg. 114, there exists B_1 basic in H with $B_1 > B$. Since H is pure and G/H is divisible, B_1 is basic in G, a contradiction. Therefore $B=H$, and G is a Σ-group by Theorem 7.

The next theorem is a result concerning the Σ-groups of a torsion group.

THEOREM 10. *Every torsion group G contains a Σ-subgroup R pure in G such that $R^1 = G^1$.*

Proof. First if $G^1 = 0$, put $R = B$ basic in G. Also if G is a Σ-group, but $R = G$. So suppose that $G^1 \neq 0$ and G is not a Σ-group. Let B be a basic subgroup of G. Embed B in a high subgroup H of G. By Theorem 8 and the assumptions on G, $H/B \neq 0$. B is basic in H so that $G/B = H/B \oplus R/B$, where the divisibility of H/B guarantees that R/B may be chosen to contain $(G^1 \oplus B)/B$. Hence R contains G^1. The purity of R/B in G/B gives us that R is pure in G. Hence $R^1 = G^1$. Now $H \cap R = B$, so that $G[p] = H[p] \oplus G^1[p]$ and

$$R[p] = (R \cap H)[p] \oplus G^1[p] = B[p] \oplus R^1[p] .$$

By Lemma 2, B is high in R so that by Theorem 7, R is a Σ-group, and the proof is complete.

We do not know whether every subgroup of a Σ-group is a Σ-group. However, every pure subgroup of a torsion Σ-group is a Σ-group. In fact, we have

THEOREM 11. *Every subgroup L of a torsion Σ-group G with $L^1 = L \cap G^1$ is a Σ-group.*

Proof. Embed a high subgroup H_L of L in a high subgroup H G. Since G is a Σ-group, H is a direct sum of cyclic groups and hence so is H_L. Now apply Theorem 7 to L to get that L is a Σ-group.

COROLLARY. *Every pure subgroup R of a torsion Σ-group G is a Σ-group.*

Proof. $R^1 = R \cap G^1$, and Theorem 11 then yields the desired result.

COROLLARY. *Every pure subgroup of a direct sum of countable torsion groups is a Σ-group.*

High subgroups in mixed groups. In this section we will discuss some properties of high subgroups of arbitrary Abelian groups and generalize some of our results for the torsion case. A lemma which is useful is

LEMMA 8. *Let S be any subgroup of an Abelian group G with $S \cap G^1 = 0$. Then for any subgroup T with $(T/S) \cap (G/S)^1 = 0$, we have $T \cap G^1 = 0$.*

Proof. Suppose $T \cap G^1 \neq 0$. Then $(T/S) \cap (G/S)^1 \neq 0$.

The following theorem was proved in [2] for the torsion case. The fact that G^1 is divisible in the torsion free case makes this case easy, so we proceed directly to the general case.

THEOREM 12. *Let H and K be any two high subgroups of a group G. Then*

(a) *G/H is divisible*

(b) *G/H is a divisible hull of $(G^1 \oplus H)/H \cong G^1$*

(c) *$G/H \cong G/K$.*

Proof. (a) Let T/H be the torsion subgroup of G/H. Now T/H is divisible, for if not, T/H would have a non-zero cyclic direct summand L/H. But L/H would be a direct summand of G/H since T/H is pure in G/H. Hence $(L/H) \cap (G/H)^1 = 0$, and Lemma 8 gives us that $L \cap G^1 = 0$. Consequently H is not high in G, a contradiction. Thus we have $G/H = T/H \oplus F/H$, where F/H is torsion free. This means that $(F/H)^1$ is divisible, whence $F/H = (F/H)^1 \oplus R/H$. Now clearly, $(R/H) \cap (G/H)^1 = 0$, so that by Lemma 8, $R = H$, and G/H is divisible as stated.

(b) As a divisible group, G/H must contain a divisible hull D/H of $(G^1 \oplus H)/H$. Put $G/H = D/H \oplus L/H$. Clearly $L \cap G^1 = 0$, hence $L/H = 0$ and (b) is proved.

(c) This follows from $(G^1 \oplus H)/H \cong G^1 \cong (G^1 \oplus K)/K$ and the fact that divisible hulls of isomorphic groups are isomorphic. Thus we see that G/H is a structural invariant of G.

We shall now discuss a generalization to arbirary Abelian groups of a theorem proved in [2] for the torsion case. Here again, the torsion free case is easy (G^1 is divisible), and for a torsion free group G we see easily that all high subgroups are isomorphic. First we need

THEOREM 13. *Let T be the torsion subgroup of an Abelian group G, H be a high subgroup of G, and T_H be the torsion subgroup of H. Then T_H is high in T.*

Proof. We need only consider the case of a mixed group G. Let E be a divisible hull of G with D_H and D_{G^1} divisible hulls in E of H and G^1. Then $E = D_H \oplus D_{G^1}$ (see [2]). Next let D_{T_H} and $D_{T_{G^1}}$ be divisible hulls of T_H and T_{G^1} in D_H and D_{G^1} respectively. Then

$$E = D_H \oplus D_{G^1} = D_{T_H} \oplus D_1 \oplus D_{T_{G^1}} \oplus D_2 = D_{T_H} \oplus D_{T_{G^1}} \oplus D_1 \oplus D_2 \,.$$

Applying Lemmas 3 and 4 to D_H and D_{G^1}, we have that $T_E = D_{T_H} \oplus T_{T_{G^1}}$ is the torsion subgroup of E, and $D_1 \oplus D_2$ is torsion free. Clearly (Lemma 3) T_E is a divisible hull of T in E. Now by Theorem 3, it remains to verify that $T \cap D_{T_H} = T_H$. To this end put $L = T \cap D_{T_H}$. That $T_H \subset L$ is clear. By Lemma 4, $H \cap D_{T_H} = T_H$. So suppose there exists $t \in (G\backslash H) \cap L$. Then by the definition of H, there exists $h \in H$ with $h + mt = g_1 \neq 0$, where $g_1 \in G^1$. But since L is torsion we have that $h \in T_H$. Hence $(h + mt) \in D_{T_H}$, and $h + mt = g_1 \neq 0$ together with $D_{T_H} \subset D_H$ contradict $D_H \cap D_{G^1} = 0$, and T_H is high in T as desired.

COROLLARY. *Let H be a high subgroup of G, and let T_H be the torsion subgroup of H. Then T_H is pure in G.*

Proof. By Theorem 13, T_H is high in T, and consequently pure in T. Since T is pure in G, it follows that T_H is pure in G.

THEOREM 14. *Let H be a high subgroup of an Abelian group G. Then H is pure in G.*

Proof. Let the notation be the same as in Theorem 13. Now by Theorem 13 T/T_H is divisible, so that $G/T_H = T/T_H \oplus R/T_H$, where R is chosen such that R/T_H contains H/T_H. Since T_H is pure in G, R is pure in G, and since H is neat in G, H/T_H is neat in R/T_H. But R/T_H is torsion free and since a neat subgroup of a torsion free group is pure we have that H/T_H is pure in R/T_H. Thus H is pure in R, so that H is pure in G, and the proof is complete.

The following embedding theorem is a generalization to arbitrary Abelian groups of the solution to Fuchs' Problem 4 (see [1]).

THEOREM 15. *Let S be any infinite subgroup of an Abelian group G with $S \cap G^1 = 0$. Then there exists a subgroup K pure in G with $S \subset K$, $K \cap G^1 = 0$, and $|K| = |S|$.*

Proof. Embed S in H high in G. By [1] pg. 78 N, there exists a pure subgroup K of H with $S \subset K$ and $|S| = |K|$. The purity of

H implies the purity of K in G, and $K \cap G^1 \subset H \cap G^1 = 0$, so that $K \cap G^1 = 0$, completing the proof.

An unsolved problem. To conclude the present paper we shall make a few remarks concerning the question of whether all high subgroups of an Abelian torsion group are isomorphic. The reader may have observed, from the proof of Theorem 6, that this question is a special case of the more general open question: Given two pure subgroups A and B of a primary group G with $A[p] = B[p]$, is it true in general that $A \cong B$? The authors feel that an affirmative answer to this question would have important consequences in the theory of Abelian torsion groups. A step in this direction is

THEOREM 16. *Let A and B be pure subgroups of primary group G with $A[p] = B[p]$. Then $G = A \oplus C$ implies $G = B \oplus C$ and $A \cong B$.*

Proof. Let $G = A \oplus C$. Then $G[p] = A[p] \oplus C[p] = B[p] \oplus C[p]$. We will show that $G = B \oplus C$. First notice that $A[p] = B[p]$ gives us that $B \cap C = 0$. To prove $G = B \oplus C$, it suffices to verify that $G \subset B \oplus C$. For this purpose it is sufficient that $G[p^n] \subset B \oplus C$ for each n. But this is true if and only if $A[p^n] \subset B \oplus C$ for each n. Now we use induction to show that $G[p^n] \subset B \oplus C$ for each n. First, $G[p] \subset B \oplus C$ by hypothesis. Next suppose that $G[p^n] \subset B \oplus C$, and let $a \in A$ with $o(a) = p^{n+1}$. Then $p^n a = b \in A[p] = B[p]$. By the purity of B, $p^n a = p^n b_1$ with $b_1 \in B$, and $p^n(a - b_1) = 0$, so that $a - b_1 \in B \oplus C$ by the induction hypothesis. Hence $a \in B \oplus C$, therefore

$$A[p^{n+1}] \subset B \oplus C ,$$

which means that $G[p^{n+1}] \subset B \oplus C$. Thus $G = B \oplus C$. Finally $A \cong G/C \cong B$, and the proof is complete.

The foregoing theorem suggests the following generalization.

THEOREM 17. *Let G be a direct sum of torsion groups, $G = \sum_{\alpha \in A} G_\alpha$, and let $\{T_\alpha\}_{\alpha \in A}$ be a family of subgroups pure in G with $T_\alpha[p] = G_\alpha[p]$ for each relevant prime p and each $\alpha \in A$. Then for any subfamily $\{T_\alpha\}_{\alpha \in S}$, $G = \sum_{\alpha \in S} T_\alpha \oplus \sum_{\alpha \notin S} G_\alpha$. In particular, $G = \sum_{\alpha \in A} T_\alpha$ and $G_\alpha \cong T_\alpha$ for each $\alpha \in A$.*

Proof. Put $T = \sum_{\alpha \in A} T_\alpha$. It suffices to show that $G = T$. We show as before that for each n we have $G[p^n] \subset T$. This is true if for each $\alpha \in A$ we have for the primary components G_{α_p}, that $G_{\alpha_p}[p^n] \subset T$ for each n. This is accomplished as in the proof of Theorem 16. Finally, that $T_\alpha \cong G_\alpha$ for each α follows as before.

REFERENCES

1. L. Fuchs, *Abelian Groups*, Budapest, 1958.
2. John M. Irwin, *High subgroups of Abelian torsion groups*, Pacific J. Math., **11**, (1961).
3. Irving Kaplansky, *Infinite Abelian Groups*, Ann Arbor, 1954.

NEW MEXICO STATE UNIVERSITY (UNIVERSITY PARK, N.M.)

HIGH SUBGROUPS OF ABELIAN TORSION GROUPS

JOHN M. IRWIN

The results in this paper were part of a doctor's thesis completed in February 1960 under Professor W. R. Scott at the University of Kansas. The author wishes to express his gratitude to Professor Scott for his advice and for checking the results.

In what follows, all groups considered are Abelian. Let G^1 be the subgroup of elements of infinite height in an Abelian group G (see [2]). A subgroup H of G maximal with respect to disjointness from G^1 will be called a *high* subgroup of G. If N is a subgroup of G, H will be called *N-high* if and only if H is a subgroup of G maximal with respect to disjointness from N. Zorn's lemma guarantees the existence of N-high subgroups for any subgroup N of G. A group E minimal divisible among those groups containing G will be called a *divisible hull* of G. Unless otherwise specified, the notation and terminology will be essentially that of L. Fuchs in [1].

The main theorem says that high subgroups of Abelian torsion groups are pure. After proving some preparatory lemmas, we will prove the main theorem. Then we will discuss Fuchs' Problem 4 and list some of the more important properties of high subgroups. Finally we will state some generalizations.

A lemma describing N-high subgroups is

LEMMA 1. *Let G be a primary group with H an N-high subgroup of G, D a divisible hull of G, A any divisible hull of H in D* (this means that $A \subset D$), *and B any divisible hull of N in D.*

Then

(a) $D = A \oplus B$.

(b) $A \cap G = H$, and H and $B \cap G$ are neat in G.

(c) Any complementary direct summand of A in D containing N is a divisible hull of N.

(d) Any complementary direct summand of B in D containing H is a divisible hull of H.

(e) D is a divisible hull of any subgroup M with $(H \oplus N)[p] \subset M \subset G$.

(f) $D[p] = (H \oplus N)[p] = H[p] \oplus N[p] = G[p]$.

(g) All N-high subgroups H of G may be obtained as $E \cap G$, where E is a complementary direct summand of a divisible hull F of N in D.

Proof. When $N = 0$ there is nothing to prove, so suppose $N \neq 0$.

Received October 31, 1960.

(a) To see this, we first show that $A \cap B = 0$. If $0 \neq x \in A \cap B$, then by Kulikov's lemma ([1], p.66) there exist positive integers r and s such that $p^r x \neq 0 \neq p^s x$ and $p^r x \in H$, $p^s x \in N$. But $0 \neq p^{\max(r,s)} x \in H \cap N = 0$, which is impossible. The divisibility of $A \oplus B$ provides a decomposition $D = A \oplus B \oplus C$. If $C \cap G \neq 0$, H will not be N-high in G, whence $C \cap G = 0$. By Kulikov's lemma, $C = 0$, and we have $D = A \oplus B$ as stated.

(b) That $A \cap G = H$ is clear. The neatness follows from [1], p. 92, h,

(c) and (d) follow from (a) and the definition of H,

(e) and (f) follow from Kulikov's lemma,

(g) follows from (a) and (b). This concludes the proof of Lemma 1.

In what follows $\langle x \rangle$ will denote the cyclic subgroup of G generated by $x \in G$. An interesting and helpful lemma is

LEMMA 2. *Let N be a subgroup of a primary group G, H an N-high subgroup of G, and let H contain a basic subgroup B of G. Then H is pure in G.*

Proof. The group G/B is divisible since B is basic in G. Now $H/B \subset G/B$, and by [1], p. 66, Theorem 20.2, there exists a divisible hull E/B of H/B in G/B. Suppose $E/B > H/B$. Then $E > H$, and hence $E \cap N \neq 0$. Thus there exists a nonzero element $g \in N$ with $0 \neq \langle g + B \rangle \subset E/B$. Now $\langle g + B \rangle \cap (H/B) = 0$. To see this, suppose $0 \neq m(g + B) = mg + B = h + B$. Then $mg - h = b \in B$, and $0 \neq mg = h + b \in H$, contradicting $H \cap N = 0$. Thus we have $\langle g + B \rangle \cap (H/B) = 0$. By Kulikov's lemma, $\langle g + B \rangle = 0$, and therefore $g \in B$, which implies that $g = 0$, contrary to the choice of g. Thus $E/B = H/B$ is divisible, and therefore is pure in G/B. Then the purity of B in G together with [1], p. 78, M imply that H is pure in G.

A useful lemma with a standard proof is

LEMMA 3. *If $G = S \oplus T$, where each element of S has finite height, then $G^1 \subset T$ and $T^1 = G^1$.*

A lemma which displays an inheritance property is

LEMMA 4. *If $G = S \oplus T$, where $S \subset H$ and H is high* in G, *then $H = S \oplus H \cap T$, and $H \cap T$ is high in T.* (Note: This implies that $H \cap T$ is maximal with respect to disjointness from G^1 in T by Lemma 3.)

Proof. Put $M = H \cap T$, and suppose that there exists $0 \neq t \in T \backslash H$, with $\{M, t\} \cap T^1 = 0$. But this means that $\{M, t\} \cap G^1 = 0$, and hence $[S \oplus \{M, t\}] \cap G^1 = 0$; for otherwise we would have $s + (m + kt) = g \neq 0$ with $s \in S, m \in M, g \in G^1$. Then $s = 0$ and $m + kt = g \neq 0$. But $m + kt \in \{M, t\}$, which is not possible. Thus $[S \oplus \{M, t\}] \cap G^1 = 0$ and

$[S \oplus \{M, t\}] > H$, contrary to the assumption that H is high in G.

A lemma on making new basic subgroups out of old ones is

LEMMA 5. *Let $B_1 \oplus \cdots \oplus B_n \oplus G_n = G$, where $B = \Sigma B_n$ is basic in G. Let $T_n = B'_{n+1} \oplus B'_{n+2} \oplus \cdots$ be basic in G_n. Then $C = B_1 \oplus \cdots \oplus B_n \oplus T_n$ is basic in G.*

Proof [1], p. 109, Exercise 9a.

An $x \in G$ will be called a *pure element* of G if and only if $\langle x \rangle$ is a pure subgroup (and therefore is a direct summand) of G.

The next lemma is the kingpin in the proof that if H is high in G, then H contains B basic in G. It is not altogether obvious that H contains nonzero pure subgroups of G. The proof of the next lemma will be carried out in several steps. We will consider special cases which are perhaps unnecessary, but which will help to clarify the method of proof.

LEMMA 6. *Let H be a high subgroup of a primary group G. If G contains nonzero pure elements of order p^n, but not of smaller order, then H contains pure elements of G of order p^n.*

Proof.

Case 1. $n = 1$. Let $b \in G$ be pure of order p with $b \notin H$. Then there exists $h \in H$ such that $h + b = g \neq 0$, where $g \in G^1$. Clearly this means that the orders of h and b are the same. Now h and b both have finite height, and hence their heights must be equal (since their sum is an element of infinite height in G). Here we are making use of the fact that if $o(h) = p$ and $h(h) = 0$, then $\langle h \rangle$ is pure in G. The fact that b is a pure element of order p in G necessarily means that $h(b) = 0$; whence $h(h) = 0$, and h is a pure element of G.

Case 2. $n > 1$. Let b be a pure element of G of order p^n such that $b \notin H$. Then there exists an $h \in H$ such that $h + p^j b = g \neq 0$, where $g \in G^1$ and $0 \leqq j < n$.

Case 2.1. $j = 0$. Then we have $h + b = g$ and $p^{n-1}h + p^{n-1}b = p^{n-1}g \in G^1$. Clearly $p^{n-1}h$ has order p and height $n-1$ in G and in $\langle h \rangle$. Thus by [1], p. 78, J, we have that h is a pure element of G.

Case 2.2. $1 \leqq j < n$. Now the equation $h + p^j b = g \neq 0$ clearly implies that the height of h in G is j. If the height of h in H were also j, that is if $h = p^j h'$ for some $h' \in H$, then h' would be a pure element of G of order p^n. To see this simply consider the equation $p^{n-1}h' + p^{n-1}b = p^{n-j-1}g \in G^1$. $o(p^{n-1}h') = p$ and obviously $h(p^{n-1}h') = n-1$. The height of $p^{n-1}h'$ in $\langle h' \rangle = n-1$, so that by [1], p. 78, J, h' is a pure element of G of order p^n in H. Thus it remains to verify that the height of h in H is j.

From the neatness of H and the fact that $g \in G^1$, it follows that $h = ph_1$ for some $h_1 \in H$. Now if $h(h_1) > 0$, we again have by the neatness of H that $h_1 = ph_2$ for some $h_2 \in H$. Continuing in this way, we must eventually arrive at $h = p^{j-k}h_k$, $h_k \in H$ where the height of h_k in G is 0. If $k > 0$, then let m be the least positive integer such that $p^m h_k = p^{m+1}z$ for $z \in G$ (if worst comes to worst $m = j-k$ will do). Then clearly $0 < m \leqq j-k < j < n$, $p^m(h_k - pz) = 0$, and $p^{m-1}(h_k - pz) \neq 0$ has height $m-1$ in G by the choice of m and h_k. Thus since $o(p^{m-1}(h_k - pz)) = p$, and the height of $p^{m-1}(h_k - pz)$ is $m-1$ in $\langle h_k - pz \rangle$, we have by [1], p. 78, J, that $h_k - pz$ is a pure element of G of order $p^m < p^n$. This contradicts the hypotheses on G. Hence we must have $k = 0$, $h = p^j h_k$, and h_k is a pure element of G in H of order p^n.

If $B = \Sigma B_n$ is a basic subgroup of G where B_n is a direct sum of cyclic groups of order p^n, then such a subgroup B_n which does not consist of 0 alone will be referred to as a B_n of G.

LEMMA 7. *Let G be a primary group, H a high subgroup of G, and n the least positive integer such that G contains a B_n. Then H contains a B_n of G.*

Proof. By Lemma 6, H contains pure elements of G of order p^n. The fact that the union of an ascending chain of pure subgroups is pure together with [1], p. 80, Theorem 24.5 allows us to apply Zorn's lemma to obtain a p^n-bounded direct summand H_n of G, maximal with respect to the property of being contained in H. We wish to show that H_n is a B_n for G. To see this write $G = H_n \oplus R_n$ and $H = H_n \oplus H \cap R_n$ where by Lemma 4, $H \cap R_n$ is high in R_n. Suppose that H_n is not a maximal p^n-bounded direct summand (a B_n) of G. Then there exists a B_n of G with $H_n < B_n$. Now $G = H_n \oplus R_n$, so that $B_n = H_n \oplus B_n \cap R_n$. Now the transitivity of purity tells us that $B_n \cap R_n \neq 0$ is pure in G. Thus R_n contains pure elements of order p^n since G contains no pure elements of order less than p^n. This means by Lemma 6 that $H \cap R_n$ as a high subgroup of R_n must contain a pure element h of order p^n. Then $G = H_n \oplus R_n = H \oplus \langle h \rangle \oplus R_n'$, and $(H_n \oplus \langle h \rangle) > H_n$ implies that H_n is not a maximal p^n-bounded direct summand of G contained in H, contrary to the choice of H_n. This means that H_n is a B_n of G contained in H after all, and this concludes the proof.

LEMMA 8. *Let G be a primary group, and let H be a high sub-*

group of G. Then H contains a basic subgroup of G.

Proof. By a theorem of Baer ([1], p.62), it suffices to consider the reduced case. Lemma 7 provides a start for the induction. Let B'_{n_1} be a first B_n of G. By lemma 7, H contains a B_{n_1} and $G = B_{n_1} \oplus R_1$ with $H = B_{n_1} \oplus H \cap R_1$. Let B_{n_2} be the next B_j of G. By Szele's theorem ([1], p. 99) and Lemma 5, R_1 contains a B_{n_2} but no preceding B_j. We apply Lemma 7 to $H \cap R_1$ as a high subgroup of R_1 to see that $H \cap R_1$ contains a B_{n_2} of G. By successive application of this procedure, we have by induction, Szele's theorem, Lemmas 5 and 7 that H contains a basic subgroup of G.

We are now ready to state and prove our main theorem.

THEOREM 1. *Let G be a primary group and H a high subgroup of G. Then H is pure in G.*

Proof. As in the proof of Lemma 8, it suffices to consider the case where G is reduced. Lemmas 2 and 8 complete the proof.

In his book [1], L. Fuchs asks the following question: "Let G be a primary group and H an infinite subgroup without elements of infinite height. Under what conditions can H be imbedded in a pure subgroup of the same power and again without elements of infinite height?" Theorem 1 allows us to give the best possible solution to this problem.

THEOREM 2. *Let G be an Abelian primary group. If S is any infinite subgroup of G with $S \cap G^1 = 0$, then S can be embedded in a pure subgroup K of G so that $K \cap G^1 = 0$ and $|K| = |S|$.*

Proof. By Zorn's lemma, there exists a subgroup H high in G with $H \supset S$. By Theorem 1, H is pure in G. Szele has shown that every infinite subgroup can be embedded in a pure subgroup of the same power ([1], p. 78). So let K be a pure subgroup of H containing S and of the same power as S. Then by the transitivity of purity, we have that K is pure in G. Since $K \subset H$, it follows that $K \cap G^1 = 0$. This concludes the proof.

The following discussion yields the solution to Fuchs' question in the torsion case. The proofs of the next two lemmas are standard and consequently will be omitted.

LEMMA 9. *Let G be a torsion group. If $G = \Sigma G_\alpha$, then $G^1 = \Sigma G^1_\alpha$.*

LEMMA 10. *Let G be a torsion group. Then an internal direct sum of pure subgroups of the direct summands of a given direct decomposition of G is a pure subgroup of G.*

Concerning the primary decomposition of a torsion group G, we have, LEMMA 11. *If H is a high subgroup of a torsion group G, then writing G and H in terms of their primary components $G = \Sigma G_p$ and $H = \Sigma H_p = \Sigma H \cap G_p$, we have that H_p is a high subgroup of G_p for each relevant prime p in the primary decomposition of G.*

Proof. Clearly $H_p \cap G_p^1 = 0$. So suppose for some p, H_p is not high in G_p. Then there exists an $x \in G_p \backslash H_p$ with $\{H_p, x\} \cap G_p^1 = 0$. Replacing H_p by $S_p = \{H_p, x\}$ in $H = \Sigma H_p$, we obtain from Lemma 9 a subgroup $S > H$ with $S \cap G^1 = 0$. But this is contrary to H high in G.

A generalization of Theorem 1 is

THEOREM 3. *If H is a high subgroup of a torsion group G, then H is pure in G.*

Proof. Write $G = \Sigma G_p$ and $H = \Sigma H_p$ and by Lemma 11, we have that H_p is high in G_p so that by Theorem 1 we have H_p is pure in G_p. Now by Lemma 10, H is pure in G.

The generalization of the solution to Fuchs' question to torsion groups is

THEOREM 4. *Any infinite subgroup S of a torsion group G with $S \cap G^1 = 0$ can be embedded in a pure subgroup K of G so that $|K| = |S|$ and $K \cap G^1 = 0$.*

Proof. Use Theorem 3 and the proof of Theorem 2.

We mention for completeness that Lemma 8 has a suitable generalization to torsion groups.

LEMMA. 12. *Let G be a torsion group and let H be a high subgroup of G. Then H contains a basic subgroup of G.*

Proof. Use Lemma 8, the primary decomposition of H, and [**1**], p. 109, Exercise 9a.

Some of the more interesting properties of high subgroups are contained in

THEOREM 5. *Let G be a reduced primary group with $G^1 \neq 0$, and let H and K be high subgroups of G. Then*

(a) *H contains B basic in G*

(b) *H is pure in G*

(c) *G/H is a divisible hull of $(G^1 \oplus H)/H \cong G^1$*

(d) *$G/K \cong G/H$*

(e) *p^nH is high in p^nG for all $n \in I$* (I is the set of positive integers.)

(f) *p^nH is pure in p^nG for all $n \in I$*

(g) *$G = \{H, p^nG\}$ for all $n \in I$*

(h) *H is infinite*

(i) *H is of unbounded height in G*

(j) *$p^nG = \{p^nH, p^{n+k}G\}$ for all $n, k \in I$*

(k) *$p^nH/p^{n+k}H \cong p^nG/p^{n+k}G$ for all $n, k \in I$.*

(l) *$p^nG/p^nH \cong G/H$ for all $n \in I$.*

(m) *G is minimal pure containing $H \oplus G^1$*

(n) *$H \oplus G^1 < G$*

(o) *$|H \oplus G^1| = |G|$*

(p) *$|G^1| < |G|$ implies $|H| = |G|$*

(q) $|K| = |H|$ (This also holds for N-high subgroups of infinite rank.)

(r) $|G| \leqq |H|^{\aleph_0}$

(s) *$G/p^nH = H/p^nH \oplus p^nG/p^nH$ for all $n \in I$.*

(t) *H is not always basic*

(u) *If H is countable, then H is basic in G, and $H \cong K$.*

Proof. (a) and (b) have already been proved.

(c) Is easy.

(d) Follows from (c) and the fact that isomorphic groups have isomorphic divisible hulls (see [1], p. 66, Theorem 20.2).

(e) To see that p^nH is high in p^nG, suppose that there exists $x \in G$ with $\{p^nH, p^nx\} \cap G^1 = 0$ and $p^nx \notin p^nH$. (Here we are using the fact that $(p^nG)^1 = G^1$.). Now by purity of H, $p^nx \notin p^nH$ implies $p^nx \notin H$. Thus we have some $h \in H$ with $h + mp^nx = g \neq 0, g \in G^1$. But then h must be in p^nH contrary to $\{p^nH, p^nx\} \cap G^1 = 0$.

(f) The purity of p^nH in p^nG follows from (e), and Theorem 1 applied to p^nG.

(g) This is an immedite consequence of (c).

(h) And (i) both follow from (g) and the fact that a high subgroup of a reduced group is not a direct summand.

(j) Follows from (e) and (g).

(k) Follows from (j), the second isomorphism theorem, and (f).

(l) Is an immediate consequence of the fact that both quotient groups are divisible hulls of G^1. This is also a straightforward application of (g).

(m) This follows from Lemma 1 (f) and [1], p. 78, *K*.

(n) Follows from the fact that (c) holds and hence H is not a direct summand of G.

(o) Follows from Lemma 1 (f) and an easy set theoretic argument.

(p) Is an easy consequence of (o).

(q) Here some cases are taken care of by (d), but a proof for the general case is not difficult. To show that $|H| = |K|$, it suffices (by an easy set theoretic argument) to verify that $H[p] \cong K[p]$. For this purpose let D be a divisible hull of G, and C be a divisible hull of G^1 in D. By Lemma 1, if A and B are divisible hulls in D of H and K respectively, then A and B are complementary direct summands of C in D. Finally $A \cong D/C \cong B$ and $H[p] = A[p] \cong B[p] = K[p]$. The same argument shows the result for N-high subgroups of infinite rank.

(r) Follows trivially from (a) and [1], p. 102, Theorem 30.1.

(s) To see this, use (g) and the purity of H.

(t) Let G be the direct sum of an unbounded closed primary group and any primary group with nonzero elements of infinite height.

(u) This follows from (b), (c), (q), the fact that a countable H is a direct sum of cyclic groups, and that any two basic subgroups of G are isomorphic.

For a comparison with the properties of basic subgroups see [1], p. 101. The reader will notice that (d) is an interesting property of high subgroups which basic subgroups do not possess.

We are now ready to discuss the question of whether or not any two high subgroups of a reduced primary group are isomorphic. Let A be a subgroup of G, and let $\tilde{A}$ be the image under the natural homomorphism from G onto G/G^1. It is a simple matter to verify that $\tilde{G}$ is a reduced primary group without elements of infinite height. Thus if H is a high subgroup of G, we have that $H \cong \tilde{H}$. This provides us a natural way to study the properties of high subgroups without actually looking at these subgroups themselves.

A result concerning Ulm invariants as defined by Kaplansky in [2], and providing another proof that two countable high subgroups of a group G are isomorphic is the following

THEOREM 6. *Let H and K be high subgroups of a primary group G. Then $(p^n\tilde{H})[p]/(p^{n+1}\tilde{H})[p] = (p^n\tilde{K})[p]/(p^{n+1}\tilde{K})[p]$. In particular, H and K have the same Ulm invariants. Moreover, their nth Ulm invariants are the same as the nth Ulm invariant of G.*

Proof. Consider $\tilde{H}$ and $\tilde{K}$. First we notice

(i) $\tilde{H}[p] = \tilde{K}[p]$.

To see this we observe that $o(h) = o(h + G^1)$. Suppose $h \in H[p]\backslash H \cap K$. Then there exists $k \in K$ with $h - k = g \neq 0$ where $g \in G^1$. Clearly $o(k) = p$ and we have $h = k + g$. This proves that $\tilde{H}[p] \subset \tilde{K}[p]$. Thus by symmetery $\tilde{H}[p] = \tilde{K}[p]$. Next we have

(ii) $p^n\tilde{H}[p] = p^n\tilde{K}[p]$ for $n \in I$.

To see this use Theorem 5 (e), and the foregoing (i).

Now from (ii) we have that $(p^n\tilde{H})[p]\backslash(p^{n+1}\tilde{H})[p] = (p^n\tilde{K})[p]/(p^{n+1}\tilde{K})[p]$ since the numerators are equal and the denominators are equal, and hence the Ulm invariants of $\tilde{H}$ and $\tilde{K}$ are equal. Finally the fact that

$H \cong \tilde{H}$ gives us that H and K have the same nth Ulm invariants. The last part of the theorem follows from $(p^nG)[p]/(p^{m+1}G)[p] \cong (p^n\tilde{H})[p]/(p^{n+1}\tilde{H})[p]$ which is obtained with the help of Lemma 1 (f), Theorem 5 (e), and the second isomorphism theorem.

We will now mention a few generalizations to modules. In what follows, R will denote a principal ideal ring. This means that R is an integral domain (commutative ring with an identity and no divisors of zero) in which every ideal is principal. By an R-module we mean a unitary left R-module, and by submodule of an R-module we mean a sub-R-module. An R-module M is called *primary* if and only if the order ideal of every element of M is generated by a power of the same prime element p of R. We shall be content with a generalization to primary modules of our main results for primary groups. We rely heavily on the generalizations of Theorems 1 to 14 in [2].

We make a blanket assertion: All of our lemmas and theorems for primary groups are true for primary modules. Only minor, straightforward modifications of the definitions and proofs are necessary, and these can be easily carried out by imitating all the previous definitions and proofs. When referring to orders of elements in a primary module, we say that $o(x)$ is smaller than $o(y)$ if and only if the generator of the order ideal of x divides the generator of the order ideal of y.

In conclusion we state without proof the most worthwhile lemmas and theorems.

LEMMA 13. *Let M be a primary R-module. Let L, N be submodules of M with L containing a basic submodule B of M, and L maximal with respect to disjointness from N. Then L is pure in M.*

THEOREM 7. *Let H be a high submodule of a primary R-module M. Then H is pure in M.*

The solution of Fuchs' question for primary modules is

THEOREM 8. *Let S be an infinitely generated submodule of the primary R-module M with R countable and $S \cap M^1 = 0$. Then S can be embedded in a pure submodule K of M such that $K \cap M^1 = 0$ and $|K| = |S|$.*

The only essential difference between this theorem and Theorem 2 is that the word infinite has been replaced by the words infinitely generated to make $|K| = |S|$ true in all cases. The proof is the same as before. The countability assumption on R makes the proof of [1] p. 78 N easy.

The author conjectures that all high subgroups of a given primary group are isomorphic, and also wishes to pose the questions:

For what subgroups N of a primary group G is it true that
(a) all N-high subgroups are pure
(b) all N-high subgroups are isomorphic
(c) all N-high subgroups are endomorphic images of G
(d) G/N divisible implies N contains B basic in G?

References

1. L. Fuchs. *Abelian Groups.* Publishing House of The Academy of Sciences, Budapest, 1958.

2. Irving Kaplansky. *Infinite Abelian Groups.* University of Michigan Press, Ann Arbor, 1954.

NEW MEXICO STATE UNIVERSITY AND THE UNIVERSITY OF KANSAS

QUOTIENT RINGS OF RINGS WITH ZERO SINGULAR IDEAL

R. E. JOHNSON

Many papers have been written recently (see [2]–[14] of bibliography) on extensions of rings to rings of quotients. In most of these papers, strong enough conditions are imposed on the given rings to insure that each has a vanishing singular ideal (first defined in [5]). It seems appropriate at this time to collect these results and present them in as general a form as possible. In this paper, it is assumed that each ring has a zero right singular ideal. A subsequent paper will give the quotient structure of a ring having a vanishing right and left singular ideal.

1. **Introduction.** If R is a ring and M is an R-module, then $L(R)$ and $L(M, R)$ will designate the lattices of right ideal of R and R-submodules of M, respectively. Superscripts "r" and "l" will be used in designating the right and left annihilators, respectively, of an element or subset of a ring or module. The context will always make it clear from what set the annihilators are to be chosen.

In a lattice L with 0 and I, an element B is called an *essential extension* of element A, and we write $A \subset' B$, if and only if $A \subset B$ and $C \cap A \neq 0$ for every C in L for which $C \cap B \neq 0$. An element A of L is called *large* if $A \subset' I$. The sublattice of L of all large elements is designated by $L^\blacktriangle$.

If R is a ring and M is a right R-module, then let

$$M^\blacktriangle(R) = \{x \mid x \in M, x^r \in L^\blacktriangle(R)\}, \quad R^\blacktriangle = \{x \mid x \in R, x^r \in L^\blacktriangle(R)\}.$$

It is easily shown that $M^\blacktriangle(R)$ is a submodule of M and $R^\blacktriangle$ is a (two-sided) ideal of R. The ideal $R^\blacktriangle$ is called the *singular ideal* [5; p. 894] of R.

A ring R with zero singular ideal has the unusual property, proved in [7; Section 6], that each $A \in L(R)$ has a unique maximal essential extension A^s in $L(R)$. The mapping $s: A \to A^s$ of $L(R)$ is shown there to be a closure operation on $L(R)$ having the following properties:

(1) $0^s = 0$,

(2) $(A \cap B)^s = A^s \cap B^s$ for each $A, B \in L(R)$, and

(3) $(x^{-1}A)^s = x^{-1}A^s$ for each $x \in R$ and $A \in L(R)$, where $x^{-1}B = \{y \mid y \in R, xy \in B\}$. The set $L^s(R)$ of closed right ideals (i.e., $A = A^s$) may be made into a lattice in the usual way by defining the union of a set of

Received December 10, 1960. This research was supported in part by a grant from the National Science Foundation.

elements of $L^s(R)$ to be the least upper bound of the set. The resulting lattice $L^s(R)$, which is not in general a sublattice of $L(R)$, is proved *to be a complete complemented modular lattice in* [7; *Section 6*]. *If* M is a right R-module for which $M^\blacktriangle(R) = 0$, then the closure operation s may be defined in a similar way on $L(M, R)$. The resulting lattice $L^s(M, R)$ has similar properties to those of $L^s(R)$, as was shown in [7; Section 6].

For $A, B \in L(R)$, B is called a *complement* of A if $B \cap A = 0$ whereas $C \cap A \neq 0$ for every $C \supset B$, $C \neq B$. If B is a complement of A, then clearly $A + B \in L^\blacktriangle(R)$. Furthermore, if $R^\blacktriangle = 0$, then $B \in L^s(R)$.

If A is a two-sided ideal of R for which $A \cap A^l = 0$, then evidently A^l is the unique complement of A in $L(R)$. Since $(A + A^l)^l = A^l \cap A^{ll}$, clearly A^{ll} is the unique complement of A^l in case $R^\blacktriangle = 0$. In this case, both A^l and A^{ll} are in $L^s(R)$. By [7; 6.7], $C^s(R) = \{A \mid A$ ideal of $R, A \cap A^l = 0, A = A^{ll}\}$ is the *center of the* lattice $L^s(R)$. For each $A \in C^s(R)$, it is easily seen that $A^\blacktriangle = 0$, that $L^s(A) = \{B \cap A \mid B \in L^s(R)\}$, and that $C^s(A) = \{B \cap A \mid B \in C^s(R)\}$. Of course, $L^s(A) \subset L^s(R)$ and $C^s(A) \subset C^s(R)$.

Every regular ring R has a zero singular ideal. This is evident because $e^r \cap eR = 0$ for each idempotent $e \in R$. Since $R = eR + e^r$, evidently eR and e^r are complements of each other and each is in $L^s(R)$. Consequently, each principal right ideal $aR \in L^s(R)$.

A ring R for which $R^\blacktriangle = 0$ and $C^s(R) = \{0, R\}$ is called (right) *irreducible*. An irreducible ring need not be prime. For example, the ring of all $n \times n$ triangular matrices over the ring Z of integers is irreducible by [8; 3.5]. Clearly this ring has a nonzero nilpotent ideal. By [8; 2.1], an irreducible ring is prime if and only if it contains no nonzero nilpotent ideal.

If R is a subring of ring Q then Q is called a (right) *quotient ring* of R, and write $R \leqq Q$, if and only if $qR \cap R \neq 0$ each nonzero $q \in Q$. It was proved in [5] that each ring R for which $R^\blacktriangle = 0$ has a unique maximal quotient ring $\hat{R}$. By [5; Theorem 2], $\hat{R}$ is a regular ring with unity. Essentially, the definition of $\hat{R}$ in [5] was as follows:

$$\hat{R} = \bigcup_{A \in L^\blacktriangle(R)} \mathrm{Hom}_R(A, R) .$$

If $x, y \in \hat{R}$, then we take $x = y$ if and only if $xa = ya$ for every a in some large right ideal $A \subset \mathrm{Dom}\, x \cap \mathrm{Dom}\, y$.

In case R is a subring of a ring Q, then we may consider Q as a right R-module. If we do so, then the assumption $R \leqq Q$ implies that $R \subset' Q$, considering R and Q as right R-modules. It is easily verified

The more general definition of a quotient ring in [**12**] and [**2**] is equivalent to ours in case $R^\blacktriangle = 0$.

that if $R \leqq Q$ then $Q^{\blacktriangle}(R)=0$ if and only if $R^{\blacktriangle}=0$.

2. Some basic lemmas. The rest of this paper will be concerned only with a ring R for which $R^{\blacktriangle}=0$. We shall prove in this section that if Q is a quotient ring of such a ring R, then the lattices of closed right ideals of R and Q are isomorphic.

2.1 LEMMA. *If $R \leqq Q$ and $A \in L(Q)$, then $A \in L^{\blacktriangle}(Q)$ if and only if $A \cap R \in L^{\blacktriangle}(R)$.*

Proof. If $A \in L^{\blacktriangle}(Q)$ and $b \in R, b \neq 0$, then $A \cap bQ \neq 0$ and $a = bq \neq 0$ for some $a \in A$ and $q \in Q$. Now $qC \subset R$ for some $C \in L^{\blacktriangle}(R)$ by [7; 6.1]. Since $Q^{\blacktriangle}(R)=0$, $bqC \neq 0$ and therefore $A \cap bR \neq 0$. Hence $(A \cap R) \cap bR \neq 0$ and $A \cap R \in L^{\blacktriangle}(R)$.

On the other hand, let us assume that $A \in L(Q)$ and $A \cap R \in L^{\blacktriangle}(R)$. For each nonzero $q \in Q, qC \subset R$ for some $C \in L^{\blacktriangle}(R)$. If we let $B = C \cap (A \cap R)$, then $B \in L^{\blacktriangle}(R)$ and $qB \neq 0$ since $Q^{\blacktriangle}(R)=0$. Hence $qB \cap (A \cap R) \neq 0$ and we conclude that $qQ \subset A \neq 0$ for each nonzero $q \in Q$. Thus, $A \in L^{\blacktriangle}(Q)$.

2.2 LEMMA. *If $R \leqq Q$ and M is a right Q-module, then M is a right R-module and $M^{\blacktriangle}(R) = M^{\blacktriangle}(Q)$.*

Proof. If $x \in M$ and $A = x^r(\text{in } Q)$ then $A \in L^{\blacktriangle}(Q)$ if and only if $A \cap R \in L^{\blacktriangle}(R)$ by 2.1. Therefore, $M^{\blacktriangle}(R) = M^{\blacktriangle}(Q)$.

2.3 COROLLARY. *If $R \leqq Q$, then $Q^{\blacktriangle}=0$.*

This follows from 2.2 if we let $M = Q$ and use the assumption that $R^{\blacktriangle}=0$.

2.4 LEMMA. *If $R \leqq Q$ and M is a right Q-module such that $M^{\blacktriangle}(Q)=0$, then $L^s(M, R) = L^s(M, Q)$.*

Proof. If $A \in L^s(M, R)$ and $q \in Q$, then $qB \subset R$ for some $B \in L^{\blacktriangle}(R)$. Therefore $(Aq)B \subset A$ and $Aq \subset A$ by [7; 6.4]. Hence, $A \in L(M, Q)$ and we conclude that $L^s(M, R) \subset L(M, Q)$.

If $A \in L(M, Q)$, $x \in M$ and $B_x = \{b \mid b \in Q, xb \in A\}$, then $x \in A^s$ if and only if $B_x \in L^{\blacktriangle}(Q)$ by [7; 6.4]. Therefore, in view of 2.1, the closure of A relative to Q is the same as its closure relative to R. Thus, $L^s(M, R) = L^s(M, Q)$.

2.5 THEOREM. *If $R \leqq Q$, if M is a right Q-module for which $M^{\blacktriangle}(Q)=0$ and if $N \in L^{\blacktriangle}(M, R)$, the $L^s(M, Q) \cong L^s(N, R)$ under the* *correspondence* $A \to A \cap N$, $A \in L^s(M, Q)$.

Proof. By [7; 6.8], $L^s(M, R) \cong L^s(N, R)$. Thus 2.5 follows from 2.4.

2.6 COROLLARY. *If* $R \leqq Q$, *then* $L^s(Q) \cong L^s(R)$ *under the correspondence* $A \to A \cap R$, $A \in L^s(Q)$.

If R is an irreducible ring, so that $C^s(R) = \{0, R\}$, then $C^s(\hat{R}) = \{0, \hat{R}\}$ by 2.6. Hence $\hat{R}$ also is irreducible. Actually, since $\hat{R}$ is regular, $\hat{R}$ is a prime ring by [8; 2.1]. We state this result as follows.

2.7 THEOREM. *If R is an irreducible ring, then $\hat{R}$ is a prime ring.*

3. **$L^s(R)$ atomic.** Let us assume in this section that R is a ring for which $R^\blacktriangle = 0$ and the lattice $L^s(R)$ is *atomic*. We define this to mean that $L^s(R)$ has minimal nonzero elements, called atoms, and that each element of $L^s(R)$ contains at least one atom. It is proved in [7; 6.9] that a nonzero element x of R is contained in an atom if and only if x^r is a maximal element of $L^s(R)$. Incidentally, $(xR)^s$ is the atom containing x.

Two atoms A and B are said to be *perspective* [1; p. 118], and we write $A \sim B$, if and only if A and B have a common complement. It is easily shown in our case that $A \sim B$ if and only if $A \cup B$ contains a third atom [1; p. 120, Lemma 3]. We proved in [7; 6.10] that $A \sim B$ if and only if $a^r = b^r$ for some nonzero $a \in A$ and $b \in B$. If $A \sim B$ and $B \sim C$ then $a^r = b^r$ and $b_1^r = c^r$ for some nonzero $a \in A, b, b_1 \in B$ and $c \in C$. Since B is an atom, $bR \cap b_1R \neq 0$ and there exist $x, x_1 \in R$ such that $bx = b_1x_1 \neq 0$. Hence, $(ax)^r = (bx)^r = (b_1x_1)^r = (cx_1)^r$. It follows that perspectivity is an equivalence relation on the set of all atoms of $L^s(R)$. Clearly for a finite set $\{A_1, \cdots, A_n\}$ of perspective atoms, there exist nonzero $a_i \in A_i$ such that $a_i^r = a_j^r$ for each i and j.

For each atom A of $L^s(R)$, let A^* be the union in $L^s(R)$ of all atoms perspective to A. It is proved in [7] that A^* is an ideal of R [7; 6.7] and that A^* is an atom of $C^s(R)$ [7; 6.12]. Conversely, each atom of $C^s(R)$ is of the form A^* for some atom A of $L^s(R)$.

Since $C^s(R)$ is a Boolean algebra, R is the direct union of all atoms of $C^s(R)$. Hence, if $\{A_i^*; i \in \Delta\}$ is the set of all distinct atoms of $C^s(R)$, then the ring-union S of the atoms of $C^s(R)$ is a discrete direct sum of these atoms,

$$S = \sum_{i \in \Delta} A_i^* .$$

Since $S^l = 0$, evidently $S \leqq R$. Consequently, the maximal quotient ring of R is just the maximal quotient ring of S.

The following theorem characterizes $\hat{R}$ in terms of left full rings. We shall call a ring R a *left full ring* if there exists a division ring D and a right D-module M such that

$$R \cong \mathrm{Hom}_D(M, M) .$$

Evidently we may consider M as a (R, D)-module.

3.1 THEOREM. *If R is a right irreducible ring, then $\hat{R}$ is a left full ring. If R is right reducible, then $\hat{R}$ is a complete direct sum of left full rings.*

Proof. Consider first the case in which R is irreducible. Since $\hat{R}$ is regular and $L^s(R) \cong L^s(\hat{R})$, the lattice $L^s(\hat{R})$ is atomic and its atoms are principal and hence minimal right ideals of $\hat{R}$. Since $\hat{R}$ is prime and has minimal right ideals, it is primitive. Let e be an idempotent element of $\hat{R}$ such that $e\hat{R}$ is a minimal right ideal. Then $M = \hat{R}e$ is a minimal left ideal of $\hat{R}$ and $D = e\hat{R}e$ is a division ring. Since $x\hat{R}e \neq 0$ for each nonzero $x \in \hat{R}$ by the primeness of $\hat{R}$, evidently $\hat{R}$ is a right quotient ring of M. However, $\hat{R}$ is a maximal right quotient ring so that we must have $\hat{M} = \hat{R}$. Besides being a ring, M may be considered to be a $(\hat{R}, D)$-module. Clearly the right ideals of M are its D-submodules. Thus, M is the only large right ideal of M. Consequently,

$$\mathrm{Hom}_M(M, M) ,$$

considering M as a right M-module, is the maximal right quotient ring of M. Since $x(ae) = x(eae)$ for each $x \in M$ and $a \in \hat{R}$, evidently

$$\mathrm{Hom}_M(M, M) = \mathrm{Hom}_D(M, M) .$$

Since $\hat{M} = \hat{R}$, this proves that $\hat{R}$ is a left full ring.

If R is not irreducible, then there exists a set $\{R_i; i \in \Delta\}$ of irreducible rings, each having an atomic lattice of closed right ideals, such that

$$\sum_{i \in \Delta} R_i \leqq R$$

by our previous results. We shall not give the details, but it is easily seen that if

$$S = \sum_{i \in \Delta} R_i , \quad \text{then } \hat{S} = \sum_{i \in \Delta}{}' \hat{R}_i$$

where $\sum'$ designates the complete direct sum. Since $\hat{S} = \hat{R}$, this proves the second part of 3.1.

The important special case of this theorem when R is a primitive ring was proved by Utumi [12; 5.1] and Wong [13; 4.1]. Both Utumi and Lambek [10] have independently proved the theorem if R is prime.

4. $L_s(R)$ **finite-dimensional.** The usual assumption that $R^{\blacktriangle} = 0$ is made for each ring R of this section. If either the a.c.c. or the d.c.c. holds for $L^s(R)$ then so does the other one. In fact, each is equivalent to the assumption that $L^s(R)$ contains a maximal chain of finite length. When this condition is satisfied, a *dimension function* d may be defined on $L^s(R)$ as follows [1; p. 67]: for each $A \in L^s(R)$, $d(A)$ is the length of the longest chain joining 0 to A. Incidentally, every maximal chain joining 0 to A has the same length $d(A)$. We shall assume in this section that such a dimension function d is defined on $L^s(R)$ and that $d(R)$ is finite. Since the lattice $L^s(R)$ is also complemented, each $A \in L^s(R)$ is a direct union of $d(A)$ atoms [1; p. 105].

It is proved in [9; 3.4] that if $d(R)$ is finite then for each $a \in R$, $aR \in L^{\blacktriangle}(R)$ if and only if $a^r = 0$. Of course, $a^l = 0$ whenever $aR \in L^{\blacktriangle}(R)$. Thus, $D(R) = \{a \mid a \in R, aR \in L^{\blacktriangle}(R)\}$ is the set of *regular* elements of R. Each $a \in D(R)$ has an inverse in $\hat{R}$. For, by the regularity or $\hat{R}$, $(ab-1)a = a(ba-1) = 0$ for some $b \in \hat{R}$. Since $(ab-1)^r \supset aR$, a large element of $L^{\blacktriangle}(R)$, $ab - 1 = 0$ in view of 2.1 and 2.3. Also, $ba - 1 = 0$ since $a^r = 0$ in $\hat{R}$ as well as in R. Consequently, $b = a^{-1}$.

4.1 THEOREM. *If R is irreducible and $d(R) = n$, then $\hat{R}$ is a full ring of dimension n.*

By a full ring of dimension n we mean a ring isomorphic to $\mathrm{Hom}_D(M, M)$ where D is a division ring and M is a right D-module of dimension n.

If we select $M = \hat{R}e$ as in the proof of 3.1, then $M \leqq \hat{R}$ and the lattices $L^s(R)$, $L^s(M)$ and $L^s(\hat{R})$ are isomorphic by 2.6. Since the right ideals of M are its D-submodules, M is an n-dimensional vector space over D. Hence 4.1 follows from 3.1.

A different proof of 4.1 was given in [9; 3.6].

If R is a prime ring for which $d(R)$ is finite, then it was proved in [3; Theorem 10] and in [9; 3.5] that every large right ideal of R contains a regular element. Since $B = \{b \mid b \in R, qb \in R\}$ is a large right ideal of R for each $q \in \hat{R}$, clearly $qb = a$ for some $b \in D(R)$ and $a \in R$; that is, $q = ab^{-1}$. This proves the following theorem of Goldie[2] [3] (also proved in [11] and [9]).

[2] That each ring considered by Goldie has a zero singular ideal is proved in [4; 3.2].

4.2 THEOREM. *If R is a prime ring for which $d(R) = n$, then not only is $\hat{R}$ the full ring of linear transformations of an n-dimensional vector space over a division ring but also* $R=\{ab^{-1} \mid a \in R, b \in D(R)\}$.

From 3.1 and 4.1, we easily deduce the following theorem.

4.3 THEOREM. *If R is a ring for which $d(R)$ is finite, then $\hat{R}$ is a direct sum of a finite number of finite-dimensional full rings.*

A ring R is called *semiprime* if it contains no nonzero nilpotent ideal. We recall that if S is the direct sum of the atoms of $C^s(R)$, then $S \leqq R$. Since each nonzero ideal of R has nonzero intersection with some atom of $C^s(R)$, evidently R is semiprime if and only if each atom of $C^s(R)$ is prime. The following theorem was recently proved by Goldie [4].

4.4 THEOREM. *If R is a semiprime ring for which $d(R)$ is finite, then not only is $\hat{R}$ a direct sum of a finite number of finite-dimensional full rings but also* $R = \{ab^{-1} \mid a \in R, b \in D(R)\}$.

The first part of 4.4 follows directly from 4.3. To prove the second part, let $S = R_1 \oplus \cdots \oplus R_k$ be the sum of the atoms of $C^s(R)$. Then $\hat{R} = \hat{S} = \hat{R}_1 \oplus \cdots \oplus \hat{R}_k$. If $q_i \in \hat{R}$, then $q_i = a_i b_i^{-1}$ for some $a_i \in R_i$ and $b_i \in D(R_i)$ by 4.2. Thus, if $q = q_1 + \cdots + q_k$, $a = a_1 + \cdots + a_k$, and $b = b_1 + \cdots + b_k$, $q = a\, b^{-1}$. This proves the second part of 4.4.

A converse of 4.4 has been given by Goldie [5; 4.4]. He proved that if R is a ring for which $d(\hat{R})$ is finite and $\hat{R} = \{ab^{-1} \mid a \in R, b \in D(R)\}$, then R is semiprime. Naturally, this implies the following converse of 4.2: If R is a ring for which $\hat{R}$ is a finite-dimensional full ring and $\hat{R} = \{ab^{-1} \mid a \in R, b \in D(R)\}$ then R is prime.

BIBLIOGRAPHY

1. G. Birkhoff, *Lattice theory*, Amer. Math. Soc. Colloquium Publications, rev. ed. vol. 25. New York, 1948.
2. G. D. Findlay and J. Lambek. *A generalized ring of quotient*, I *and* II, Can. Math, Bull., **1** (1958), 77-85, 155-167.
3. A. W. Goldie, *The structure of prime rings under ascending chain conditions*, Proc. London Math. Soc., **8** (1958), 589-608.
4. ———, *Semi-prime rings with maximum condition*, ibid **10** (1960), 201-220.
5. R. E. Johnson, *The extended centralizer of a ring over a module*, Proc. Amer. Math. Soc., **2** (1951), 891-895.
6. ———, *Structure theory of faithful rings I. Closure operations on lattices*, Trans. Amer. Math. Soc., **84** (1957), 508-522.
7. ———, II. *Restricted rings*, ibid, p.p. 523-544.
8. ———, III. *Irreducible rings*, Proc. Amer. Math. Soc., **11** (1960), 710-717.

9. R. E. Johnson and E. T. Wong, *Quasi-injective modules and irreducible rings*, J. London Math. Soc., **36** (1961), 260-268.

10. J. Lambek, *On the structure of semi-prime rings and theirs of quotients*, Can. J. Math. **13**(1961), 392-417.

11. L. Lesieur and R. Croisot, *Sur les anneaux premiers noethériéns á gauche*, Ann. Sci. Ec. Norm. Sup., **76** (1959), 161-183.

12. Y. Utumi, *On quotient rings*, Osaka Math. J., **8** (1956), 1-18.

13. E. T. Wong, *Quotient rings*, Ph. D. Thesis, U. of Rochester (1956).

14. E. T. Wong and R. E. Johnson, *Self-injective rings*, Can. Math. Bull., **2** (1959), 167-173.

UNIVERSITY OF ROCHESTER

THE ASYMPTOTIC DISTRIBUTION OF THE TIME-TO-ESCAPE FOR COMETS STRONGLY BOUND TO THE SOLAR SYSTEM

DAVID G. KENDALL AND J. L. MOTT.

1. Introduction This paper is one of a series (Hammersley and Lyttleton [1], [2], Kerr [3], Kendall [4], [5], [6]) concerned with the statistical-dynamical properties of the sun's family of comets. For the astronomical background, terminology, conventions, units, etc., we refer the reader to [5].

We consider a comet in the energy-state $x > 0$ (so that the total energy per unit mass is equal to $-x$) which is approaching perihelion, not necessarily for the first time, and we write T for the total time spent by the comet in describing complete circuits subsequent to this perihelion. We ignore the low energy (high x) catastrophes (capture by Jupiter, falling into the sun, etc.) and consider the fate of the comet subject to independent energy-perturbations at perihelion, the magnitudes of which we suppose to be distributed according to the probability law

$$\tfrac{1}{2}e^{-|w|/b}dw/b \qquad (-\infty < w < \infty),$$

the so-called 'double-exponential law'. It is then known [5] that T is almost certainly finite.

The probability distribution of T cannot be found explicitly, but its Laplace-Stieltjes transform ϕ satisfies a differential equation which we treat by a perturbation method. At first sight it seems unlikely that a perturbation procedure followed by a Laplace inversion could yield any positive information about the distribution being studied, but in fact by a careful arrangement of the argument we are able to calculate the exact limit-law

$$\lim_{x\to\infty} \Pr\left\{\frac{T}{\sqrt{x}} \leqq c \,|\, x\right\}$$

for the reduced random variable $T/\sqrt{x}$; the result is given at (15) below.

If we are chiefly interested in the origin of comets we can identify the given perihelion with the comet's first, and x is then its initial energy-state. There are indications ([5], [6]) that this value of x is small when compared with the average size of the perturbations, but information about solutions for large x can be extracted from Hammersley [2], and the present result thus forms a useful complement to some of his results, with which it is consistent: in fact, the same (limit-) law was obtained by Hammersley in his exact solution to the corresponding problem involv-

Received November 16, 1960.

ing Brownian motion.

If we do not identify the given perihelion with the comet's first, then our result tells us the distribution of the remaining time-to-escape T for a comet which happens to have entered a high energy-state. From this point of view the result is of value whatever opinion we may hold about the origin of comets, but it is of course limited by the fact that when x is large (i.e., the comet is strongly bound) then one cannot properly neglect the low-energy (high x) catastrophes.

The justification for the use of the double-exponential perturbation law will be found in [3], [4], [5], [6]. Because of the asymptotic character of the present result one might expect the detailed form of the perturbation law to be unimportant, and one might hope that identically the same result would follow for any perturbation with zero mean and a finite variance. The proof that this is so is the object of an investigation by C. Stone and J. Lamperti, who will in a forthcoming paper discuss the appropriate invariance theorem.

In the course of our work we shall make use of some Bessel function formulae given by Watson [7]. We shall refer to these formulae as (1W), $\cdots$, (5W), where (1W) is given on p. 80 of [7] at (19), (2W)-77(2), (3W)-202(1), (4W)-203(2) and (5W)-80(15).

2. The asymptotic distribution of $T/\sqrt{x}$. T is the total time during which a comet remains in the system, measured from (say first) perihelion; thus T is the total time spent in describing complete circuits. The comet is subject to energy perturbations at perihelion distributed according to the double-exponential law, and there is also a chance k $(0 \leqq k < 1)$ of disintegration at each perihelion passage; for the moment we retain the possibility of disintegration but our main results depend on a method which would not be very easy to handle when $k > 0$, and we shall shortly put $k = 0$.

Define

$$\phi(s \mid x) \equiv \mathscr{E}(e^{-sT} \mid x);$$

here x is the energy-state of the comet during the approach to first perihelion, so that $x > 0$, and $s \geqq 0$. $V(y) = y^{-3/2}$ $(y > 0)$ gives the periodic time of an orbit in state y, but (following Hammersley [2]) we shall first set $V(y) = y^{-\alpha}$. We shall later put $\alpha = 3/2$ to give our main result, and afterwards remark briefly on the more general case.

Consideration of the possible events at first perihelion leads to the integral equation for ϕ:

$$\phi(s \mid x) = k + (1-k)\left\{\tfrac{1}{2}\, e^{-x/b} + \int_0^\infty \tfrac{1}{2}\, b^{-1}\, e^{-w/b}\, e^{-sV(x+w)}\, \phi(s \mid x+w)\, dw \right.$$
$$\left. + \int_0^x \tfrac{1}{2}\, b^{-1}\, e^{-w/b}\, e^{-sV(x-w)}\, \phi(s \mid x-w)\, dw \right\},$$

whence

$$(1)\quad \phi(s\,|\,x) = k + (1-k)\left\{\tfrac{1}{2}\, e^{-x/b} + \tfrac{1}{2}\, b^{-1}\, e^{x/b} \int_x^\infty e^{-w/b}\, e^{-sV(w)}\, \phi(s\,|\,w)\, dw \right.$$
$$\left. + \tfrac{1}{2}\, b^{-1}\, e^{-x/b} \int_0^x e^{w/b}\, e^{-sV(w)}\, \phi(s\,|\,w)\, dw \right\}.$$

Since $0 \leqq \phi \leqq 1$ for $x > 0$ and $s \geqq 0$, and $\phi(s\,|\,\cdot)$ is measurable for $s \geqq 0$, we see from (1) that $\phi(s\,|\,\cdot)$ is continuous on the interval $0 < x < \infty$. But then $\phi(s\,|\,\cdot)$ is also differentiable, and with $D \equiv \partial/\partial x$,

$$(2)\quad (1-k)^{-1}\, D\phi = -\tfrac{1}{2}\, b^{-1}\, e^{-x/b} + \tfrac{1}{2}\, b^{-2}\, e^{x/b} \int_x^\infty e^{-w/b}\, e^{-sV(w)}\, \phi(s\,|\,w)\, dw$$
$$- \tfrac{1}{2}\, b^{-2}\, e^{-x/b} \int_0^x e^{w/b}\, e^{-sV(w)}\, \phi(s\,|\,w)\, dw;$$

also

$$(3)\quad (1-k)^{-1}\, D^2\phi = \tfrac{1}{2}\, b^{-2}\, e^{-x/b} + \tfrac{1}{2}\, b^{-3}\, e^{x/b} \int_x^\infty e^{-w/b}\, e^{-sV(w)}\, \phi(s\,|\,w)\, dw$$
$$- b^{-2}\, e^{-sV(x)}\, \phi(s\,|\,x) + \tfrac{1}{2}\, b^{-3}\, e^{-x/b} \int_0^x e^{w/b}\, e^{-sV(w)}\, \phi(s\,|\,w)\, dw.$$

Thus from (1) and (3)

$$(4)\quad D^2\phi = b^{-2}\{1 - (1-k)\, e^{-sV(x)}\}\, \phi(s\,|\,x) - kb^{-2} \qquad (x > 0);$$

and now we have that, in fact, $\phi \in C^\infty$.

We now put $k = 0$, and then write (4) as

$$(5)\quad D^2\phi - g\phi = f\phi,$$

where

$$g = sb^{-2}\, V(x) \quad \text{and} \quad f = \{1 - sV(x) - e^{-sV(x)}\}\, b^{-2}.$$

We discuss the nature of ϕ by using the standard method of variation of parameters, and so postulate as a solution of (5) (for the moment we suppress the variable s)

$$\phi(x) = A(x)\, \theta_1(x) + B(x)\, \theta_2(x),$$

where $\theta_1(x)$, $\theta_2(x)$ are independent solutions of

$$(6)\quad D^2\phi - g\phi = 0.$$

We find that

$$(7)\quad \phi(x) = \theta_1(x) \int_c^x f(y)\, \phi(y)\, \theta_2(y)\, dy / W - \theta_2(x) \int_d^x f(y)\, \phi(y)\, \theta_1(y)\, dy / W,$$

where $W = \theta_1'(y)\, \theta_2(y) - \theta_2'(y)\, \theta_1(y)$ (actually a nonzero constant) and c, d

here are constants. When $\alpha = 3/2$, which is the case we shall work through in detail, we can take (cf. Watson [7], p. 96)

$$(8) \qquad \theta_1(x) = x^{1/2} K_2(4b^{-1} s^{1/2} x^{1/4}) \quad \text{and} \quad \theta_2(x) = x^{1/2} I_2(4b^{-1} s^{1/2} x^{1/4}) ;$$

also $W = -\frac{1}{4}$, using ($1W$). We now rewrite (7) in the more convenient form:

$$(9) \qquad \phi(x) = A\theta_1(x) + B\theta_2(x) - 4\theta_1(x)\int_{0+}^{x} f(y)\phi(y)\,\theta_2(y)\,dy - 4\theta_2(x)\int_{x}^{\infty} f(y)\,\phi(y)\,\theta_1(y)\,dy ,$$

where A and B are constants (possibly involving s) to be found later; this we can do because, for fixed s, $\theta_2(y) \sim Cy$ $(y \to 0)$ and $\theta_1(y) \sim Cy^{3/8} e^{-cy^{1/4}}$ $(y \to \infty)$, by ($2W$) and ($3W$). Here (and elsewhere) C is some positive constant (often depending on s), but not necessarily the same each time it occurs. Thus the effect of our work so far has been to replace the natural integral equation, (1), by a second, (9); but (9) is the easier to handle.

We now find A and B. We first note that $B = 0$. For

(i) $\theta_1(y) \sim Cy^{3/8}\exp(-4b^{-1} s^{1/2} y^{1/4})$ $(y \to \infty)$, by ($3W$),
(ii) $\theta_2(y) \sim Cy^{3/8}\exp(4b^{-1} s^{1/2} y^{1/4})$ $(y \to \infty)$, by ($4W$) and
(iii) $f(y) = O(y^{-3})$ $(y \to \infty)$; thus the two integral terms of (9) tend to zero as $x \to \infty$. Since $0 \leqq \phi \leqq 1$ for all x, there is (for $x \to \infty$) just one unbounded term in (9) if $B \neq 0$, and so $B = 0$.

To find A we need further boundary conditions on ϕ. From (1) (with $k = 0$) we have

$$(10) \qquad \phi(s\,|\,0+) = \tfrac{1}{2} + \tfrac{1}{2} b^{-1} J ,$$

where

$$J = \int_0^\infty e^{-w/b} e^{-sV(w)} \phi(s/w)\,dw ;$$

and from (2) (with $k = 0$)

$$D\phi(s\,|\,0+) = -\tfrac{1}{2} b^{-1} + \tfrac{1}{2} b^{-2} J .$$

Thus

$$(11) \qquad \phi(s\,|\,0+) = 1 + bD\phi(s\,|\,0+)$$

which, with (9), allows us (after some detailed calculation) to evaluate A. We find, by elaboration of the methods used below, that

$$A = \left\{\frac{b^2}{8s} + \frac{s}{b}\log\left(\frac{4s}{b^2}\right) + (2\gamma - 1)\frac{s}{b}\right\}^{-1};$$

we shall not give the details because the asymptotic distribution of $T/\sqrt{x}$ is obtainable without this complete treatment of A. We shall show later (in § 4) that

$$\phi = Ab^2/8s + O(x^{1/2}) \qquad (x \to 0), \tag{12}$$

when $s > 0$, whence $\phi(s\,|\,0+) = Ab^2/8s$. Then from (10), recalling that A depends on s, we see that

$$\begin{aligned}\lim_{s\to 0}(Ab^2/8s) &= \tfrac{1}{2} + \tfrac{1}{2}b^{-1}\lim_{s\to 0}\int_0^\infty e^{-w/b}e^{-sV(w)}\phi(s\,|\,w)dw \\ &= \tfrac{1}{2} + \tfrac{1}{2}b^{-1}\int_0^\infty e^{-w/b}\,dw\,\Pr\{T<\infty\} = 1\end{aligned}$$

(because almost certainly there will be only finitely many complete circuits), so that

$$A \sim 8s/b^2 \qquad (s \to 0).$$

Now put $s\sqrt{x} = \sigma > 0$ in (9) and let $x \to \infty$ and $s \to 0$, σ being fixed. Then

$$\mathscr{E}(e^{-\sigma T/\sqrt{x}}\,|\,x) = \phi(s\,|\,x) \to 8b^{-2}\sigma K_2(4b^{-1}\sqrt{\sigma}), \ (\sigma > 0,\ x \to \infty). \tag{13}$$

if both integral terms of (9) tend to zero; this is in fact the case, as we show in §3. The (honest) probability distribution

$$\frac{16}{b^4}\exp\left(-\frac{4}{b^2\tau}\right)\frac{d\tau}{\tau^3} \qquad (0 < \tau < \infty) \tag{14}$$

has the expression on the right-hand side of (13) as its Laplace transform, so that

$$\lim_{x\to\infty}\mathscr{E}(e^{-\sigma T/\sqrt{x}}\,|\,x) = \mathscr{E}(e^{-\sigma\tau})$$

for all $\sigma > 0$. It follows by the continuity theorem for the Laplace-Stieltjes transforms of probability distributions of nonnegative random variables that

$$\lim_{x\to\infty}\Pr\left\{\frac{T}{\sqrt{x}} \leqq c\,|\,x\right\} = \frac{16}{b^4}\int_0^c \exp\left(-\frac{4}{b^2\tau}\right)\frac{d\tau}{\tau^3}, = \left(1+\frac{4}{b^2c}\right)\exp\left(-\frac{4}{b^2c}\right). \tag{15}$$

This is our main result; but it is clear that (to some extent at least) the precise value of α affects the detail of (15) rather than its essential nature. For $\alpha \neq 2$ we can take as independent solutions of (9)

$$\theta_1(x) = x^{1/2}K_\nu(2\nu b^{-1}s^{1/2}x^{1-(1/2)\alpha}) \quad \text{and} \quad \theta_2(x) = x^{1/2}I_\nu(2\nu b^{-1}s^{1/2}x^{1-(1/2)\alpha}),$$

where $\nu = |\alpha - 2|^{-1}$ (for $\alpha = 2$ the solutions are powers of x), and then find as before that W is a constant independent of y. But some limitations on α are imposed by the need for suitable behaviour of various integral terms, and we merely note here one analogue of (15):

if $V(x) = x^{-1}$, *then*

$$\lim_{x\to\infty} \Pr\left\{\frac{T}{x} \leqq c \,|\, x\right\} = \frac{1}{b^2}\int_0^c \exp\left(-\frac{1}{b^2\tau}\right)\frac{d\tau}{\tau^2}, = \exp\left(-\frac{1}{b^2 c}\right). \tag{16}$$

3. Analytical details. Consider the behaviour of the integral terms in (9) when $x \to \infty$ and $s\sqrt{x} = \sigma > 0$. These terms are (apart from constant factors)

$$\theta_1(x)\int_0^x f(y)\,\phi(s\,|\,y)\,\theta_2(y)\,dy \tag{a}$$

and

$$\theta_2(x)\int_x^\infty f(y)\,\phi(s\,|\,y)\,\theta_1(y)\,dy. \tag{b}$$

Write (a) as $\theta_1(x)\left\{\int_0^1 + \int_1^x f(y)\,\phi(s\,|\,y)\,\theta_2(y)\,dy\right\} = A_1 + A_2$, say, and consider separately the terms A_1 and A_2. Using $|f| < 2b^{-2}sy^{-3/2}\,(y > 0)$, and noting that $\theta_1(x) = \sigma s^{-1}K_2(4b^{-1}\sqrt{\sigma})$, we have

$$|A_1| < C\frac{\sigma}{s}\int_0^1 \frac{s}{y^{3/2}}\cdot y^{1/2}I_2\left[\frac{4\sqrt{\sigma}}{b}\left(\frac{y}{x}\right)^{1/4}\right]dy$$

$$< C\int_0^1 \frac{1}{y}\left(\frac{y}{x}\right)^{1/2} dy$$

for x sufficiently large, since $I_2(\lambda^{1/4}) = O(\lambda^{1/2})$ $(\lambda \to 0)$. Thus $A_1 = O(x^{-1/2})$ $(x \to \infty)$. For A_2 we use $|f| < \frac{1}{2}s^2/y^3b^2$ and find

$$|A_2| < C\frac{\sigma}{s}\int_1^x \frac{s^2}{y^3}y^{1/2}I_2\left[\frac{4\sqrt{\sigma}}{b}\left(\frac{y}{x}\right)^{1/4}\right]dy < Cs\int_1^x y^{-5/2}\,dy,$$

since $I_2(\theta)$ is bounded for $0 \leqq \theta \leqq 4b^{-1}\sqrt{\sigma}$. Thus $A_2 = O(x^{-1/2})\,(x \to \infty)$, so that (a) $\to 0$ as $x \to \infty$ with $s\sqrt{x} = \sigma > 0$.

For (b) have

$$|(b)| < C\,x^{1/2}\int_x^\infty y^{1/2}\frac{s^2}{y^3}\,dy = O(x^{-2}) \qquad (x \to \infty),$$

which completes the proof that the integral terms tend to zero.

4. Analytical details (continued). We now prove (12). To do this we need (part of)

(i) $\theta_1(x) = b^2/8s - \frac{1}{2}x^{1/2} + O(x \log x) \qquad (x \to 0)$,

(ii) $\theta_2(x) = 2sb^{-2}x + O(x^{3/2}) \qquad (x \to 0)$

and

(iii) $1 - sV(x) - e^{-sV(x)} = -sx^{-3/2} + O(1) \qquad (x \to 0)$.

Proofs. By (5 W)

$$K_2(z) = 2z^{-2} - \tfrac{1}{2} + O(z^2 \log z) \qquad (z \to 0),$$

whence (i) follows. Likewise, using (2 W), (ii) follows from

$$I_2(z) = z^2/8 + O(z^4) \qquad (z \to 0).$$

Finally, (iii) follows from $0 < e^{-sV(x)} < 1$ for $x > 0$, $s > 0$.

We now note that (12) follows at once from (9) if we show that

(iv) $\theta_1(x) \int_{0+}^{x} f(y)\,\phi(y)\,\theta_2(y)\,dy = O(x^{1/2}) \qquad (x \to 0)$

and

(v) $\theta_2(x) \int_{x}^{\infty} f(y)\,\phi(y)\,\theta_1(y)\,dy = O(x^{1/2}) \qquad (x \to 0)$.

These results follow from those already given. We have

$$f(y) = O(y^{-3/2}) \ (y \to 0) \quad \text{and} \quad \theta_2(y) = O(y) \ (y \to 0)$$

so that

$$\int_0^x f(y)\,\phi(y)\,\theta_2(y)\,dy = O(x^{1/2}) \qquad (x \to 0).$$

Since $\theta_1(x) = O(1)$ $(x \to 0)$, this gives (iv). The proof of (v) is similar.

References

1, 2. J. M. Hammersley and R. A. Lyttleton, *On the statistical loss of long period comets from the solar system*, I, II, 4th Berkeley Symp. on Math. Stat. and Prob., Berkeley, University of California Press, 1961.

3. R. H. Kerr, *Perturbations of cometary orbits*, 4th Berkeley Symp. on Math. Stat. and Prob., Berkeley, University of California Press, 1961.

4. D. G. Kendall, *The distribution of energy perturbations for Halley's and some other comets*, 4th Berkeley Symp. on Math. Stat. and Prob., Berkeley, University of California Press, 1961.

5, 6. D. G. Kendall, *Some problems in the theory of comets*, I, II, 4th Berkeley Symp. on Math. Stat. and Prob., Berkeley, University of California Press, 1961.

7. G. N. Watson, *A treatise on the theory of Bessel functions*, second edition, Cambridge, 1944.

MAGDALEN COLLEGE, OXFORD.
UNIVERSITY OF EDINBURGH.

THE SPECTRUM OF SINGULAR SELF-ADJOINT ELLIPTIC OPERATORS

KURT KREITH

This note deals with the Dirichlet problem for the second order elliptic operator

$$L = -\frac{1}{r(x)} \sum_{i,j=1}^{n} \frac{\partial}{\partial x_j}\left(a_{ij}(x)\frac{\partial}{\partial x_i}\right) + c(x)$$

whose coefficients are defined in a bounded domain $G \subset E^n$. We suppose the following:

(a) The $a_{ij}(x)$ are complex valued and of class C' in G; $a_{ij} - \bar{a}_{ji}$.
(b) $c(x)$ is real valued, continuous, and bounded below in G.
(c) $r(x)$ is continuous and positive in G.
(d) There exists a function $\sigma(x)$, continuous and positive in G satisfying

$$\sum_{i,j=1}^{n} a_{ij}\xi_i\bar{\xi}_j \geqq \sigma \sum_{i=1}^{n} |\xi_i|^2$$

for all x in G and all complex n-tuples $\vec{\xi} = (\xi_1, \xi_2, \cdots, \xi_n)$.

Under these assumptions it is easy to show that L is formally self-adjoint in the Hilbert space $\mathscr{L}^2_r(G)$ of functions which satisfy

$$\int_G r\,|u|^2\,dx < \infty\,.$$

We denote by $C_0^\infty(G)$ the set of infinitely differentiable functions with compact support in G. The operator L defined on $C_0^\infty(G)$ is a semi-bounded symmetric operator in $\mathscr{L}^2_r(G)$ and therefore has a Friedrichs extension which is self-adjoint in $\mathscr{L}^2_r(G)$. This operator, to be denoted by $\bar{L}$, will be referred to as the Dirichlet operator associated with L on G. It is well known that $\bar{L}$ is unique, has the same lower bound as the symmetric operator L, and that in sufficiently regular cases, $\bar{L}$ can be obtained by imposing Dirichlet boundary conditions on the domain of L^*. The purpose of this note is to state a criterion for the discreteness of the spectrum of $\bar{L}$.

We shall say that the spectrum of an operator A is discrete if the spectrum of A consists of a set of isolated eigenvalues of finite multiplicity. The complex number λ belongs to the essential spectrum of A if there exists an orthonormal sequence $\{u_n\}$ it the domain of A for which $(A - \lambda I)u_n \to 0$. If A is self-adjoint, then it can be shown (see

Received December 6, 1960. This research was partially supported by a grant of the National Science Foundation NSF G 5010.

[2]) that λ belongs to the essential spectrum of A if and only if λ belongs to the spectrum of A and is not an isolated eigenvalue of finite multiplicity. Thus the spectrum of a self-adjoint operator is discrete if and only if the essential spectrum is empty.

In case G is bounded and the conditions (a)–(b) are satisfied in $\bar{G}$ as well as G, then it is well known that $\bar{L}$ has a discrete spectrum. Here we shall allow the possibility that $\sigma(x)$ and $r(x)$ tend to 0 or ∞ on a set $S \subset \partial G$. With this generalization the spectrum of $\bar{L}$ need not be discrete.

In order to state criteria for the discreteness of the spectrum of $\bar{L}$, it is convenient to express the problem in the canonical form where

$$\begin{aligned} &G \subset \{x \mid x_n > 0\} \\ &S \subset \{x \mid x_n = 0\} \\ &L = \frac{\partial}{\partial x_n}\left(a_{nn}\frac{\partial}{\partial x_n}\right) + \sum_{i,j=1}^{n-1}\frac{\partial}{\partial x_j}\left(a_{ij}\frac{\partial}{\partial x_i}\right) + c \end{aligned}$$

Mihlin [1] has shown that this canonical form can in general be attained by a change of variables. Previous criteria for discreteness derived by Mihlin [1], Wolf [2], and others depend on the behavior of a_{nn} near S. The criterion to be derived here is independent of the behavior of a_{nn}; with minor modification, the method can also be applied if G is an unbounded domain.

We define

$$\begin{aligned} G_t &= G \cap \{x \mid x_n < t\} \\ E_t &= G \cap \{x \mid x_n = t\} , \end{aligned}$$

and denote by $\bar{x}$ the coordinates $(x_1, \cdots, x_{n-1})$ in E_t. Let $\bar{L}_t$ denote the Dirichlet operator associated with L on G_t. Then the following is a special case of an invariance principle due to Wolf [2].

LEMMA 1. *For $t > 0$ the essential spectrum of $\bar{L}_t$ is identical with the essential spectrum of $\bar{L}$.*

LEMMA 2. *If* $\lim\limits_{t\to 0} \inf\limits_{u \in C_0^\infty(G_t)} \dfrac{(Lu, u)}{\| u \|^2} = \infty$, *then the spectrum of $\bar{L}$ is discrete.*

Proof. Suppose to the contrary that there is a $\lambda_0 < \infty$ which belongs to the essential spectrum of $\bar{L}$. We can choose $t_0 > 0$ sufficiently small so that

$$\inf_{u \in C_0^\infty(G_{t_0})} \frac{(Lu, u)}{\| u \|^2} \geqq \lambda_0 + 1 .$$

Then, by the definition of $\bar{L}_{t_0}$

$$\frac{(\bar{L}_{t_0}u, u)}{\| u \|^2} \geqq \lambda_0 + 1$$

for all u in the domain of $\bar{L}_{t_0}$, and λ_0 does not belong to the spectrum of $\bar{L}_{t_0}$. By Lemma 1 this is a contradiction.

For $t > 0$ the operator

$$T_t = -\frac{1}{r(\bar{x}, t)} \sum_{i,j=1}^{n-1} \left(a_{ij}(\bar{x}, t) \frac{\partial}{\partial x_i} \right) + c(\bar{x}, t)$$

is a nonsingular elliptic operator defined on E_t. Therefore $\bar{T}_t$, the Dirichlet operator associated with T_t on E_t, has a discrete spectrum. Let $m(t)$ denote the smallest eigenvalue of $\bar{T}_t$.

THEOREM. *If $\lim_{t \to 0} m(t) = \infty$, then the spectrum of $\bar{L}$ is discrete.*

Proof. If $u \in C_0^\infty(G)$, then clearly $u(\bar{x}, t) \in C_0^\infty(E_t)$. Thus for all $u \in C_0^\infty(G)$

$$\begin{aligned} m(t)\int_{E_t} | u |^2 r d\bar{x} &\leqq \int_{E_t} \left[\sum_{i,j=1}^{n-1} a_{ij} \frac{\partial u}{\partial \bar{x}_i} \frac{\partial \bar{u}}{\partial x_j} + rc \, | u |^2 \right] d\bar{x} \\ &\leqq \int_{E_t} \left[a_{nn} \left| \frac{\partial u}{\partial x_n} \right|^2 + \sum_{i,j=1}^{n-1} \frac{\partial u}{\partial x_i} \frac{\partial \bar{u}}{\partial x_j} + rc \, | u |^2 \right] dx \, . \end{aligned}$$

Defining $\bar{m}(t) = \inf_{\tau \leqq t} m(\tau)$ and integrating both sides from $x_n = 0$ to $x_n = t$ we obtain

$$\bar{m}(t)\int_{G_t} | u |^2 r dx \leqq \int_{G_t} \left[a_{nn} \left| \frac{\partial u}{\partial x_n} \right|^2 + \sum_{i,j=1}^{n-1} a_{ij} \frac{\partial u}{\partial x_i} \frac{\partial \bar{u}}{\partial x_j} + rc \, | u |^2 \right] dx \, .$$

Since $\lim_{t \to 0} \bar{m}(t) = \infty$ we have

$$\lim_{t \to 0} \inf_{u \in C_0^\infty(G_t)} \frac{(Lu, u)}{\| u \|^2} = \infty \, .$$

The desired result now follows from Lemma 2.

We give two simple applications of the preceding theorem.

COROLLARY 1. *Let V_t denote the $(n-1)$-dimensional Lebesgue measure of E_t. Let $\phi(t)$ and $\rho(t)$ be continuous positive functions satisfying*

(i) $\rho(t) \geqq r(\bar{x}, t)$

(ii) $\phi(t) \sum_{i=1}^{n-1} | \xi_i |^2 \leqq \sum_{i,j=1}^{n-1} a_{ij}(x, t) \xi_i \xi_j$ for all $\vec{\xi} = (\xi_1, \cdots, \xi_{n-1})$.

If $\lim_{t\to 0} \phi(t)/\rho(t)V_t^{2/n-1} = \infty$, then the spectrum of $\bar{L}$ is discrete.

Proof. Let $\mu(t)$ denote the smallest eigenvalue of the Dirichlet operator associated with $-\Delta = -\sum_{i=1}^{n-1}\partial^2/\partial x_i^2$ on E_t. By (i) and (ii) $m(t) \geqq \phi(t)\mu(t)/\rho(t)$. It is well known that $\mu(t)$ is minimized if E_t is a $(n-1)$-dimensional sphere of volume V_t and that then $\mu(t) = C/V_t^{2/n-1}$, C being a constant. Therefore $m(t) \geqq C\phi(t)/\rho(t)V_t^{2/n-1}$ and the result follows from the preceding theorem.

The preceding corollary made no use of the shape of E_t. The following corollary gives stronger results in case E_t becomes "narrow" in the proper sense.

COROLLARY 2. *Suppose we can find functions* $\alpha_1(x_n), \cdots, \alpha_{n-1}(x_n)$, $\gamma(x_n)$ *and* $\rho(x_n)$ *which satisfy conditions* (a)–(d) *and*

(i) $\sum_{i=1}^{n-1} \alpha_i(x_n) \mid \xi_i \mid^2 \leqq \sum_{i,j=1}^{n-1} a_{ij}\xi_i\bar{\xi}_j$ *for all* $\xi = (\xi_1, \cdots, \xi_{n-1})$ *and all* x *in* G.

(ii) $\gamma(x_n) \leqq c(x)$ *for all* x *in* G.

(iii) $\rho(x_n) \geqq r(x)$ *for all* x *in* G.

Suppose also that we can enclose G *in a region*

$$\Gamma = \{x \mid f_i(x_n) < x_i < g_i(x_n),\ i = 1, \cdots, n-1;\ 0 < x_n < b < \infty\}\ .$$

If for some $i < n$

$$\lim_{t\to 0} \frac{\alpha_i(t)}{\rho(t)[g_i(t) - f_i(t)]^2} + \gamma(t) = \infty$$

then the spectrum of $\bar{L}$ *is discrete.*

Proof. Denote by $\mu(t)$ the smallest eigenvalue of the Dirichlet operator associated with

$$\tau(t) = -\frac{1}{\rho(t)} \sum_{i=1}^{n-1} \alpha_i(t) \frac{\partial^2}{\partial x_i^2} + \gamma(t)$$

on $\Gamma \cap \{x \mid x_n = t\}$. By classical variational principles $\mu(t) \leqq m(t)$. Since we can compute

$$\mu(t) = \pi^2 \sum_{i=1}^{n-1} \frac{\alpha_i(t)}{\rho(t)[g_i(t) - f_i(t)]^2} + \gamma(t)\ ,$$

the discreteness of the spectrum of $\bar{L}$ follows from the preceding theorem.

BIBLIOGRAPHY

1. S. G. Mihlin, *Degenerate Elliptic Equations*, Vestnik Lengrad Univ. **8** (1954), 19-48.

2. F. Wolf, *On the essential spectrum of partial differential boundary problems*, Comm. on Pure and App. Math., **12** (1959), no. 2.

UNIVERSITY OF CALIFORNIA, DAVIS

THE SEMICONTINUITY OF THE MOST GENERAL INTEGRAL OF THE CALCULUS OF VARIATIONS IN NON-PARAMETRIC FORM.*

LIONELLO LOMBARDI

Summary. The positive quasi-regularity of integrals depending upon any number of surfaces in non-parametric form, each with any number of dimensions, is defined. Positive quasi regularity is proved to be sufficient for lower semicontinuity.

1. Let $D_i (i = 1, 2, \cdots, m)$ be a closed bounded set of the n-dimensional euclidean space of the variable vector $x_i \equiv \{x_i^j\} (j = 1, 2, \cdots, n)$, bounded by surfaces which are absolutely continuous in the sense of Tonelli [60, 62, 63], without multiple points, and let D be the cartesian product $\prod_{1_i}^{m} D_i$. Let $y \equiv \{y_i\} (i = 1, 2, \cdots, m)$ denote a vertical m-vector, and let p denote *a mxn* matrix, whose row-vectors are $p_i \equiv \{p_i^j\} (j = 1, 2, \cdots, n)$. Let x be the $m \cdot n$ matrix whose row-vectors are x_i and $\phi[x, y, p]$ a real function, defined for $x_i \in D_i$ $(i = 1, 2, \cdots, m)$ and for any y and p, which is continuous with all its partial derivatives of the types

$$\frac{\partial \phi[x, y, p]}{\partial p_r^s}, \quad \frac{\partial^2 \phi[x, y, p]}{\partial p_r^s \partial p_r^t} \qquad (r = 1, \cdots, m; s, t = 1, \cdots n).$$

Let $q = m$ be a positive integer and let U_q denote a set of distinct positive integers out of the first m; let ζ be an index ranging over U_q, and let $\mu(\delta)$ be a mapping of U_q into the set of the first n integers. It will be assumed throughout that, for every q, every U_q and every $\mu(\zeta)$, all the partial derivatives

$$\text{(1.1)} \qquad \frac{\partial^{2q} \phi[x, y, p]}{\prod_1^q \partial x_\zeta^{\mu(\zeta)} \partial p_\zeta^{\mu(\zeta)}}$$

exist and are continuous for every $x \in D$ and for every y and p.

Let $y(x) \equiv \{y_i(x_i)\}$ $(i = 1, 2, \cdots, m)$ denote a vector-valued function of the matrix x, such that each *component* $y_i(x_i)$ depends only upon the row vector x_i. We assume that each $y_i(x_i)$ is absolutely continuous, in the sense of Tonelli [63]; we shall call *Variety* V the set of m surfaces represented by $y(x)$.

Received July 29, 1960. The preparation of this paper was sponsored in part by the Office of Naval Research and the Office of Ordnance Research, U. S. Army. The author also wishes to acknowledge the help given Gerald W. Kimble in the preparation of this paper. The author is now with Scientific Information Treatment Centre, EURATOM.

We shall say that V is of *class* 1 if all the first partial derivatives of all the $y_i(x_i)$ exist and are continuous; we shall say the V is of *class* 2 if the same is also true for all the partial derivatives of the second order.

Let

$$p_i^j(x) \equiv \frac{\partial y_i(x_i)}{\partial x_i^j} \qquad (i = 1, 2, \cdots m;\ j = 1, 2, \cdots, n)\,,$$

and

$$dx \equiv \prod_1^m {}_i dx_i \equiv \prod_1^m {}_i \prod_1^n {}_j dx_i^j\,.$$

The $m \cdot n$ integral[1]

$$I_V = \int_D \phi[x, y(x), p(x)]dx$$

is called *variety integral in non-parametric form;* all the varieties V where I exists and is finite are called *ordinary.*

REMARK 1.1. *Varieties* V *of class* 1 *and* 2 are ordinary for any function $\phi[x, y, p]$.

Let $\bar{p} \equiv \{\bar{p}_i\} \equiv \{\bar{p}_i^j\}$ denote another variable in the space of the matrix p, $\bar{y} \equiv \{\bar{y}_i\}$ another variable in the space of the vector y, $V \equiv \bar{y}(x) \equiv \{\bar{y}_i(x_i)\}$ another variety V; let

$$\bar{p}_i^j(x_i) = \frac{\partial \bar{y}_i(x_i)}{\partial x_i^j}\ ;$$

the distance $\rho(V, \bar{V})$ between V and $\bar{V}$ is defined by the formula

$$\rho(V, \bar{V}) = \sup_{x,\,i} |\, y_i(x) - \bar{y}_i(x)\,|\ .$$

Continuity and semicontinuity of the real function I_V will be considered throughout with respect to this *m-uniform* metric.

In one of our previous papers [33] the following theorem was proved:

CONTINUITY THEOREM 1.2. *Necessary and sufficient conditions for the continuity of* I_V *at every* V *is that the function* $\phi[x, y, p]$ *is linear with respect to each one of the vectors* p_i.

REMARK 1.3. As a consequence of Theorem 1.2, the most general function $\phi[x, y, p]$, such that $\int_D \phi[x, y(x), p(x)]\, dx$ is continuous at every

1 For the relation between this integral and non local field theories see bibliography [**1, 6, 27, 28, 29, 40, 41, 42, 46, 47, 48, 58**].

V, may be written in the form

$$\sum_{q=1}^{m} \sum_{U_q} \sum_{\mu} \{A_{U_q,\mu}(x, y) \prod_{\zeta \in U_q} p_{\zeta}^{\mu(\zeta)}\}, \tag{1.2}$$

where we assume by convention that, if η is a variable integer ranging over a set S and $\{\alpha_\eta\}$ is a sequence of numbers, then

$$\prod_{\eta \in S} \alpha_\eta = 0, \text{ whenever } S \text{ is empty}.$$

Let $L[x, y, p, \bar{p}]$ denote a polynomial in the indeterminates

$$[\bar{p}_i^{(j)} - p_i^{(j)}] \tag{1.3}$$

of degree not exceeding 1 in any of the vectors $[\bar{p}_i - p_i]$, whose coefficients $W_{U_q,\mu}(x, y, p)$ are functions of (x, y, p) which are continuous together with all their derivatives of the form

$$\frac{\partial^q W_{U_q,\mu}(x, y, p)}{\prod_{\zeta \in U_q} \partial x_{\zeta}^{\mu(\zeta)}}, \tag{1.4}$$

$L[x, y, p, \bar{p}]$ may be written in the form

$$\sum_{q=1}^{m} \sum_{U_q} \sum_{\mu} \{W_{U_q,\mu}(x, y, p) \prod_{\zeta \in U_q} [\bar{p}_{\zeta}^{\mu(\zeta)} - p_{\zeta}^{\mu(\zeta)}]\}. \tag{1.5}$$

Let us define the *generalized Weierstrass function* $E_L[x, y, p, \bar{p}]$ *of* L_V *with respect to* $L[x, y, p, \bar{p}]$, by the formula

$$E_L(x, y, p, \bar{p}) = \phi[x, y, \bar{p}] - L[x, y, p, \bar{p}]. \tag{1.6}$$

The integral $I_V = \int \phi[x, y(x), p(x)]dx$ is said to be *positively quasi-regular with respect to* L (abbreviation: $LPQR$) if both the relations

$$E_L[x, y, p, p] = 0 \tag{1.7}$$

$$E_L[x, y, p, \bar{p}] \geqq 0 \tag{1.8}$$

hold for every $x \in D$ and for every $y, p, \bar{p}$.

REMARK 1.4. Notice that if I_V is $LPQR$, then the element of degree 0 of the polynomial $L[x, y, p, \bar{p}]$ must be $\phi[x, y, p]$, and the vector consisting of the coefficients of the elements of degree 1 is the gradient of $[x, y, p]$ with respect to p: therefore, if $m = 1$, i.e., if I_V is a usual multiple integral [60, 62], the fact that I_V is $LPQR$ completely determines the function $L[x, y, p, \bar{p}]$. This does not happen if $m > 1$, as was shown by an appropriate example [30], referring to Fubini-Tonelli integrals, i.e., to the case $(m = 2, n = 1)$.

We say that I_V is *positively quasi-regular* (abbreviation PQR) if there exists at least one function $L[x, y, p, \bar{p}]$ such that I_V is $LPQR$.

REMARK 1.5. Let us say that I_V is *negatively quasi-regular* with respect to L (abbreviation: $LNQR$) if $\int_D -\phi[x, y(x), p(x)]dx$ is $LPQR$. Then it is easy to prove that, if I_V is both L_1PQR and L_2NQR, then $L_1[x, y, p, \bar{p}] \equiv L_2[x, y, p, \bar{p}]$, and $\phi[x, y, p]$ is a polynomial of degree not exceeding 1 in each p_i; i.e., by Theorem 1.2, I_V is continuous. Theorem 1.2 also implies that every continuous I_V is both $LPQR$ and $LNQR$ for some $L[x, y, p, \bar{p}]$.

REMARK 1.6 In the case $m = 1$, our definition of *positive quasi-regularity* reduces to the one which was given by Tonelli [59, 60] and Cinquini [1] for simple and multiple integrals. In this particular case, the *positive quasi-regularity* of an integral is equivalent to the *lower convexity* of its *figurative*, i.e., of $\phi[x, y, p]$ considered as a function of p only.

In the case $n = 1$, the definition of *positive quasi-regularity* reduces to the one given by this author for the Fubini-Tonelli integrals [30].

REMARK 1.7. If I_V is PQR, then its value is $+\infty$ at every non-ordinary variety.

2. Let us prove the following

THEOREM 2.1. *If I_V is PQR, then it is lower semicontinuous at every variety V of class 2; i.e., if V is of class 2, there exists a positive function $\rho(\varepsilon)$ of the positive variable ε such that, if $\bar{V} \equiv \bar{y}(x)$ is any variety, then*

$$I_{\bar{V}} - I_V > -\varepsilon, \text{ whenever } \rho(V, \bar{V}) < \rho(\varepsilon), \tag{2.1}$$

regardless of whether or not V is of class 2.

Proof. Let $L[x, y, p, \bar{p}]$ be a function, such that I_V is $LPQR$. By (1.6) we may write

$$\begin{aligned} I_{\bar{V}} - I_V = &\int_D E_L[x, \bar{y}(x), p(x), \bar{p}(x)]dx - \int_D E_L[x, y(x), p(x), p(x)]dx \\ &+ \int_D L[x, \bar{y}(x), p(x), \bar{p}(x)]dx - \int_D L[x, y(x), p(x), p(x)]dx . \end{aligned} \tag{2.2}$$

Let $V \equiv y(x)$ be a variety of *class* 2;

$$P[x, \bar{y}, \bar{p}] \equiv L[x, \bar{y}, p(x), \bar{p}]$$

is a polynomial of a degree not exceeding 1 in each $\bar{p}$, and all of the derivatives

$$\frac{\partial P[x, \bar{y}, \bar{p}]}{\partial \bar{p}_r^s}, \quad \frac{\partial^2 P[x, \bar{y}, \bar{p}]}{\partial \bar{p}_r^s \partial \bar{p}_r^t}, \quad \frac{\partial^{2q} P[x, \bar{y}, \bar{p}]}{\prod_1^q {}_\zeta \partial x_\zeta^{\mu(\zeta)} \partial \bar{p}_\zeta^{\mu(\zeta)}} \tag{2.3}$$

$$(r = 1, 2, \cdots, m; s, t = 1, 2, \cdots n)$$

exist and are continuous for every $[x, \bar{y}, \bar{p}]$ and for every q, U_q, $\mu(\zeta)$, r, s, t as a consequence of the existence and continuity of the functions (1.4) and of the partial derivatives of the first two orders of the functions $y_r(x)$, $(r = 1, 2, \cdots, m)$.

By the continuity Theorem 1.2,

$$J_V = \int_D P[x, \bar{y}(x), \bar{p}(x)]dx$$

is continuous; hence the difference of the last two integrals on the right side of (2.2) is smaller than any predetermined real positive ε, whenever $\rho(V, \bar{V})$ is less than a certain positive number $\rho(\varepsilon)$. Since the first integral on the right side of (2.2) is not negative by (1.8) and the second vanishes by (1.7), (2.1) holds: the theorem is thus proved.

3. (a) In this section the concept of *asymptotic evaluability of the integral* I_V is defined; the lower semicontinuity on every very variety V of any *positively quasi-regular* and *asymptocally evaluable* integral is proved. The results of this chapter may be regarded as extensions of Tonelli's theorems on usual multiple integrals [59, 60], and of our results on Fubini-Tonelli integrals [30].

(b) Suppose that $I_V = \int_D \phi[x, y(x), p(x)]dx$ is PQR, and let $L[x, y, p, \bar{p}]$ be one of the functions, such that I_V is $LPQR$.

The function

$$\bar{\Phi}[x, y, p] \equiv E_L[x, y, \Omega, p] , \tag{3.b.1}$$

where Ω is a $m \cdot n$ matrix whose elements are all 0, is never negative. Furthermore,

$$\bar{I}_V = \int_D \bar{\Phi}[x, y(x), p(x)]dx \tag{3.b.2}$$

is $\bar{L}PQR$, where

$$\bar{L}[x, y, p, \bar{p}] = L[x, y, p, \bar{p}] - L[x, y, \Omega, \bar{p}]. \tag{3.b.3}$$

By (1.7), the equation

$$\bar{\Phi}[x, y, \Omega] = 0$$

holds for every $x \in D$ and every y.

Let R denote a positive real number and let $\varphi^R[x, y, p]$ denote a function such that the following conditions are satisfied:

I. $\varphi^R[x, y, \mathrm{p}]$ is continuous with all its partial derivatives of any of the forms

$$\frac{\partial\varphi^R[x, y, p]}{\partial p_r^s}, \quad \frac{\partial^2\varphi^R[x, y, p]}{\partial p_r^s \partial p_r^t}, \quad \frac{\partial^{2q}\varphi^R[x, y, p]}{\prod_1^q{}_\zeta \partial x_\zeta^{\mu(\zeta)} \partial p_\zeta^{\mu(\zeta)}}.$$

II. The integral

$$Y_V = \int_D \varphi^R[x, y, (x), p(x)]dx \tag{3.b.4}$$

is PQR.

III. The relation

$$0 \leqq \varphi^R[x, y, p] \leqq \bar{\Phi}[x, y, p] \tag{3.b.5}$$

holds for every y, p and for every $x \in D$; furthermore

$$\varphi^R[x, y, p] = \bar{\Phi}[x, y, p], \text{ whenever } \sum_{i=1}^m \sum_{j=1}^n (p_1^j)^2 \leqq R. \tag{3.b.6}$$

IV. There exists at least one function $\Lambda[x, y, p, \bar{p}]$, such that Y_V is ΛPQR, and such that, for each $T > 1$, there exists a number Q, for which the following condition is satisfied:

Let q, U_q, ζ, $\mu(\zeta)$ be defined as they were in § 1; *let $\bar{U}_q$ denote the complement of U_q with respect to the set of the first m positive integers, and let $\bar{\zeta}$ be an index ranging over $\bar{U}_q$. Then the inequality*

$$\left| W_{U_q,\mu}^R[x, y, p] \right| < Q\left(1 + \prod_{\bar{\zeta} \in \bar{U}_q} \left| p_{\bar{\zeta}}^{\mu(\bar{\zeta})} \right| \right)$$

where $W_{U_q,\mu}^R[x, y, p]$ denotes the coefficient of the element

$$\prod_{\zeta \in U_q} [\bar{p}_\zeta^{\mu(\zeta)} - p_\zeta^{\mu(\zeta)}]$$

of the expression $\Lambda[x, y, p, \bar{p}]$, holds for every q, U_q, p, for every $x \in D$ and for every y such that

$$|y_i| < T \qquad (i = 1, 2, \cdots, n).$$

REMARK 3.1. In the case of the usual multiple integrals ($m = 1$), Condition IV reduces to the boundedness of the derivatives

$$\frac{\partial \varphi^R[x, y, p]}{\partial p_1^s} \qquad (s = 1, 2, \cdots, n)$$

in any domain where $y(x)$ is bounded; this condition is exactly the one considered by Tonelli [59, 60].
In the case of Fubini-Tonelli integrals ($n = 1$), this condition reduces to the one that this author considered in [30, § 1, page 132].

REMARK 3.2. Y_V exists and is finite on every variety V, i.e., every variety V is ordinary for the integral Y_V.

(c) LEMMA 3.3. *The integral Y_V defined by* (3.b.4) *is lower semicontinuous at every variety V.*

Proof. Let $V \equiv y(x) \equiv \{y_i(x_i)\}$ be any variety; and let $1 > \varepsilon > 0$ and $R > 0$ be given, and let $\pi \equiv \pi(x) \equiv \{\pi_i(x_i)\}$ denote a variety of *class* 2, such that

$$\rho(\pi, V) < \varepsilon \tag{3.c.1}$$

Let $T = \sup_{x_i, i} | y_i(x_i) | + 2$.

Let $\pi'(x) \equiv || \pi'^j_i(x) || \equiv \left|\left| \dfrac{\partial \pi_i(x_i)}{\partial x_i^j} \right|\right|$, $(i = 1, 2, \cdots, m; j = 1, 2, \cdots, n)$,

and let $\bar{D}_i \subset D_i$ denote set of the points x_i, such that, for some j, either $p_i^j(x_i)$ does not exist or it is such that

$$| \pi'^j_i(x_i) - p_i^j(x_i) | \geqq \varepsilon . \tag{3.c.2}$$

Suppose further that, for each i $(i = 1, 2, \cdots, m)$,

$$\int_{\bar{D}_i} \sum_1^n{}_j [| \pi'^j_i(x_i) | + | p_i^j(x_i) |] dx_i < \varepsilon . \tag{3.c.3}$$

The construction of such a variety π is possible for any V [68]. If $\bar{V} = \bar{y}(x) \equiv \{\bar{y}_i(x)\}$ is any other variety, we may write

$$\begin{aligned} Y_{\bar{V}} - Y_V = & \int_D E_\Delta^\varphi[x, \bar{y}(x), \pi'(x), \bar{p}(x)] dx \\ & - \int_D E_\Delta^\varphi[x, y(x), \pi'(x), p(x)] dx \\ & + \int_D \Delta[x, \bar{y}(x), \pi'(x), \bar{p}(x)] dx \\ & - \int_D \Delta[x, y(x), \pi'(x), p(x)] dx \end{aligned} \tag{3.c.4}$$

where

$$E_\Delta^\varphi[x, y, p, \bar{p}] \equiv \varphi[x, y, \bar{p}] - \Delta[x, y, p, \bar{p}] \tag{3.c.5}$$

The first integral on the right side of (3.c.4) may not be negative because Y_V is *PQR*; since π is a variety of *class* 2, we may show in the same way as we did for proving Theorem 2.1, that there exists a $0 < \rho_1 < 1$, such that, if $P(V, \bar{V}) < \rho_1$, then the difference between the last two integrals on the right side of (3.c.4) is less than ε.

Let us consider the expression

$$\int_D | E^{\varphi}_{A}[x, y(x), \pi'(x), p(x)] | \, dx \tag{3.c.6}$$

by (3.c.5), (3.b.7) and (3.c.3), recalling the defininition of $A[x, y, p, \bar{p}]$, i.e., the definition of $L[x, y, p, \bar{p}]$, since $\pi'(x_i)$ $(i = 1, 2, \cdots, m)$ is bounded, there exists a number k, which depends upon m, n, the variety V and the diameters of the sets D_i $(i = 1, 2, \cdots, m)$, but which depends neither of π nor of ε, such that the expression (3.c.6) is less than $\varepsilon \cdot k$ ([59, vol. 1, § 11, # 142]; [60, § 3, # 9]; [30, § 3, c]). Consequently the absolute value of the integral on the right side of (3.c.4) is also less than $\varepsilon \cdot k$; hence

$$Y_{\bar{V}} > Y_V - \varepsilon(1 + k), \text{ whenever } \rho(V, \bar{V}) < \rho_1 .$$

Thus the theorem is proved.

(d) DEFINITION 3.4. Then integral I_V is said to be *asymptotically evaluable* (abbreviation: *AE*), if it is *PQR* and if there exists a function $L[x, y, p, \bar{p}]$ such that I_V is *LPQR* and if, for every positive R, there exists a function $\varphi^R[x, y, p]$ (as described in § 3.b).

REMARK 3.5. Tonelli [59, vol. 1, page 398–9] gave a procedure by which $\varphi^R[x, y, p]$ may be constructed starting from any *simple* integral $(m = n = 1)$, which is *PQR*: he thus proved that, if a *simple* integral is *PQR*, it is necessarily *AE*. Some criteria of *asymptotic evaluability* are exhibited in [30, § 2, page 140]; although it appears intuitively that every *PQR* integral is also *AE*, this fact was never proved, except in the case $(m = n = 1)$; therefore the statement of any theorem of semi-continuity in whose proof the function $\varphi_R[x, y, p]$ is used, has to contain the hypothesis that this function can be constructed, i.e., that the integral considered is *AE*.

THEOREM 3.6. *If the integral I_V is PQR and AE, it is lower semicontinuous on every ordinary variety.*

Proof. Let us first point out that existence and lowers emi-continuity on any variety of I_V, and those of the integral $\bar{I}_V$ defined by (3.b.2), are equivalent, since the integral

$$\int_D \{\Phi[x, y(x), p(x)] - \bar{\Phi}[x, y(x), p(x)]\}dx \equiv \int_D L[x, y(x), \ \Omega, p(x)]dx$$

exists and is continuous at every variety, by the Continuity Theorem 1.2.

Let $V = y(x) \equiv y_i(x_i)$ be an ordinary variety, and let $\varepsilon > 0$ be given.

Since $\bar{\Phi}[x, y, p]$ is never negative, it is possible to find a positive number R, such that, if D_i' $(i = 1, 2, \cdots, m)$ is the subset of D_i consisting of the points x_i such that, for each j $(j = 1, 2, \cdots, n)$, the partial derivative $\partial y_i(x_i)/\partial x_i^j$ exists and its absolute value does not exceed R, the inequality

$$\bar{I}_V - \int_{\Pi_1^m {}_i D_i'} \bar{\Phi}[x, y(x), p(x)]dx < \varepsilon/2 \tag{3.d.1}$$

holds.

The integral Y_V, that we associate with I_V and R (see §3.b) is lower semicontinuous on V by Lemma 3.3; i.e., there exists a positive number $\bar{\rho}$ such that, for each variety $\bar{V}$

$$Y_{\bar{V}} > Y_V - \varepsilon/2, \text{ whenever } \rho(V, \bar{V}) > \bar{\rho}\ . \tag{3.d.2}$$

From (3.b.5) and (3.d.1) we have

$$\bar{I}_V - Y_V < \varepsilon/2$$

whence, by (3.d.2)

$$\bar{I}_{\bar{V}} > \bar{I}_V - \varepsilon, \text{ whenever } \rho(V, \bar{V}) < \bar{\rho}$$

i.e., $\bar{I}_V$ is lower semicontinuous at any ordinary variety, and so is I_V.

DEFINITION 3.7. We shall say that the integral I_V is *lower semicontinuous* at a variety V, such that $I_V = +\infty$, if there exists a positive function $\rho(\varepsilon)$, defined for each positive ε, such that, if $\bar{V}$ is any ordinary variety, then

$$I_{\bar{V}} > \varepsilon, \text{ whenever } \rho(V, \bar{V}) < \rho(\varepsilon)\ .$$

THEOREM 3.8. *An integral I_V, which is PQR and AE, is lower semicontinuous at every variety V.*

In the case in which V is ordinary, Theorem 3.6 states the lower semicontinuity of I_V on V. If V is not ordinary, the value of I_V on V is $+\infty$ (see Remark 1.7).

Let us again consider $\bar{I}_V$ instead of I_V. Let ε be a given number, and let R be another positive number, such that, if $\bar{\bar{D}}_i$ $(i = 1, 2, \cdots, m)$ denotes the subset of D_i consisting of the points x_i where all the partial derivatives of $y_i(x_i)$ exist and are less than R in absolute value, then

$$\int_{\overset{m}{\underset{1}{\Pi}}_i \bar{\bar{D}}_i} \bar{\Phi}[x, y(x), p(x)]dx > \varepsilon + 1 . \tag{3.e.1}$$

Like in the proof of Theorem 3.6, we consider again $\varphi^R[x, y, p]$. Y_V exists finite and is lower semicontinuous at V: hence we may find a positive number $\bar{\rho}$, such that, if V is any variety such that

$$\rho(V, \bar{V}) < \bar{\rho} , \tag{3.e.2}$$

then

$$Y_{\bar{V}} > Y_V - 1 .$$

By (3.b.5) and (3.b.6) ,

$$\begin{aligned} Y_V &\geqq \int_{\overset{m}{\underset{1}{\Pi}}_i \bar{\bar{D}}_i} \varphi^R[x, y(x), p(x)]dx \\ &= \int_{\overset{m}{\underset{1}{\Pi}}_i \bar{\bar{D}}_i} \bar{\Phi}[x, y(x), p(x)]dx , \end{aligned}$$

hence, considering (3.e.1), if (3.e.2) is satisfied,

$$\bar{I}_{\bar{V}} > \varepsilon .$$

Therefore $\bar{I}_V$ is semicontinuous at V, and so is I_V. The theorem is thus completely proved.

Conclusion. Let us list four problems which are still open in the area of the study of the semicontinuity of the integrals of the Calculus of Variations in non-parametric form:

Problem 1. No example of any lower semicontinuous integral which is not PQR is known: it appears worth while to investigate whether or not *positive quasi-regularity* is also necessary for lower semicontinuity.

Problem 2. For proving Theorems 3.6, 3.8, we used the construction of the function $\varphi^R[x, y, p]$, and we had to assume that this construction could be made for every R (see § 3.b). It would be interesting to prove Theorem 3.8 without using this construction, i.e., dropping the hypothesis that I_V is AE.

REMARK C.1. The semicontinuity at any variety V of *class* 1, or even just such that all the functions $y_i(x_i)$ are Lipschitzian, can easily be proved for any I_V, which is PQR, without any hypothesis of *asymp-*

totic evaluability, by generalizing the procedure followed in [30, § 3, *First Theorem of Semicontinuity*].

Problem 3. No example of any integral I_V, which is PQR without being AE, is known. It would be useful to devise a general method by which it would be possible to construct $\varphi^R[x, y, p]$ from R and $\phi[x, y, p]$: thus proving that if I_V is PQR, it is necessarily AE.

Problem 4. Only varieties which are absolutely continuous in the sense of Tonelli [63] and the *m-uniform* metric were considered in this paper; however, it appears that *positively quasi-regular* integrals are lower semicontinuous even with respect to weaker metrics, on more general classes of varieties. Generalization of the results contained in this paper may be considered.

BIBLIOGRAPHY

1. C. Bloch, *On field theories with non-localized interactions*, Det Kongelige Danske Videnskabernes Selskab, Matematisk-fysiske Meddeleser, (1952).
2. S. Cinquini, *Condizioni necessarie per la semicontinuitá degli integrali doppi del Calcolo delle Variazioni*, Annali Matem. Pura ed Appl., **10** (1932), 233.
3. ______, *Nuovi teoremi di esistenza dell'estremo in campi illimitati per i problemi di ordine n*, Ann. Scuola Norm. Super. Pisa, (1930).
4. ______, *L'estremo assoluto degli integrali doppi dipendenti dalle derivate di ordine superiore*, Ann. Scuola Norm. Super. Pisa (2) **10** (1941), 215.
5. ______, *Sopra l'estremo assoluto degli integrali doppi in forma ordinaria*, Ann. Mat. Pura Appl. **30** (1949), 249.
6. M. E. Ebel, *Casual Behaviour of Field Theories with Non-Localized Interactions*, (Magazine quoted in [**1**], 1954).
7. S. Faedo, *Condizioni necessarie per la semicontinuitá di un nuovo tipo di funzionali*, Ann. Mat. Pura Appl. (4) **23** (1944), 69.
8. ______, *Un nuovo tipo di funzionali continui*, Rend. Mat. Appl. (5) **4** (1943), 1.
9. ______, *Sulle condizioni di Legendre e di Weierstrass per gli integrali di Fubini-Tonelli*, Ann. Scuola Norm. Super. **15** (1950), 127.
10. L. Giuliano, *Sulle condizioni sufficienti per la semicontinuitá degli integrali doppi del Calcolo delle Variazioni*, Ann. Scuola Norm. Super. Pisa (3) **1** (1941), 161.
11. ______, *Osservazioni sopra alcuni teoremi di semicontinuitá degli integrali doppi* (Idem, 1941).
12. ______ *La variazione prima nei problemi di Lagrange*, (Rend. di Mat. App. (1947).
13. ______, *Sulla continuitá degli integrali curvilinei del Calcolo delle Variazioni* (Nota I), (Rend. dell'Acc. Naz. Lincei, (1947).
14. ______, *Idem* (Nota II), (Idem, 1947).
15. ______, *Idem* (Nota III), (Idem, 1949).
16. ______, *Sulla continuitó degli integrali curvilinei del Calcolo delle Variazioni*, (Ann. della Scuola Normale Superiore, Pisa, (1947).
17. M. R. Hestenes, *A sufficiency Proof for Isoperimetric Problems in the Calculus of Variations*, Bull. Am. Math. Soc. (1938).
18. ______, *Sufficient Conditions for Multiple Integral Problems in the Calculus of Variations* Amer. J. Math. **70** (1948), 239.

19. _______, *Application of the theory of Quadratic Forms in Hilbert space to the Calculus of variations*, Pacific J. Math. **1** (1951), 525.
20. _______, *Sufficient Conditons for the General Problem of Mayer with Variable End. Points*, University of Chicago Press, (1952).
21. _______, *On sufficient conditions in the problems of Langrange and Bolza*, Ann. Math. **37** (1936), 543.
22. _______, *A Direct Sufficiency proof for the Problem of Bolza in the Calculus of Variations*, Trans. Amer. Math. Soc. **42** (1947), 141.
23. _______, *Generalized problem of Bolza in the calculus of variations*, Duke Math. J. **5** (1939), 309.
24. _______, *The problem of Bolza in the Calculus of Variations*, Bull. Amer. Math. Soc. **48** (1942), 50.
25. _______, *An Alternate Sufficiency Proof for the Normal Problem of Bolza*, Trans. Amer. Math. Soc. **61** (1947), 256.
26. _______, *An indirect Sufficiency Proof for the Normal Problem of Bolza in Non-Parametric Form* Trans. Amer. Math. Soc. **62** (1947), 509.
27. P. Kristensen, and C. Moller, *On a Convergent Meson Theory*, Journal quoted in [**1**], (1952).
28. A. Kryzwiki, I. Rzewuski, I. Zamorski, and A. Zieba, *Non local problems in the Calculus of Variations*, Annales Polonici Matematici, (1955).
29. A. Krzywiki, I. Rzewuski, I. Zamorski, and A. Zieba, *Idem, Second Paper*, Idem, (1957)
30. L. Lombardi, *Sulla semicontinuitá degli integrali di Fubini-Tonelli*, Annali della Scuola Normale Superiore, Pisa **312** (1958), 129-153.
31. _______, *Sull'esistenza del minimo degli integrali di Fubini-Tonelli*, Rendiconti dell'Istituto Lombardo Sci. A **92** (1958), 446-458.
32. _______, *Sur la Méthode du Gradient dans le Calcul des Variations*, Comptes-Rendus du Congrés *Les Mathématiques de l'Ingénieur*, 216-225 Mons-Bruxelles, (1958).
33. _______, *The continuity of the most general integral of the calculus of variations in non-parametric form*, J. of Mathematical Analysis and Applications (in press).
34. E. Magenes, *Intorno agli integrali di Fubini-Tonelli: condizioni sufficienti per la semicontinuitá*, Annali della Scuola Normale Superiore, Pisa (3) **2** (1948), 1.
35. _______, *Intorno agli integrali di Fubini-Tonelli: teoremi di esistenza dell'estremo*, (Idem, (3) **3** 65, 1949).
36. _______, *Sul minimo degli integrali di Fubini-Tonelli*, Giornale di Battaglini, Napoli, (1949-50).
37. _______, *Sulle condizioni di Eulero relative agli integrali di Fubini-Tonelli*, Rend. Sem. Math. Padova. **19** (1950), 62.
38. _______, *Un'osservaione sulle condizioni necessarie per la semicontinuitá degli integrali di Fubini-Tonelli*, (Idem, **17** 44, 1950).
39. _______, *Sul minimo semiforte degli integrali di Fubini-Tonelli*, (Idem, **20** 401 1950).
40. C. Moller, *Non-local Field Theory* (Intern. Conference on Elementary Particles, Bombay, (1950).
41. _______, *On the Problem of Convergence in Non-local Field Theories*, Proc. of the Intern. Congress on Theoret. Physics, Kyoto and Tokyo, (1953).
42. _______, *Motion of free particles in discontinuous gravitational fields*, Amer. J. of Physics, (1959).
43. C.B. Morrey, *Functions of several variables and absolute continuity*, Duke Math, J. **6** (1940), 187.
44. _______, *Multiple Integral Problems in the Calculus of Variations and Related Topics*, University of California Publ. in Math, (1943).
45. _______, *Quasi-convexity and the lower semicontinuity of multiple integrals and related topics*, Pacific. J. Math. **2** (1952).
46. W. Pauli, *On the Hamiltonian structure of Non-local field theories*, Nuovo Cimento, **10** (1953), 648.
47. I. Rzewuski, *On differential structure of non-local field theories*, Bull. de l'Academie

Polonaise des Sciences, **2** (1954), 429.
48. _______, *differential structure of non-local theories,* Acta Physica Polonica, **14** (1955), 121.
49. B. Schweizer, A. Sklar, *Statistical metric spaces,* Pacific. J. Math. **1** (1960).
50. _______, *Statistical metric spaces arising from sets of random variables in Euclidean n-space,* (Forthcoming).
51. G. I. Silova, *The existence of the absolute minimum of the multiple integrals of the calculus of variations,* Doklady Akad. Nauk SSSR, **102** (1955), 669.
52. _______. *Two dimensional problems of the calculus of variations,* Uspechi Akad. Nauk SSSR, **6** (1951), 16.
53. A. G. Sigalov, *Conditions for the existence of a minimum of double integrals in an unbounded region,* Doklady Ak. Nauk SSSR, **8** (1951), 741.
54. _______, *Regular Double Integrals of the Calculus of Variations in Non-Parametric Form,* (Idem, **73** 891, 1950).
55. _______, *Two dimensional problems of the calculus of variations transformed from non-parametric to parametric form,* Mat. Sbornik, **73** (1954), 385.
56. G. Stampacchia, *Sopra una classe di funzioni di due variabili; applicazioni agli integrali doppi del Calcolo delle Variazioni,* Giorn. di Battaglini, **79** (1950), 169.
57. _______, *Gli integrali doppi del Calcolo delle Variazioni in forma ordinaria,* Atti Accad. Naz. dei Lincei, VIII **8** (1950), 21.
58. E. C. G. Stueckelberg, G. Wanders, *Acausalité de l'intéraction non-locale,* Helv. Phys. Acta, **27** (1954), 667.
59. L. Tonelli, *Fondamenti di Calcolo delle Variazioni,* Zanichelli, Bologna, 1921, vol. 1 and 2).
60. _______, *Sur la semicontinuité des intégrales doubles du calcul des variations,* Acta Math., **53** (1939), 323.
61. _______, *Gli integrali del Calcolo delle Variazioni in forma ordinaria,* Annali della Scuola Superiore, Pisa, (2) **3** (1934), 401.
62. _______, *L'estremo assoluto degli integrali doppi,* (Idem, (2) **2** 89, 1933).
63. _______, *Sull'approssimazione analitica delle funzioni di piú variabili,* Rend. Circ. Mat. Palermo, **32** (1910), 297.
64. _______, *Sui massimi e minimi assoluti del Calcolo delle variazioni,* (Idem, 1911).
65. _______, *Sul caso regolare nel calcolo delle variazioni,* (Idem, **35** 49, 1913).
66. _______, *Sur une méthode directe du Calcul des Variations,* (Idem, **39** 233, 1915).
67. _______, *La semicontinuitá nel Calcolo delle Variazioni,* (Idem, **44** 167, 1920).
68. _______, *Sopra alcune proprietá di un polinomio di approssimazione,* Rend. Acc. Naz. dei Lincei, (1926).

UNIVERSITY OF CALIFORNIA, LOS ANGELES

GAME THEORETIC PROOF THAT CHEBYSHEV INEQUALITIES ARE SHARP

ALBERT W. MARSHALL AND INGRAM OLKIN

1. Summary. This paper is concerned with showing that Chebyshev inequalities obtained by the standard method are sharp. The proof is based on relating the bound to the solution of a game. An optimum strategy yields a portion of the extremal distribution, and the remainder is obtained as a solution of the relevant moment problem.

2. Introduction. Let X be a random vector taking values in $\mathscr{X} \subset R^k$, and suppose that $Ef(X) \equiv E(f_1(X), \cdots, f_r(X)) = (\varphi_1, \cdots, \varphi_r) \equiv \varphi$ is given, where f_j is a real valued function on $\mathscr{X}$. For convenience, we suppose $f_1 \equiv 1$. An upper bound for $P\{X \in \mathscr{T}\}$, $\mathscr{T} \subset \mathscr{X}$, may be obtained as follows. If $a = (a_1, \cdots, a_r) \in R^r$ and $\chi_{\mathscr{T}}$ is the indicator of $\mathscr{T}$ then $af' \geqq \chi_{\mathscr{T}}$ on $\mathscr{X}$ implies $P\{X \in \mathscr{T}\} \leqq a\varphi'$, and if $\mathscr{A}_0 = \{a\colon af' \geqq \chi_{\mathscr{T}}$ on $\mathscr{X}\}$, a "best" bound is given by

$$P\{X \in \mathscr{T}\} \leqq \inf_{a \in \mathscr{A}_0} a\varphi' . \tag{2.1}$$

In general, a bound is called sharp if it cannot be improved. For some cases, when $\mathscr{T}$ is assumed to be closed, the bound can actually be attained by a distribution satisfying the moment hypotheses.

The main result of this paper is

THEOREM 2.1. *Inequality* (2.1) *is sharp in the following cases.*

(I) $X = (X_1, \cdots, X_k)$ with EX_iX_j or EX_i and EX_iX_j given, $i, j = 1, \cdots, k$.

(II) X has range $(-\infty, \infty)$, $[0, \infty)$, or $[0, 1]$, and EX^j is given, $j = 1, \cdots, m$.

(III) X is a random angle in $[0, 2\pi)$ and the trigonometric moments $Ee^{i\alpha X}$, $\alpha = \pm 1, \cdots, \pm m$ are given.

Sharpness has been shown in (I) by Marshall and Olkin [6] when $\mathscr{T}$ is convex, and by Isii [3, 4] in the unbounded cases of (II). Sharpness has also been proved in a number of specialized situations.

In § 3 the proof for (I) will be given in detail. The necessary alterations for each of the remaining cases will be given in § 4, 5, 6, 7. The solution of certain moment problems depend on conditions on Hankel matrices, i.e., matrices of the form $H = (h_{i+j})$, and some results concerning these matrices are given in § 8.

Received August 31, 1960. Research sponsored by the Office of Naval Research at Stanford University.

The notation $A > 0\,(\geqq 0)$ is used to mean that the matrix A is symmetric and positive definite (p.s.d).

3. **The multivariate case.** The relation between inequality (2.1) and a game can be greatly simplified if we use matrix theoretic arguments. This is true in part because functions of the form af', $a \in \mathscr{A}_0$, can be written very naturally as quadratic or bilinear forms.

Let $X = (X_1, \cdots, X_k)$ be a random vector on R^k with $EX = \mu$ and moment matrix $EX'X = \Sigma$. If $u \equiv u(x) = (1, x)$ for $x \in R^k$, then $Eu'(X)u(X) = \begin{pmatrix} 1 & \mu \\ \mu' & \Sigma \end{pmatrix} = \Pi$. We assume $\Pi > 0$, for otherwise the dimensionality of X can be reduced.

Functions of the form af', $a \in \mathscr{A}_0$ can be written as uAu', A: $k+1 \times k+1$, $A \in \mathscr{A} = \{A;\, A \geqq 0,\, uAu' \geqq 1 \text{ for } x \in \mathscr{T}\}$. Hence

$$P\{X \in \mathscr{T}\} \leqq \inf_{a \in \mathscr{A}_0} a\varphi' = \inf_{A \in \mathscr{A}} \operatorname{tr} A\Pi \,. \tag{3.1}$$

Let $x_1, \cdots, x_m$ be points (row vectors) in R^k, $u_i = u(x_i)$, $p_1, \cdots, p_m$, $\Sigma p_i = 1$ be probabilities, $T = (u_1', \cdots, u_m')$, $D_p = \operatorname{diag}(p_1, \cdots, p_m)$, and $H = TD_pT'$. By $H \sim \mathscr{T}$ we mean that all $x_i \in \mathscr{T}$. The condition $uAu' \geqq 1$ for $x \in \mathscr{T}$ can then be written as $\operatorname{tr} AH \geqq 1$ for $H \sim \mathscr{T}$, so that $\mathscr{A} = \{A\colon A \geqq 0, \operatorname{tr} AH \geqq 1 \text{ for } H \sim \mathscr{T}\}$.

With this notation, we can rewrite the bound (3.1) in a form which is suggestive of a game.

$$\begin{aligned} \inf_{A \in \mathscr{A}} \operatorname{tr} A\Pi &= \inf_{\{A:\ \inf_{H \sim \mathscr{T}} \operatorname{tr} AH \geqq 1,\ A \geqq 0\}} \operatorname{tr} A\Pi \\ &= \inf_{S \geqq 0} \left(\frac{\operatorname{tr} S\Pi}{\inf_{H \sim \mathscr{T}} \operatorname{tr} SH} \right) = \left(\sup_{S \geqq 0} \inf_{H \sim \mathscr{T}} \frac{\operatorname{tr} SH}{\operatorname{tr} S\Pi} \right)^{-1} \\ &= \Big(\sup_{\{S:\, S \geqq 0,\, \operatorname{tr} SH \leqq 1\}} \ \inf_{\Pi \sim \mathscr{T}} \operatorname{tr} SH \Big)^{-1} \equiv \nu^{-1} \,. \end{aligned} \tag{3.2}$$

In view of (3.2) it is natural to consider the game $G = (\mathscr{S}, \mathscr{H}, g)$, where $\mathscr{S} = \{S\colon S \geqq 0, \operatorname{tr} S\Pi \leqq 1\}$ and $\mathscr{H} = \{H\colon H \sim \mathscr{T}\}$ are the strategy spaces for players I and II, respectively, and $g(S, H) = \operatorname{tr} SH$ is the payoff to player I.

Clearly $\mathscr{S}$ and $\mathscr{H}$ are closed and convex. Further, $\mathscr{S}$ is bounded since

$$\|S\|^2 \equiv (\operatorname{tr} SS') \leqq (\operatorname{tr} S) c_M(S) \leqq (\operatorname{tr} S)^2 \leqq (\operatorname{tr} S\Pi)^2 / c_m^2(\Pi) \leqq 1/c_m^2(\Pi) \,,$$

where $c_m(A)$, $c_M(A)$ are the minimum and maximum characteristic roots of A. For the present we assume that $\mathscr{H}$ is bounded, then by [2, Section 2.5], G has a value and there exist optimal strategies $S_0 \in \mathscr{S}$, $H_0 \in \mathscr{H}$, such that

$$\operatorname{tr} SH_0 \leqq \operatorname{tr} S_0H_0 = \nu \leqq \operatorname{tr} S_0H\,, \qquad \text{for all } S \in \mathscr{S}, H \in \mathscr{H}. \tag{3.3}$$

The optimal strategy S_0 has the property that $\inf_{A\in\mathscr{A}} \operatorname{tr} A\Pi = \operatorname{tr} A_0\Pi$, where $A_0 = S_0/\nu$.

To prove sharpness of (3.1), we must show that there exists a distribution for X such that $P\{X \in \mathscr{T}\} = 1/\nu$, and $Eu'u = \Pi$. H_0 is the moment matrix of a distribution F_1 on points in $\mathscr{T}$. If we can prove the existence of a probability distribution F for X of the form $F = F_1/\nu + F_2$, and with moment matrix Π, then this distribution attains equality in (3.1). To see this, note that F assigns at least probability ν to $\mathscr{T}$, and by (3.1) it can assign at most probability ν to $\mathscr{T}$.

To show the above, we need only show that a distribution F_2 exists with total variation $1 - 1/\nu$ and moment matrix $\Psi = \Pi - H_0/\nu$. The following Lemma yields this result.

LEMMA 3.1. *Let* $\Pi > 0$, $\mathscr{S} = \{S\colon S \geqq 0, \operatorname{tr} S\Pi \leqq 1\}$.
(i) *If* $\operatorname{tr} SH \leqq \nu$ *for all* $S \in \mathscr{S}$, *then* $\Psi = \Pi - H/\nu \geqq 0$.
(ii) *If* $\operatorname{tr} SH = \nu$ *for some* $S_0 \in \mathscr{S}$, *then* Ψ *is not strictly* > 0.
(iii) *If* $\operatorname{tr} SH < \nu$ *for all* $S \in \mathscr{S}$, *then* $\Psi > 0$.

Proof. There exists a representation $\Pi = WW'$, $H = WD_\theta W'$, $|W| \neq 0$, $D_\theta = \operatorname{diag}(\theta_0, \cdots, \theta_k)$, and hence $\Psi \geqq 0$ if and only if $\theta_i \leqq \nu$, $i = 0, \cdots, k$. If $W'SW = \begin{pmatrix}1&0\\0&0\end{pmatrix}$, then $S \in \mathscr{S}$, and from $\operatorname{tr} SH = tr\, W'SWD_\theta \leqq \nu$, we obtain $\theta_0 \leqq \nu$. Part (i) follows using permutations. If $\operatorname{tr} SH < \nu$, then in the above argument, each $\theta_i < \nu$. If $\operatorname{tr} S_0H = \operatorname{tr}(W'S_0W)D_\theta = \nu$ and $\operatorname{tr} W'S_0W \leqq 1$, then at least one of the θ_i is equa to ν.

The condition that $\mathscr{H}$ be bounded now can be removed, since $\|H_0\|^2 \leqq (\operatorname{tr} H_0)^2 \leqq [\nu \operatorname{tr} \Pi]^2$, by Lemma 3.1.

REMARK 3.1. We note that $\operatorname{tr} S_0\Pi = 1$, for if not, αS_0 for $\alpha > 1$ would violate (3.3).

S_0 and H_0 are related by $\nu S_0\Pi = S_0H_0$. This follows from the fact that $\operatorname{tr} S_0\Psi = \operatorname{tr} S_o(\Pi - H_0/\nu) = 0$ and $\Psi \geqq 0$ implies that $S_0^{1/2}\Psi S_0^{1/2} = 0$, or equivalently that $S_0^{1/2}\Psi^{1/2} = 0$, which yields the result.

REMARK 3.2. In the above development we assumed that $EX = \mu$ was given. If this is not the case, then choose $\mathscr{S} = \{S = \begin{pmatrix}\alpha&0\\0&S_1\end{pmatrix}\colon S > 0$, $\operatorname{tr} S\Pi \leqq 1\}$, $S_1\colon k \times k$, and the entire development remains unchanged with S_1 replacing S, since $S \geqq 0$ if and only if $\alpha > 0$, $S_1 \geqq 0$ and $\operatorname{tr} S\Pi = \alpha + \operatorname{tr} S_1\Sigma$.

We now summarize the essential points of the proof which are appropriately modified in each of the remaining cases.

(1) Introduce vectors $u(x)$ and $v(x)$ ($u = v$ in the above) such that
(i) $Ev'(X)u(X) = \Pi$ is a matrix of given moments,
(ii) $af', a \in \mathscr{A}_0$ can be written as uAv' with $A \in \mathscr{A}$.

To define $\mathscr{A}$ we first must characterize $\mathscr{A}_0$.

(2) Define $\mathscr{H}$, a set of moment matrices of the same kind as Π, but corresponding to distributions on $\mathscr{T}$.

(3) Define $\mathscr{S}$ and show that $\mathscr{S}$ is bounded.

(4) Use the game to assert that H_0 exists, and show that the moment problem with moments defined by $\Psi = \Pi - H_0/\nu$ has a solution with $\psi_{11} = 1 - 1/\nu$.

4. Univariate distributions on $(-\infty, \infty)$. Let $u(x) = v(x) = (1, x, \cdots, x^n)$. Then polynomials $af'(x)$ of degree $\leqq 2n$ which are nonnegatve in $(-\infty, \infty)$ can be expressed as $uAu', A \geqq 0$, [7, p. 82]. Hence $\mathscr{A} = \{A: A \geqq 0, uAu' \geqq 1 \text{ for } x \in \mathscr{T}\}$, and (3.1) holds. Note that $\Pi = (\pi_{i+j-2}) = (EX^{i+j-2})$, $i, j = 1, \cdots, n+1$. Let $-\infty < t_i < \infty$, $u_i = u(t_i)$, $i = 1, \cdots, m$, $T = (u_1', \cdots, u_m')$, $D_p = \operatorname{diag}(p_1, \cdots, p_m) \geqq 0$, $\operatorname{tr} D_p = 1$, $H = TD_pT' = (h_{i+j-2})$, $i, j = 1, \cdots, n+1$. Define $\mathscr{H} = \{H: t_i \in \mathscr{T}, i = 1, \cdots, m\}$, $\mathscr{S} = \{S: S \geqq 0, \operatorname{tr} S\Pi \leqq 1\}$. We assume that the moment problem corresponding to the given moments $\{\pi_0, \cdots, \pi_{2n}\}$ is not determined so that $\Pi > 0$, [8, Th. 3.3], and the previous argument that $\mathscr{S}$ is bounded holds. Assuming that $\mathscr{H}$ is bounded, there exists an S_0 and $H_0 = (h^0_{i+j-2})$ satisfying (3.3), and with Lemma 3.1 we conclude as before that the boundedness condition on $\mathscr{H}$ can be removed.

Since $\pi_0 = h_0^0 = 1$, $\psi_0 = 1 - 1/\nu$. Define $\Delta_r = |\psi_{i+j-2}|_{i,j=1}^{r+1}$; then since $\Psi \geqq 0$, by Theorem 8.1 it follows that $\Delta_1 > 0, \cdots, \Delta_{r-1} > 0, \Delta_r = 0, \cdots, \Delta_n = 0$, for some r. The reduced (Hamburger) moment problem has a solution if and only if $\Psi \geqq 0$, in which case there exists a (unique) representation $\psi_j = \sum_{i=1}^r p_i \xi_i^j$, $j = 0, 1, \cdots, 2n-1$, and $\psi_{2n} = \sum_{i=1}^r p_i \xi_i^{2n} + c$, $c \geqq 0$, and $c = 0$ if $r = n$, [8, p. 85].

In the event $c > 0$, by using an ε-good strategy for player II to guarantee Ψ strictly > 0, we obtain a distribution with moments $\{\pi_0, \cdots, \pi_{2n}\}$, which assigns probability $1/(\nu + \varepsilon)$ to $\mathscr{T}$.

REMARK 4.1. The representation obtained from [7, p. 82] is of the form $(\Sigma u_i c_i)^2 + (\Sigma u_i d_i)^2$, which is expressible as uAu', where $A = c'c + d'd$. However, the same class of polynomials is obtained if we include all $A \geqq 0$.

REMARK 4.2. If $\mathscr{T}$ is bounded, there exists an extremal distribution with a spectrum consisting of at most $2(n+1)$ points. This follows from the fact that the least number of points contributing to H_0 is at most $(n+1)$, [2, § 2.5], and to Ψ is at most $(n+1)$ points by the previous argument.

5. Univariate Case on $[0, \infty)$. Consider first the case $m = 2n-1$, and let $u(x) = (1, x, \cdots, x^{n-1})$, $v(x) = (1, x, \cdots, x^n)$. Then polynomials $af'(x)$ of degree $\leqq 2n-1$ can be expressed as $u[(B, 0) + (0, C)]v' \equiv uAv'$, where $B \geqq 0, C \geqq 0$ are $n \times n$ matrices (See [7, p. 82] and Remark 4.1). Hence $\mathscr{A} = \{A: B \geqq 0, C \geqq 0, \ uAv' \geqq 1 \ \text{ for } \ x \in \mathscr{T}\}$, and (3.1) holds. Now $\Pi = (\pi_{i+j-2}) = (EX^{i+j-2})$, $i = 1, \cdots, n+1; j = 1, \cdots, n$. Let $0 < t_i < \infty$, $u_i = u(t_i)$, $v_i = v(t_i)$, $i = 1, \cdots, m$, $T_1 = (u_1', \cdots, u_m')$, $T_2 = (v_1', \cdots, v_m')$, $D_p = \operatorname{diag}(p_1, \cdots, p_m) \geqq 0$, $\operatorname{tr} D_p = 1$, $H = T_1 D_p T_2' = (h_{i+j-2})$, $i = 1, \cdots, n+1; j = 1, \cdots, n$. Define $\mathscr{H} = \{H : t_i \in \mathscr{T}, \ i = 1, \cdots, m\}$, $\mathscr{S} = \{S = (S_1, S_2) : S_1 \geqq 0, S_2 \geqq 0, \operatorname{tr}[(S_1, 0) + (0, S_2)]\Pi \leqq 1\}$, $S_1, S_2 : n \times n$, $0 : n \times 1$. Assuming that the moment problem corresponding to Π is not determined, i.e., $\Pi_{(1)} = (\pi_{i+j-2})$, $i, j = 1, \cdots, n$, $\Pi^{(1)} = (\pi_{i+j-1})$, $i, j = 1, \cdots, n$, are positive definite, [8, p. 6], the argument of § 3 that $\mathscr{S}$ is bounded holds, with $||S|| \equiv ||(S_1, 0) + (0, S_2)||$.

Assuming that $\mathscr{H}$ is bounded, there exists an $S_0 = (S_{10}, 0) + (0, S_{20})$ and $H_0 = (h^0_{i+j-2})$, $i = 1, \cdots, n+1 ; j = 1, \cdots, n$, satisfying (3.3). Define $H_{0(1)}$ and $H_0^{(1)}$ in the some manner as $\Pi_{(1)}$ and $\Pi^{(1)}$. An application of Lemma 3.1 yields $\Psi_{(1)} = \Pi_{(1)} - H_{0(1)}/\nu \geqq 0$ and $\Psi^{(1)} = \Pi^{(1)} - H_0^{(1)}/\nu \geqq 0$. The boundedness condition of $\mathscr{H}$ can now be removed since $||H_0||^2 \leqq ||H_{01}||^2 + ||H_0^{(1)}||^2 \leqq \nu \, tr(\Pi_{(1)} + \Pi^{(1)})$. Also $\psi_0 = \pi_0 - h_0/\nu = 1 - 1/\nu$.

In order for the reduced (Stieltjes) moment problem to have a solution, it is necessary that both $\Psi_{(1)}$ and $\Psi^{(1)}$ be $\geqq 0$.[1]

Recall from § 4 that $\Delta_r = |\psi_{i+j-2}|$, $i, j = 1, \cdots, r+1$. Now define $\Delta_r^{(1)} = |\psi_{i+j-1}|$, $i, j = 1, \cdots, r+1$. From Theorem 8.1 it follows that either

(i) $\Delta_0 > 0, \cdots, \Delta_r > 0$, $\Delta_{r+1} = \cdots = \Delta_n = 0$ and $\Delta_0^{(1)} > 0, \cdots, \Delta_r^{(1)} > 0$, $\Delta_{r+1}^{(1)} = \cdots = \Delta_n^{(1)} = 0$, or

(ii) $\Delta_0 > 0, \cdots, \Delta_r > 0$, $\Delta_{r+1} = \cdots = \Delta_n = 0$ and $\Delta_0^{(1)} > 0, \cdots, \Delta_{r-1}^{(1)} > 0$, $\Delta_r^{(1)} = \cdots = \Delta_n^{(1)} = 0$, for some r. But these are the conditions that there exist a distribution whose spectrum consists of $r+1$ points distinct from 0 in case (i) and including 0 in (ii).

If $m = 2n$, let $u(x) = v(x) = (1, x, \cdots, x^n)$. Then polynomials $af'(x)$ of degree $\leqq 2n$ can be expressed as $v\left[B + \begin{pmatrix} 0 & C \\ 0 & 0 \end{pmatrix}\right]v'$, where $B: n+1 \times n+1$, $C: n \times n$, $B \geqq 0$, $C \geqq 0$, [7, p. 82]. The remainder of the proof is essentially the same as for the case $m = 2n-1$ above.

6. Univariate distribution on $[0, 1]$. We first deal with the case when an odd number of moments is given. Let $u(x) = (1, x, \cdots, x^{n-1})$, $v(x) = (1, x, \cdots, x^n)$. Now $\Pi = (\pi_{i+j-2}) = (EX^{i+j-2})$, $i = 1, \cdots, n+1$; $j = 1, \cdots, n$. Then polynomials $af'(x)$ of degree $\leqq 2n-1$ which are

[1] This result was communicated to the authors by S. Karlin. The proof is similar to that for the reduced Hausdorff moment problem given in [5].

nonnegative in $[0, 1]$ can be expressed as $u[(B, 0) + (0, C - B)]v' \equiv uAv'$, where B and C are $n \times n$ matrices, $B \geqq 0$, $C \geqq 0$, (See [7, p. 82] and Remark 4.1). Hence $\mathscr{A} = \{A : B \geqq 0,\ C \geqq 0,\ uAv' \geqq 1 \text{ for } x \in \mathscr{T}\}$, and (3.1) holds. We assume that the moment problem corresponding to the given moments $\{\pi_0, \cdots, \pi_{2n-1}\}$ is not determined. This means that $\Pi^{(1)} = (\pi_{i+j-1})$, $i, j = 1, \cdots, n$, and $\Pi_{(2)} = (\pi_{i+j-2} - \pi_{i+j-1})$, $i, j = 1, \cdots, n$, are both positive definite, [5, p. 55] or [8, p. 77]. (In the latter reference the conditions are presented for the interval $[-1, 1]$.)

Let $0 \leqq t_i \leqq 1$, $u_i = u(t_i)$, $v_i = v(t_i)$, $i = 1, \cdots, m$, $T_1 = (u'_1, \cdots, u'_m)$, $T_2 = (v'_1, \cdots, v'_m)$, $D_p = \operatorname{diag}(p_1, \cdots, p_m) \geqq 0$, $\operatorname{tr} D_p = 1$, $H = T_1 D_p T'_2 = (h_{i+j-2})$, $i = 1, \cdots, n+1$; $j = 1, \cdots, n$. Define $\mathscr{H} = \{H : t_i \in \mathscr{T}, i = 1, \cdots, m\}$, $\mathscr{S} = \{(S_1, S_2) : S_1 \geqq 0,\ S_2 \geqq 0,\ \operatorname{tr}[(S_1\, 0) + (0, S_2 - S_1)]\Pi \leqq 1\}$. We first show that $\mathscr{S}$ is bounded:

$$||S||^2 \equiv ||(S_1, 0) + (0, S_2 - S_1)|| \leqq 2 \operatorname{tr} S_1^2 + \operatorname{tr} S_2^2 \leqq 2(\operatorname{tr} S_1)^2 + (\operatorname{tr} S_2)^2 .$$

But $\operatorname{tr} S\Pi = \operatorname{tr} S_1 \Pi_{(2)} + \operatorname{tr} S_2 \Pi^{(1)} \leqq 1$, and $\Pi_{(2)} > 0$, $\Pi^{(1)} > 0$, so that $\operatorname{tr} S_1 \leqq 1/c_m(\Pi_{(2)})$, $\operatorname{tr} S_2 \leqq 1/c_m(\Pi^{(1)})$, and $\mathscr{S}$ is bounded.

Assuming that $\mathscr{H}$ is bounded, there exists an $S_0 = (S_{10}, 0) + (0, S_{20} - S_{10})$ and $H_0 = (h^0_{i+j-2})$, $i = 1, \cdots, n+1$; $j = 1, \cdots, n$, satisfying (3.3). Define $H_{0(2)}$ and $H_0^{(1)}$ as for $\Pi_{(2)}$ and $\Pi^{(1)}$; then an application of Lemma 3.1 yields

$$\Psi_{(2)} = \Pi_{(2)} - H_{0(2)}/\nu \geqq 0 \ , \ \Psi^{(1)} = \Pi^{(1)} - H_0^{(1)}/\nu \geqq 0 \ .$$

The boundedness condition on $\mathscr{H}$ can now be removed since $||H_0||^2 \leqq 2||H_{0(2)}||^2 + 2||H_0^{(1)}||^2 \leqq \nu \operatorname{tr}(\Pi_{(2)} + \Pi^{(1)})$. Also $\psi_0 = \pi_0 - ho/\nu = 1 - 1/\nu$.

In order for the reduced (Hausdorff) moment problem to have a solution, it is necessary that both $\Psi_{(2)}$ and $\Psi^{(1)}$ be $\geqq 0$, [5, p. 55].

If an even number of moments is given, we let $u(x) = v(x) = (1, x, \cdots, x^n)$. Now $\Pi = (\pi_{i+j-2})$, $i, j = 1, \cdots, n+1$. Polynomials $af'(x)$ of degree $\leqq 2n$ which are nonnegative in $[0, 1]$ can be expressed as $u\left[\begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & C \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & -C \end{pmatrix}\right] u' \equiv uAu'$, where B and C are $n \times n$ matrices, $B \geqq 0$, $C \geqq 0$, (See [7, p. 82] and Remark 4.1). Hence $\mathscr{A} = \{A : B \geqq 0,\ C \geqq 0,\ uAu' \geqq 1 \text{ for } x \in \mathscr{T}\}$, and (3.1) holds. We assume that the moment problem corresponding to the given moments $\{\pi_0, \cdots, \pi_{2n}\}$ is not determined. This means that Π and $\Pi_{(3)} = (\pi_{i+j-1} - \pi_{i+j})$, $i, j = 1, \cdots, n$, are positive definite, [5, p. 55] or [8, p. 77].

The remainder of the argument is analogous to the odd moment case.

REMARK 6.1. As in Remark 4.1, if $\mathscr{T}$ is bounded, there exists an extremal distribution with a spectrum consisting of at most $2(n+1)$ points. This follows from [2, § 2.5] and [5, § 17].

REMARK 6.2. A condition for the solution of the Hausdorff moment problem with an infinite number of moments is the condition that

$$\Delta^k \mu_j = \mu_j - \binom{k}{1}\mu_{j+1} + \binom{k}{2}\mu_{j+2} + \cdots + (-1)^k \mu_{j+k} \geqq 0 \ , \quad k, j = 0, 1, \cdots$$

However, this condition with $k, j = 0, 1, \cdots, n$ is not sufficient for a solution of the reduced moment problem. It is interesting to note that this condition enters naturally using an alternative formulation. Polynomials $af'(x)$ which are nonnegative in $[0, 1]$ may be represented as $\Sigma a_{ij}(1-x)^i x^j$, where $a_{ij} \geqq 0$. If we let $u(x) = (1, (1-x), \cdots, (1-x)^n)$, $v(x) = (1, x, \cdots, x^n)$, then the representation is uAv', $a_{ij} \geqq 0$. Now $\Pi = (E(1-X)^{i-1}X^{j-1}) = (\Delta^{i-1}\mu_{j-1})$, $i, j = 1, \cdots, n+1$. Using a similar development as before, $\mathscr{S} = \{S : s_{ij} \geqq 0,\ \mathrm{tr}\, S\Pi \leqq 1\}$, and from Lemma 3.1, $\Psi = \Pi - H_0/\nu = (\Delta^{i-1}\mu_{j-1}) - (\Delta^{i-1}\mu_{j-1}/\nu) \geqq 0$. Let $\psi_j = \mu_j - h_j/\nu$, $\Psi = (\Delta^{i-1}\psi_{j-1})$; we wish to show that $\Delta^{i-1}\psi_{j-1} \geqq 0$. By choosing S to have all zeros except $s_{ij} = 1/\Delta^{i-1}\pi_{j-1}$, $\mathrm{tr}\, S\Pi = 1$. The result follows after using (3.3).

7. **Random angle in** $[0, 2\pi)$. If $u(x) = v(x) = (1, e^{\mathrm{inx}}, \cdots, e^{\mathrm{inx}})$, then polynomials $af'(x)$ which are nonnegative in $[0, 2\pi)$ can be expressed as uAu', $A \geqq 0$, (See [7, p. 82] and Remark 4.1). Hence $\mathscr{A} = \{A \colon A \geqq 0,\ uAu' \geqq 1 \text{ for } x \in \mathscr{T}\}$, and (3.1) holds. Now $\Pi = (\pi_{j-k}) = (Ee^{i(j-k)X})$, $j, k = 1, \cdots, n+1$.

The proof is virtually that of § 4, noting only that the reduced trigonometric (Herglotz) moment problem has a solution if the Toeplitz matrix $\Pi > 0$. (See footnote, § 5.)

7.1. **An example.** The authors are unaware of any Chebyshev inequalities when trigonometric moments are available, and we present a simple illustration.

THEOREM 7.1. *If X is a random angle in $[0, 2\pi)$ and $E \sin X = \alpha$, $E \cos X = \beta$, then*

$$P\{2\theta < X < 2\varphi\} \geqq 1 - \frac{1 - \alpha \sin(\theta+\varphi) - \beta\cos(\theta+\varphi)}{1 - \cos(\varphi-\theta)} \ , \tag{7.1}$$

$$P\{2\theta \leqq X \leqq 2\varphi\} \leqq \frac{1 + \alpha \sin(\theta+\varphi) + \beta\cos(\theta+\varphi)}{1 + \cos(\varphi-\theta)} \ , \tag{7.2}$$

$$0 \leqq \theta \leqq \varphi \leqq \pi \ .$$

Proof. Choose $f(x) = c_1 + c_2 \sin x + c_3 \cos x$. The conditions $f(\theta+\varphi) = 0$, $f(2\theta) = f(2\varphi) = 1$ lead to (7.1), and the conditions $f(\theta+\varphi+\pi) = 0$, $f(2\theta) = f(2\varphi) = 1$ lead to (7.2).

8. Properties of Hankel matrices. In this section we obtain several properties of Hankel matrices which were required in §§ 4 and 5. These properties are known as a consequence of the solution of moment problems, but it may be of interest to present matrix theoretic proofs. We need the following preliminaries.

A matrix $U = (u_{i+j-2})$, $i, j = 1, \cdots, n$ is called a Hankel matrix. By the rth compound, $A^{(r)}$, of a matrix A: $n \times n$ we mean the matrix whose elements are the rth order minors of A arranged in lexicographic order; thus $A^{(r)} : \binom{n}{r} \times \binom{n}{r}$. The following properties of compound matrices are well-known, e.g., [1].

(8.1) Let A be symmetric. The characteristic roots of $A^{(r)}$ are the $\binom{n}{r}$ products of r characteristic roots of A. Thus, $A^{(r)} \geqq 0$ if and only if $A \geqq 0$.

$$|A^{(r)}| = |A|^{\binom{n-1}{r-1}} . \tag{8.2}$$

THEOREM 8.1. *If the Hankel matrix* $U = (u_{i+j-2})$, $i, j = 1, \cdots, r + 1$, *is* $\geqq 0$, *and if* $\Delta_r = |u_{i+j-2}|_{i,j=1}^{r} = 0$, *then* $\Delta_{r+1} = 0$.

Proof. Suppose $u_0 = 0$, then by nonnegativity of each 2×2 principal minor, it follows that $u_0 = u_1 = \cdots = u_{2n-1} = 0$, $u_{2n} \geqq 0$. But $U^{(r)} \geqq 0$ has first element 0, and hence its first row is 0, so that $\Delta_r = 0$.

THEOREM 8.2. *Let* $U = (u_{i+j-2})$, $i, j = 1, \cdots, r + 1$, $V = (u_{i+j-1})$, $i, j = 1, \cdots, r + 1$, $U \geqq 0$, $V \geqq 0$. *Then* $\Delta_r = 0 \Rightarrow \Delta_r^{(1)} = 0 \Rightarrow \Delta_{r+1} = 0$, *where* $\Delta_m = |u_{i+j-2}|$, $i, j = 1, \cdots, m$; $\Delta_m^{(1)} = |u_{i+j-1}|$, $i, j = 1, \cdots, m$.

Proof. In the rth compound $U^{(r)}$, $\Delta_r = u_{11}^{(r)} = 0$ implies that $u_{12}^{(r)} = \Delta_r^{(1)} = 0$. In the rth compound $V^{(r)}$, $\Delta_r^{(1)} = v_{11}^{(r)} = 0$, and hence all $v_{ij}^{(r)} = 0$, except possibly the last diagonal element, which is a function of u_{2r+1}. In $U^{(r+1)}$, the last column does not depend on u_{2r+1}, and its elements are the $v_{ij}^{(r)}$ which are zero. Hence $|U^{(r+1)}| = 0$, so that $\Delta_{r+1} = 0$.

9. Acknowledgment. We are grateful to Herman Rubin for some valuable discussions. He also pointed out that sharpness of Chebyshev inequalities can be proved quite generally without knowledge of moment problem solutions by an application of the Hahn-Banach extension theorem. However, the present proof provides considerable information concerning extremal distributions.

REFERENCES

1. A. C. Aitken, *Determinants and matrices*, 9th edition, Edinburgh: Oliver and Boyd, 1956.

2. David Blackwell and M. A. Girshick, *Theory of Games and Statistical Decisions*, John Wiley and Sons, New York, 1954.
3. Keiiti Isii, *On a method for generalizations of Tchebycheff's inequality*, Ann. Inst. Stat. Math., **10** (1959), 65-88.
4. ———, *Bounds on probability for non-negative random variables*, Ann. Inst. Stat. Math., **11** (1959), 89-99.
5. S. Karlin and L. S. Shapley, *Geometry of Moment Spaces*, Memoirs Amer. Math. Soc., No. 12, 1953.
6. Albert W. Marshall and Ingram Olkin, *Multivariate Chebyshev inequalities*, Ann. Math. Stat., Vol. 31 (1960), 1001-1014.
7. G. Pólya and G. Szegö, *Aufgaben und Lehrsätze aus der Analysis*, Vol. II, 2nd edition, Springer, Berlin, 1954.
8. J. A. Shohat and J. D. Tamarkin, *The problem of moments*, Mathematical Surveys Number 1, American Mathematical Society, 1943.

STANFORD UNIVERSITY AND BOEING SCIENTIFIC LABORATORIES;
UNIVERSITY OF MINNESOTA AND STANFORD UNIVESITY

PRIMITIVE ALGEBRAS WITH INVOLUTION

WALLACE S. MARTINDALE, 3RD

A well known theorem of Kaplansky ([1], p. 226, Theorem 1) states that *every primitive algebra satisfying a polynomial identity is finite dimensional over its center*. Related to this result is the following conjecture due to Herstein: *if A is a primitive algebra with involution whose symmetric elements satisfy a polynomial identity, then A is finite dimensional over its center*. Our main object in the present paper is to verify this conjecture in the special case where A is assumed to be *algebraic*. In the course of our proof we develop some results, which may be of independent interest, concerning the existence of nontrivial symmetric idempotents in primitive algebras with involution.

1. Some preliminary remarks.

In the present section we mention a few definitions and observations which we shall need in the remainder of this paper.

By the term *algebra over Φ* we shall mean an associative algebra (possibly infinite dimensional) over a field Φ. A *primitive* algebra over Φ is one which is isomorphic to a dense ring of linear transformations of a (left) vector space V over a division algebra Δ containing Φ (see [1], p. 32). The *rank* of an element a of a primitive algebra is the dimension of Va over Δ. We state without proof the following three remarks.

REMARK 1. Let A be a primitive algebra with identity 1 containing a set of nonzero orthogonal idempotents $e_1, e_2, \cdots, e_m$ such that

(a) $e_1 + e_2 + \cdots + e_m = 1$

(b) $\operatorname{rank} e_i = r_i < \infty$, $i = 1, 2, \cdots, m$.

Then the dimension of V over Δ is $\sum_{i=1}^{m} r_i < \infty$.

REMARK 2. Let A be a primitive algebra with center Z. If $za = 0$ for some $z \neq 0 \in Z$ and some $a \in A$, then $a = 0$.

REMARK 3. Let A be a primitive algebra. If a and b are nonzero elements of A, then $aAb \neq 0$. More generally, if $a_1, a_2, \cdots, a_n$ are nonzero elements of A, where n is any natural number, then

$$a_1 A a_2 A \cdots a_{n-1} A a_n \neq 0 .$$

An *I-algebra* is an algebra in which every non-nil left ideal contains a nonzero idempotent. An algebra over Φ is *algebraic* in case every

Received September 23, 1960.

element satisfies a non-trivial polynomial equation $f(t) = 0$, where $f(t) = \sum \alpha_i t^i$, $\alpha_i \in \Phi$. One can show that every algebraic algebra is an I-algebra. In the proof of this fact (see [1], p. 210, Proposition 1), however, the following sharper result is obtained.

REMARK 4. Let a be a non-nilpotent element of an algebraic algebra. Then the subalgebra $[[a]]$ generated by a contains a nonzero idempotent.

An *involution** of an algebra A over Φ is an anti-automorphism of A of period 2, that is,

$$\begin{aligned}(a + b)^* &= a^* + b^* \\ (\alpha a)^* &= \alpha a^* \\ (ab)^* &= b^* a^* \\ a^{**} &= a\end{aligned}$$

for all $a, b \in A$, $\alpha \in \Phi$. It is to be understood that in the rest of this paper the characteristic of Φ is assumed to be unequal to 2. An element a is *symmetric* if $a^* = a$; a is *skew* if $a^* = -a$. * is an involution *of the first kind* in case every central element is symmetric. * is an involution *of the second kind* in case there exists a nonzero central element which is skew. Every involution is of one of these two kinds.

2. S_n-algebras. The notion of an algebra satisfying a polynomial identity can be generalized according to the following

DEFINITION. A subspace R of an algebra A over Φ *satisfies a polynomial identity* in case there exists a nonzero element $f(t_1, t_2, \cdots, t_n)$ of the free algebra over Φ freely generated by the t_i such that

$$f(x_1, x_2, \cdots, x_n) = 0$$

for all $x_i \in R$. R will be called a *PI-subspace of degree d* if the degree d of $f(t_1, t_2, \cdots, t_n)$ is minimal.

The element $f(t_1, t_2, \cdots, t_n)$ is *multilinear of degree n* if and only if it is of the form

$$\sum_\sigma \alpha(\sigma) t_{\sigma_1} t_{\sigma_2} \cdots t_{\sigma_n},\ \alpha(\sigma) \in \Phi,\ \text{some } \alpha(\sigma) \neq 0\ ,$$

where σ ranges over all the permutations of $(1, 2, \cdots, n)$.

LEMMA 1. *Let R be a PI-subspace of degree n of an algebra A. Then R satisfies a multilinear polynomial identity of degree n.*

This lemma is a slight generalization of [1], p. 225, Proposition 1.

The same proof carries over directly and we therefore omit it.

Our main purpose in this paper is to study algebras of the following type.

DEFINITION. Let A be an algebra with an involution * over Φ. Suppose that the set S of symmetric elements is a PI-subspace of degree $\leq n$. Then A will be called an *S_n-algebra.* In case * is of the first (second) kind, we shall refer to A as an S_n-algebra *of the first (second) kind.*

It is surprisingly easy to analyze S_n-algebras of the second kind, as indicated by

THEOREM 1. *Let A be a primitive S_n-algebra of the second kind. Then A is finite dimensional over its center.*

Proof.[1] According to Lemma 1 S satisfies a multilinear polynomial identity of degree $n: f(t_1, t_2, \cdots, t_n) = 0$. Let z be a nonzero central element of A which is skew. If k is skew, then

$$(zk)^* = k^*z^* = (-k)(-z) = kz = zk ,$$

and hence zk is symmetric. Therefore we have

$$0 = f(zk_1, s_2, s_3, \cdots, s_n) = zf(k_1, s_2, s_3, \cdots, s_n)$$

for all $k_1 \in K$, $s_i \in S$, where K is the set of skew elements. By Remark 2 $f(k_1, s_2, s_3, \cdots, s_n) = 0$. It follows that $f(x_1, s_2, s_3, \cdots, s_n) = 0$ for all $x_1 \in A$, $s_i \in S$, since every $x \in A$ can be written $x = s + k$, $s \in S$, $k \in K$. Continuing in this fashion we finally have $f(x_1, x_2, \cdots, x_n) = 0$ for all $x_i \in A$. The conclusion then follows from the previously mentioned theorem of Kaplansky ([1], p. 226, Theorem 1).

3. Some basic theorems.

The assumption that the symmetric elements of an S_n-algebra satisfy a polynomial identity is used chiefly to prove

THEOREM 2. *Let A be a primitive S_n-algebra over Φ. Then there exist at most n orthogonal non-nilpotent symmetric elements.*

Proof. Suppose $s_1, s_2, \cdots, s_{n+1}$ are $n + 1$ orthogonal non-nilpotent symmetric elements. Using Remark 3 and the fact that the s_i are non-nilpotent we may choose elements $x_1, x_2, \cdots, x_n \in A$ so that

$$s_1^2x_1s_2^2x_2 \cdots s_n^2x_ns_{n+1} \neq 0 .$$

[1] A similar proof was communicated orally to the author by I. N. Herstein.

Now set $u_i = s_i x_i s_{i+1} + s_{i+1} x_i^* s_i$, $i = 1, 2, \cdots, n$. By Lemma 1 S satisfies a multilinear identity of degree n:

$$f(t_1, t_2, \cdots, t_n) = t_1 t_2 \cdots t_n + \sum_{\sigma \neq I} \alpha(\sigma) t_{\sigma_1} t_{\sigma_2} \cdots t_{\sigma_n}\,, \tag{1}$$

where σ ranges over all the permutations of $(1, 2, \cdots, n)$ except the identity permutation I. $f(u_1, u_2, \cdots, u_n) = 0$ since the u_i are symmetric. To analyze the right hand side of (1) we first note that if $u_i u_j u_k \neq 0$, i, j, k distinct, then either $j = i + 1$ and $k = i + 2$, or $j = i - 1$ and $k = i - 2$, because of the orthogonality of the s_i. It follows that

$$f(u_1, u_2, \cdots, u_n) = u_1 u_2 \cdots u_n + \alpha u_n u_{n-1} \cdots u_1$$

for some $\alpha \in \Phi$. Hence

$$0 = s_1 x_1 s_2^2 x_2 s_3^2 x_3 \cdots s_n^2 x_n s_{n+1} + \alpha s_{n+1} x_n^* s_n^2 x_{n-1}^* \cdots s_2^2 x_1^* s_1\,. \tag{2}$$

Multiplying (2) through on the left by s_1, we have $0 = s_1^2 x_1 s_2^2 x_2 \cdots s_n^2 x_n s_{n+1}$, a contradiction.

An idempotent e of an algebra A is called *non-trivial* in case $e \neq 1$ (if A has an identity) and $e \neq 0$.

THEOREM 3. *Let A be a primitive I-algebra with an involution*. Then:*

(a) *If there exists an $x \neq 0 \in A$ such that $xx^* = 0$, then either A contains a non-trivial symmetric idempotent or A is isomorphic to the total matrix ring Δ_2, where Δ is a division algebra. In the latter case $E_{11}^* = E_{22}$, where the E_{ij} are the unit matrices, $i, j = 1, 2$.*

(b) *If $xx^* \neq 0$ for all $x \neq 0 \in A$, then either A is a division algebra or A contains a non-nilpotent symmetric element which has no inverse in A. If $xx^* \neq 0$ for all $x \neq 0 \in A$ and A is algebraic over Φ, then either A is a division algebra or A contains a non-trivial symmetric idempotent.*

Proof. Suppose first that there exists an $x \neq 0 \in A$ such that $xx^* = 0$. We can choose an $a \in A$ such that $e = ax$ is a nonzero idempotent, because A is an I-algebra. Since $xx^* = 0$, $e \neq 1$. From the equations $ee^* = (ax)(ax)^* = axx^*a^* = 0$ it is easy to check that $e + e^* - e^*e$ is a non zero symmetric idempotent. We may thus assume that $1 \in A$ and $e + e^* - e^*e = 1$. eAe is a primitive I-algebra ([1], p. 48, Proposition 1, and p. 211, Proposition 2). If eAe is not a division algebra, then it contains an idempotent $f = ebe$, $f \neq 0$, $f \neq e$. Since $ff^* = ebee^*b^*e^* = 0$, $f + f^* - f^*f$ is a nonzero symmetric idempotent. It is unequal to 1 since otherwise $e = e(f + f^* - f^*f) = f$. We may therefore assume that eAe is a division algebra and consequently that rank $e = 1$. Since $(1 - e^*)(1 - e) = 1 - (e + e^* - e^*e) = 0$, a repetition of the above argument allows us to assume that $1-e$ is also an idempotent of rank 1. It follows from Remark 1 that A is the complete ring of linear transformations of a two dimensional vector space V over a division algebra Δ.

If $e^*e=0$ as well as $ee^*=0$ it is easy to show that relative to a suitable basis of V $e=E_{11}$ and $e^*=E_{22}$. In this case we are finished. Therefore suppose $e^*e\neq 0$. We shall sketch an argument, leaving some details to the reader, whereby a non-trivial symmetric idempotent can now be found. First find a basis (u_1, u_2) of V such that $u_1e=u_1$, $u_2e=0$, $u_1e^*=0$, $u_2e^*=\lambda u_1+u_2$, where $\lambda\neq 0\in\Delta$. By setting $v_1=\lambda^{-1}u_1$ and $v_2=u_2$ we obtain a basis (v_1, v_2) of V relative to which $e=E_{11}$ and $e^*=E_{21}+E_{22}$. From this we have

$$\begin{aligned}
E_{11}^* &= E_{21}+E_{22}\\
E_{21}^* &= [(E_{21}+E_{22})E_{11}]^*=(E_{21}+E_{22})E_{11}=E_{21}\\
E_{22}^* &= e-E_{21}^*=E_{11}-E_{21}\ .
\end{aligned}$$

Set $E_{12}^*=\alpha E_{11}+\beta E_{12}+\gamma E_{21}+\delta E_{22}$, $\alpha, \beta, \gamma, \delta\in\Delta$. From the following three equations

$$\begin{aligned}
E_{11}-E_{21} &= E_{22}^*=(E_{21}E_{12})^*=E_{12}^*E_{21}^*=\beta E_{11}+\delta E_{21}\\
E_{21}+E_{22} &= E_{11}^*=(E_{12}E_{21})^*=E_{21}^*E_{12}^*=\alpha E_{21}+\beta E_{22}\\
\alpha E_{11}+\beta E_{12}+\gamma E_{21}+\delta E_{22} &= E_{12}^*=(E_{11}E_{12})^*=E_{12}^*E_{11}^*\\
&= \beta E_{11}+\beta E_{12}+\delta E_{21}+\delta E_{22}
\end{aligned}$$

we obtain $\alpha=1$, $\beta=1$, $\gamma=-1$, and $\delta=-1$. Hence

$$E_{12}^*=E_{11}+E_{12}-E_{21}-E_{22}$$

and $-E_{12}E_{12}^*=E_{11}+E_{12}$ is then a non-trivial symmetric idempotent.

There remains the case in which $xx^*\neq 0$ for all $x\neq 0\in A$. We note that in this situation there exist no nonzero nilpotent symmetric elements, for, if $s\neq 0$ is symmetric, then $s^2=ss^*\neq 0$. If A is not already a division algebra then we can find an element $x\neq 0\in A$ such that xA is a proper right ideal. It follows that $xx^*A\subseteq xA$ is also a proper right ideal, and so xx^* is a nonzero, and hence, non-nilpotent symmetric element which has no inverse. In case A is algebraic over Φ the subalgebra $[[xx^*]]$ generated by xx^* contains a non-trivial symmetric idempotent, by Remark 4.

4. Total matrix rings with involution.

We begin by proving

THEOREM 4. *Let A be the total matrix ring Δ_m with an involution $*$, where Δ is a division algebra over Φ. Then there exists a set of orthogonal symmetric elements $e_1, e_2, \cdots, e_{m_1}, f_1 f_2, \cdots, f_{m_2}$ such that:*

(a) *The e_i are non-nilpotent elements of rank 1. In case A is* *algebraic over Φ, the e_i are idempotents of rank* 1.

(b) *The f_j are idempotents of rank* 2, *and f_jAf_j is isomorphic to Δ_2, with $E_{11}^* = E_{22}$ (see Theorem* 3).

(c) $m_1 + 2m_2 = m$.

Proof. Let $s_1, s_2, \cdots, s_h$ be a set of nonzero orthogonal symmetric idempotents, with h maximal. By the maximality of h we have

$$s_1 + s_2 + \cdots + s_h = 1 .$$

Each s_iAs_i may itself be regarded as a total matrix ring Δ_{r_i} with an involution induced by *, where r_i is the rank of s_i. We first consider those s_iAs_i having the property: there exists an $x \neq 0 \in s_iAs_i$ such that $xx^* = 0$. Theorem 3, together with the maximality of h, then says that s_iAs_i is isomorphic to Δ_2, with $E_{11}^* = E_{22}$. Relabeling these s_i as $f_1, f_2, \cdots, f_{m_2}$, we have taken care of (b).

The remaining s_i, of course, have the property that $xx^* \neq 0$ for all $x \neq 0 \in s_iAs_i$. As we have noted before, s_iAs_i can have no nonzero nilpotent symmetric elements, since $xx^* \neq 0$. Consider a typical s_iAs_i and select from it an element x_1 of rank 1. Then $y_1 = x_1x_1^* \neq 0$ is a non-nilpotent symmetric element of rank 1. Now assume that $k\,(<r_i)$ orthogonal non-nilpotent symmetric elements $y_1, y_2, \cdots, y_k$ of rank 1 have been found. Since the dimension of $W = \sum_{i=1}^k Vy_i$ is less than r_i, we can find an element x_{k+1} of rank 1 such that $Wx_{k+1} = 0$. Then $y_{k+1} = x_{k+1}x_{k+1}^*$ is a non-nilpotent symmetric element of rank 1 such that $Wy_{k+1} = 0$, that is, $y_iy_{k+1} = 0$, $i = 1, 2, \cdots, k$. Also $y_{k+1}y_i = 0$, $i = 1, 2, \cdots, k$, since $(y_{k+1}y_i)^* = y_i^*y_{r+1}^* = y_iy_{k+1} = 0$. It follows that there exists in s_iAs_i a set of r_i non-nilpotent orthogonal symmetric elements $y_1, y_2, \cdots, y_{r_i}$, each of rank 1. If A is algebraic over Φ the subalgebra $[[y_j]]$ generated by each y_j contains a nonzero idempotent z_j (necessarily of rank 1), and so we have r_i orthogonal symmetric idempotents $z_1, z_2, \cdots, z_{r_i}$, each of rank 1. Repeating the argument for all the s_iAs_i and labeling either all the y_j or all the z_j as $e_1, e_2, \cdots, e_{m_1}$, we have completed the proof of (a). (c) follows readily from the fact that rank $e_i = 1$, rank $f_j = 2$, and $\sum_i e_i + \sum_j f_j = 1$.

To illustrate Theorem 4 we consider the following simple example. Let $A = \Phi_2$, where Φ is a field, and define an involution * in A by:

$$\begin{pmatrix} \alpha_1 & \alpha_2 \\ \alpha_3 & \alpha_4 \end{pmatrix}^* = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \alpha_1 & \alpha_3 \\ \alpha_2 & \alpha_4 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \alpha_i \in \Phi .$$

The reader may verify that A contains no symmetric elements of rank 1. Similar examples of higher dimension can also be given.

In the remainder of this section we derive a result which will enable us, at least in the algebraic case, to "pass" from the total matrix ring

Δ_m to the division algebra Δ itself.

LEMMA 2. *Let A be the total matrix ring Δ_2, algebraic over Φ, with an involution* $*$, *where Δ is a division algebra over Φ. Suppose $E_{11}^* = E_{22}$. Then one of the following two possibilities must hold*:

(a) *A contains a symmetric idempotent of rank* 1.

(b) *The involution* $*$ *in Δ_2 is of the form*:

$$\begin{pmatrix}\alpha_1 & \alpha_2\\ \alpha_3 & \alpha_4\end{pmatrix}^* = \begin{pmatrix}0 & -\beta^{-1}\\ \beta^{-1} & 0\end{pmatrix}\begin{pmatrix}\bar\alpha_1 & \bar\alpha_3\\ \bar\alpha_2 & \bar\alpha_4\end{pmatrix}\begin{pmatrix}0 & \beta\\ -\beta & 0\end{pmatrix}$$

for all $\alpha_i \in \Delta$, some $\beta \neq 0 \in \Delta$, where $\alpha \to \bar\alpha$ is an involution in Δ.

Proof. It is well known (see for example [2], p. 24, Theorem 9) that the involution $*$ in A has the form:

$$\begin{pmatrix}\alpha_1 & \alpha_2\\ \alpha_3 & \alpha_4\end{pmatrix}^* = U^{-1}\begin{pmatrix}\bar\alpha_1 & \bar\alpha_3\\ \bar\alpha_2 & \bar\alpha_4\end{pmatrix}U$$

where $U = \begin{pmatrix}\gamma & \beta\\ \pm\bar\beta & \delta\end{pmatrix}$ is a nonsingular element of Δ_2 and $\alpha \to \bar\alpha$ is an involution in Δ. Consider the equation $E_{22} = E_{11}^* = U^{-1}E_{11}U$, that is,

$$\begin{pmatrix}\gamma & \beta\\ \pm\bar\beta & \delta\end{pmatrix}\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix} = \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}\gamma & \beta\\ \pm\bar\beta & \delta\end{pmatrix}.$$

It follows that $\gamma = \delta = 0$, and hence $U = \begin{pmatrix}0 & \beta\\ \pm\bar\beta & 0\end{pmatrix}$.

At this point we observe that an element $\begin{pmatrix}\gamma_1 & \gamma_2\\ \gamma_1 & \gamma_2\end{pmatrix} \in A$ is a non-nilpotent element of rank 1, unless $\gamma_1 + \gamma_2 = 0$. Now set $B = \begin{pmatrix}\pm\bar\beta & \beta\\ \pm\bar\beta & \beta\end{pmatrix}$. It is easy to check that $B^* = U^{-1}\begin{pmatrix}\pm\beta & \pm\beta\\ \bar\beta & \bar\beta\end{pmatrix}U = \pm B$, and hence B is either symmetric or skew. If $\beta \pm \bar\beta = 0$, i.e., $U = \begin{pmatrix}0 & \beta\\ -\beta & 0\end{pmatrix}$, we are finished. Therefore assume that $\beta \pm \bar\beta \neq 0$. We then apply the observation made at the beginning of this paragraph to conclude that B is a non-nilpotent element of rank 1. Since B is either symmetric or skew, it follows that B^2 is a non-nilpotent symmetric element of rank 1. The proof is complete when we note that, as A is algebraic over Φ, the subalgebra $[[B^2]]$ generated by B^2 over Φ contains a symmetric idempotent of rank 1.

THEOREM 5. *Let A be the total matrix ring Δ_m, algebraic over Φ, with an involution* $*$, *where Δ is a division algebra over Φ. Then there exists a division subalgebra D of A such that $D^* = D$ and D is isomorphic to Δ.*

Proof. Theorem 4 asserts the existence of either (a) a symmetric idempotent e of rank 1 or (b) a symmetric idempotent f of rank 2, where fAf is isomorphic to Δ_2 with the induced involution * such that $E_{11}^* = E_{22}$. In case (a) we merely set $D = eAe$ and the required conclusion follows. In case (b) Δ_2 satisfies the hypothesis of Lemma 2. If Δ_2 contains a symmetric idempotent of rank 1 we proceed as in case (a). Otherwise we conclude from Lemma 2 that the involution * in Δ_2 is given by:

$$\begin{pmatrix} \alpha_1 & \alpha_2 \\ \alpha_3 & \alpha_4 \end{pmatrix}^* = \begin{pmatrix} 0 & -\beta^{-1} \\ -\beta^{-1} & 0 \end{pmatrix}\begin{pmatrix} \bar{\alpha}_1 & \bar{\alpha}_3 \\ \bar{\alpha}_2 & \bar{\alpha}_4 \end{pmatrix}\begin{pmatrix} 0 & \beta \\ -\beta & 0 \end{pmatrix}.$$

Let D be the division subalgebra of Δ_2 consisting of all elements of the form $\begin{Bmatrix} \alpha & 0 \\ 0 & \alpha \end{Bmatrix}$, $\alpha \in \Delta$. D is obviously isomorphic to Δ. Furthermore, one verifies that

$$\begin{Bmatrix} \alpha & 0 \\ 0 & \alpha \end{Bmatrix}^* = \begin{Bmatrix} \beta^{-1}\bar{\alpha}\beta & 0 \\ 0 & \beta^{-1}\bar{\alpha}\beta \end{Bmatrix} \in D$$

and we see that $D^* = D$.

5. Division S_n-algebras.

We begin this section by stating

LEMMA 3. *Let Δ be an algebraic division algebra over its center Φ for which there exists a fixed integer h such that the dimension of $\Phi(x)$ over Φ is equal to or less than h for every separable element $x \in \Delta$. Then Δ is finite dimensional over Φ.*

Except for the restriction of separability, this lemma is virtually the same as [1], p. 181, Theorem 1. The proof appearing in [1] carries over directly, and we therefore omit it.

LEMMA 4. *Let Δ be an algebraic S_n-division algebra of the first kind over its center Φ. Suppose E is a finite dimensional field extension of Φ. Then $E \otimes_\Phi \Delta$ is isomorphic to the total matrix ring Γ_m, where Γ is a division algebra and $m \leqq 2n$.*

Proof. $E \otimes \Delta$ is well known to be a simple algebra over Φ with minimum condition on right ideals. Hence $E \otimes \Delta$ is isomorphic to Γ_m, where Γ is a division algebra and m is a natural number.

An involution τ can be defined in $E \otimes \Delta$ as follows:

$$(\alpha \otimes x)^\tau = \alpha \otimes x^*$$

for $\alpha \in E$, $x \in \Delta$. It can be verified that τ is a well-defined involution and that every symmetric element (under τ) in $E \otimes \Delta$ can be written in the form:

$$\sum_i \alpha_i \otimes s_i,\ \alpha_i \in E,\ s_i \in S\ . \tag{3}$$

Let $f(t_1, t_2, \cdots, t_n) = 0$ be the multilinear polynomial identity of degree n satisfied by S. Because this identity is multilinear and because E is the center of $E \otimes \Delta$, it follows from (3) that the set of symmetric elements of $E \otimes \Delta$ under τ also satisfies $f(t_1, t_2, \cdots, t_n) = 0$.

Now regard $E \otimes \Delta$ as the total matrix ring Γ_m, with involution τ. By Theorem 4 there exists in Γ_m a set of at least k non-nilpotent orthogonal symmetric elements, where $2k \geqq m$. Theorem 2 tells us that $k \leqq n$, and hence $m \leqq 2k \leqq 2n$.

We are now able to prove

THEOREM 6. *Let Δ be an algebraic S_n-division algebra. Then Δ is finite dimensional over its center.*

Proof. By Theorem 1 we may assume that Δ is an S_n-algebra of the first kind over its center Φ. Suppose Δ is not finite dimensional over Φ. Then by Lemma 3 there exists a separable element $x \in \Delta$ whose minimal polynomial $g(t)$ over Φ has degree $r > 2n$. Let E be a finite dimensional field extension of Φ containing the r distinct roots $\alpha_1, \alpha_2, \cdots, \alpha_r$ of $g(t)$.

We claim now that the element $x - \alpha_i$ is a zero divisor in $E \otimes \Delta$, $i = 1, 2, \cdots, r$. Indeed,

$$0 = g(x) = \prod_{j=1}^{r} (x - \alpha_j) = (x - \alpha_i) \prod_{j \neq i} (x - \alpha_j)\ ,$$

and it suffices to show that $\prod_{j\neq i}(x - \alpha_j)$ is a nonzero element of $E \otimes \Delta$. Suppose $\prod_{j\neq i}(x - \alpha_j) = 0$, that is,

$$(x^{r-1} \otimes 1) - (x^{r-2} \otimes \sum_{j\neq i} \alpha_j) + \cdots \pm (1 \otimes \prod_{j\neq i} \alpha_j) = 0\ . \tag{4}$$

Since $x^{r-1}, x^{r-2}, \cdots, 1$ are linearly independent over Φ, all the corresponding terms of E in (4) must be zero, which is clearly impossible. Therefore $x - \alpha_i$ is a zero divisor in $E \otimes \Delta$.

According to Lemma 4 $E \otimes \Delta$ is isomorphic to the total matrix ring Γ_m, where $m \leqq 2n$. We may therefore regard $E \otimes \Delta$ as the complete ring of linear transformations of an m-dimensional vector space V over the division algebra Γ. Set $V_i = \{v \in V \mid v(x - \alpha_i) = 0\}$, $i = 1, 2, \cdots, r$. V_i is a nonzero subspace of V since $x - \alpha_i$ is a zero divisor in $E \otimes \Delta$. Using the fact that the α_i are distinct elements belonging to the center E, we have that V_i are independent subspaces of V. It follows that

$$m \geqq \dim \sum_{i=1}^{r} V_i = \sum_{i=1}^{r} (\dim V_i) \geqq r > 2n .$$

A contradiction now arises since $m \leqq 2n$. We must therefore conclude that Δ is finite dimensional over its center.

6. Primitive S_n-algebras. We are now in a position to proceed with the proof of our main result.

THEOREM 7. *Let A be a primitive algebraic S_n-algebra. Then the center of A is a field, and A is finite dimensional over its center.*

Proof. Since A is primitive, A may be regarded as a dense ring of linear transformations of a vector space V over a division algebra Δ. According to Theorem 2 there exist at most n orthogonal symmetric idempotents. Let $e_1, e_2, \cdots, e_m$ be a set of m orthogonal symmetric idempotents, with $m(\leqq n)$ maximal. For each i, e_iAe_i is again a primitive algebraic algebra with involution induced by *. The same is true for $(1-e)A(1-e)$, where $e = e_1 + e_2 + \cdots + e_m$, if A should not already happen to have an identity. We now use Theorem 3 in conjunction with the maximality of m to assert that the rank of each e_i is 1 or 2, and that A does have an identity $1 = e_1 + e_2 + \cdots + e_m$. It follows that the dimension k of $V \leqq 2m$ and consequently that A is isomorphic to the total matrix ring Δ_k. The center of A is, of course, a subfield of Δ. Theorem 5 now says that Δ is an algebraic S_n-division algebra. By Theorem 6 Δ is finite dimensional over its center. Hence A is finite dimensional over its center.

COROLLARY. *Let A be a primitive algebraic algebra with an involution * such that the set K of skew elements is a PI-subspace of degree n. Then A is finite dimensional over its center.*

Proof. Let $f(t_1, t_2, \cdots, t_n) = 0$ be the multilinear polynomial identity of degree n satisfied by K, according to Lemma 1. If $s_1, s_2 \in S$, where S is the set of symmetric elements of A, then $s_1s_2 - s_2s_1 \in K$. From this it follows that $f(u_1v_1 - v_1u_1, u_2v_2 - v_2u_2, \cdots, u_nv_n - v_nu_n) = 0$ is a nontrivial polynomial identity of degree $2n$ satisfied by the elements of S. In other words, A is a primitive algebraic S_{2n}-algebra, and the conclusion follows from Theorem 7.

Note. Herstein's original conjecture was: if A is a simple ring (or algebra) with involution whose skew elements satisfy a polynomial identity, then A is finite dimensional over its center. In this paper we have verified his conjecture in the special case where A is a simple algebraic algebra which is not a nil algebra.

BIBLIOGRAPHY

1. N. Jacobson, *Structure of rings*, Amer. Math. Soc. Colloquium Publications, **37** (1956).

2. ———, *The theory of rings*, Amer. Math. Soc. Math. Surveys, vol. 2, New York 1943.

UNIVERSITY OF CHICAGO
SMITH COLLEGE

DECOMPOSITION OF HOLOMORPHS

W. H. MILLS

Let G be a group, and let H be its holomorph. There are two situations in which H is known to be decomposable into the direct product of two proper subgroups. If G is the direct product of two of its proper characteristic subgroups, say G_1 and G_2, then H is the direct product of the holomorphs of G_1 and G_2. If G is a complete group, then H is the direct product of G and G^*, where G^* is the centralizer of G in H. In this paper we will show that if G is not the direct product of two proper characteristic subgroups, and if G is not complete, then H is indecomposable. Thus we have a complete characterization of those groups whose holomorphs are indecomposable.

A decomposition of H into the direct product of indecomposable factors is known for the case where G is a finite abelian group [1]. Our present results enable us to generalize this and give a decomposition of H into the direct product of indecomposable factors, whenever G is the direct product of a finite number of characteristically indecomposable characteristic subgroups. In particular this gives a complete decomposition of H whenever G is a finite group.

Peremans [2] has shown that a necessary and sufficient condition for G to be a direct factor of H is that G be either complete or the direct product of a group of order two and a complete group that has no subgroups of index two. This result is related to the present paper. In fact Peremans' result can be deduced from Lemma 1*.

1. Preliminaries. Let G be a group, and let A be the group of all automorphisms of G. Let e and I denote the identities of G and A respectively. The holomorph H of G can be regarded as the semi-direct product of G and A, i.e., the set of all pairs (g, σ), $g \in G, \sigma \in A$, with multiplication defined by

$$(g, \sigma)(h, \tau) = (g(\sigma h), \sigma\tau) .$$

We identify g in G with (g, I) in H. Then H is a group that contains G as an invariant subgroup, and every automorphism of G can be extended to an inner automorphism of H.

For all a in G we let λ_a denote the inner automorphism of G corresponding to the element a. Thus $\lambda_a g = aga^{-1}$.

All the results of this paper depend on the following lemma:

LEMMA 1. *Let $H = H_1 \times H_2$. Then $G \cap H_1$ and $G \cap H_2$ are characteristic subgroups of G and*

Received November 1, 1960.

$$G = (G \cap H_1) \times (G \cap H_2) .$$

Proof. We note first that $G \cap H_1$ and $G \cap H_2$ are normal subgroups of H, and hence they are characteristic subgroups of G.

For $i = 1$ or 2, let ε_i denote the projection of H onto H_i corresponding to the decomposition $H = H_1 \times H_2$. Thus if $\alpha \in H_1$ and $\beta \in H_2$, then $\varepsilon_1(\alpha\beta) = \alpha$ and $\varepsilon_2(\alpha\beta) = \beta$. Put $J_i = \varepsilon_i G$. Clearly $J_i \subseteqq H_i$ and J_i is a normal subgroup of H. Let F_i and S_i denote the set of all first and second components respectively of elements of J_i. Thus $F_i \subseteqq G$ and $S_i \subseteqq A$.

Let (a, σ) be an element of J_1. Then for some g in G we have $\varepsilon_1 g = (a, \sigma)$. Put $\varepsilon_2 g = (b, \tau)$. Then $g = (a, \sigma)(b, \tau)$. Therefore $\tau = \sigma^{-1}$ and $(\sigma b^{-1}, \sigma) = (b, \tau)^{-1} \in J_2$. Hence $\sigma \in S_2$. It follows that $S_1 \subseteqq S_2$. By symmetry $S_2 \subseteqq S_1$, and hence $S_1 = S_2$.

Let σ be an element of S_1 and let ξ be an element of A. Put $\varepsilon_i(e, \xi) = (g_i, \xi_i)$, $i = 1, 2$. For some a and c in G we have $(a, \sigma) \in J_1$ and $(c, \sigma) \in J_2$. Now

$$(a, \sigma)(g_2, \xi_2) = (g_2, \xi_2)(a, \sigma)$$

and

$$(c, \sigma)(g_1, \xi_1) = (g_1, \xi_1)(c, \sigma) .$$

Comparing second components we see that σ commutes with both ξ_1 and ξ_2. Since $\xi = \xi_1\xi_2$, we have $\sigma\xi = \xi\sigma$. It follows that S_1 is contained in the center of A.

Let (a, σ) be an element of J_1 and let (d, μ) be an element of J_2. Since σ is contained in the center of A and since $(a, \sigma)^{-1} = (\sigma^{-1}a^{-1}, \sigma^{-1})$, it follows that

$$d(a, \sigma)d^{-1}(e, \lambda_{\sigma d})(a, \sigma)^{-1}(e, \lambda_{\sigma d})^{-1} = d(\sigma d)^{-1} .$$

Therefore $d(\sigma d)^{-1} \in H_1$. Moreover

$$d(\sigma d)^{-1} = (d, \mu)(e, \sigma)(d, \mu)^{-1}(e, \sigma)^{-1} \in H_2 .$$

Hence $d(\sigma d)^{-1} \in H_1 \cap H_2$. This gives us $d(\sigma d)^{-1} = e$ and $\sigma d = d$. Thus σ leaves every element of F_2 fixed. By symmetry, since $\sigma \in S_1 = S_2$, it follows that σ leaves every element of F_1 fixed. Now let g be an arbitrary element of G. Then $g = (f, \nu)(h, \zeta)$ with $(f, \nu) \in J_1$ and $(h, \zeta) \in J_2$. Since $g = f(\nu h)$, $\sigma f = f$, and $\sigma\nu h = \nu\sigma h = \nu h$, it follows that $\sigma g = g$. Hence $\sigma = I$. Therefore S_1 and S_2 consist of the identity alone. It follows that $J_1 \subseteqq G \cap H_1$, $J_2 \subseteqq G \cap H_2$, and

$$G \subseteqq J_1 \times J_2 \subseteqq (G \cap H_1) \times (G \cap H_2) \subseteqq G .$$

Therefore $G = (G \cap H_1) \times (G \cap H_2)$ and the proof is complete.

2. Some known results. Suppose $G = G_1 \times G_2 \times \cdots \times G_n$, where the G_i are characteristic subgroups of G. Let A_i denote the group of all automorphisms of G_i. We identify σ_i in A_i with the element σ_i' in A such that

$$\sigma_i' g = \begin{cases} g & \text{if } g \in G_j, j \neq i, \\ \sigma_i g & \text{if } g \in G_i. \end{cases}$$

Then $A = A_1 \times A_2 \times \cdots \times A_n$. Moreover H_i, the holomorph of G_i, becomes a subgroup of H, and $H = H_1 \times H_2 \times \cdots \times H_n$.

The centralizer of a group in its holomorph is called its conjoint. The conjoint G^* of G consists of the elements (g^{-1}, λ_g), $g \in G$. The mapping η defined by

$$\eta(g, \sigma) = (g^{-1}, \lambda_g \sigma)$$

is an automorphism of H that maps G onto G^* and maps G^* onto G. Therefore G and G^* are isomorphic, and G is the centralizer of G^* in H. Furthermore Lemma 1 is equivalent to the following:

LEMMA 1*. *Let $H = H_1 \times H_2$. Then $G^* \cap H_1$ and $G^* \cap H_2$ are characteristic subgroups of G^* and*

$$G^* = (G^* \cap H_1) \times (G^* \cap H_2).$$

If G is complete, i.e., if G is a centerless group with only inner automorphisms, then $H = G \times G^*$.

3. Decomposable and indecomposable holomorphs. If G is the direct product of two proper characteristic subgroups, then G is said to be characteristically decomposable. If not, then G is said to be characteristically indecomposable.

THEOREM 1. *Let G be a group, and let H be its holomorph. If G is either characteristically decomposable or complete, then H is decomposable. If G is characteristically indecomposable and not complete, then H is indecomposable.*

Proof. We have seen in § 2 that H is decomposable if G is either characteristically decomposable or complete. Suppose that G is characteristically indecomposable and that $H = H_1 \times H_2$. It follows from Lemma 1 that either $G \cap H_1 = G$ or $G \cap H_2 = G$. Thus either $G \subseteqq H_1$ or $G \subseteqq H_2$. Similarly it follows from Lemma 1* that either $G^* \subseteqq H_1$ or $G^* \subseteqq H_2$. Without loss of generality suppose that $G \subseteqq H_1$. Then H_2 is contained in the centralizer of G, that is $H_2 \subseteqq G^*$. If $G^* \subseteqq H_1$ we have $H_2 \subseteqq H_1$ and $H = H_1$. Thus we need only consider the case $G^* \subseteqq H_2$.

Here $G^* = H_2$ and H_1 is contained in the centralizer of G^*. Thus $H_1 \leqq G$, and hence $H_1 = G$. Now $G \cap G^*$ is the center of G, and $G \cap G^* = H_1 \cap H_2$. Hence G is centerless. Since $H = H_1 \times H_2 = G \times G^*$, it follows that G has only inner automorphisms. Therefore G is complete. This completes the proof of the theorem.

4. Decomposition of the holomorph into indecomposable subgroups. To complete our discussion we need the following result:

LEMMA 2. *If a group is complete and characteristically indecomposable, then it is indecomposable.*

Proof. Let G be a complete group and suppose $G = G_1 \times G_2$. Since every automorphism of G is inner, it follows that every automorphism of G maps G_1 and G_2 onto themselves. Hence G_1 and G_2 are characteristic subgroups of G. This establishes the lemma.

THEOREM 2. *Suppose G is the direct product of a finite number of characteristically indecomposable characteristic subgroups: $G = G_1 \times G_2 \times \cdots \times G_n$. Suppose that G_i is complete for $1 \leqq i \leqq r$, and that G_j is not complete for $r+1 \leqq j \leqq n$. Then a decomposition of H into indecomposable subgroups is given by*

$$H = \prod_{i=1}^{r} G_i \times \prod_{i=1}^{r} G_i^* \times \prod_{i=r+1}^{n} H_i \,, \tag{1}$$

where G_i^* and H_i are the conjoint and holomorph respectively of G_i, and where Π denotes a direct product.

Proof. It follows from § 2 that (1) is a decomposition of H. By Lemma 2 the groups G_i and G_i^* are indecomposable for $1 \leqq i \leqq r$, and by Theorem 1 the groups H_i are indecomposable for $r+1 \leqq i \leqq n$.

Since a characteristic subgroup of a characteristic subgroup of G is itself a characteristic subgroup of G it follows that G satisfies the condition of Theorem 2 whenever the characteristic subgroups of G satisfy the descending chain condition. In particular Theorem 2 gives us a decomposition of H into indecomposable subgroups whenever G is a finite group.

If G is the direct product of an infinite number of characteristic subgroups, then H is not the direct product of their holomorphs. Thus Theorem 2 does not hold in this case.

BIBLIOGRAPHY

1. W. H. Mills, *On the non-isomorphism of certain holomorphs*, Trans. Amer. Math. Soc., **74** (1953), 428–443.

2. W. Permans, *Completeness of holomorphs*, Nederl. Akad. Wetensch. Proc. Ser. A **60** (1957), 608–619.

YALE UNIVERSITY AND UNIVERSITY OF CALIFORNIA

ON THE REPRESENTATION THEORY FOR CYLINDRIC ALGEBRAS

DONALD MONK

The main purpose of this paper is to give some new sufficient conditions for the representability of infinite dimensional cylindric algebras. We also discuss certain problems and results in the representation theory reported on by Henkin and Tarski in [5].

In general we adopt the notation of [5]. § 1 contains some additional notation, the statement of a representation theorem of Henkin and Tarski frequently used in this paper, and an embedding theorem which throws some light on that representation result. § 2 is devoted mainly to some simple proofs for known results about the general algebraic theory of representable cylindric algebras. Then in § 3 we turn to representation theory proper. The first result of this section gives a sufficient condition for representability in terms of isomorphic reducts of an algebra (this result was independently obtained by Alfred Tarski). Then follows the definition of a new class of cylindric algebras, diagonal cylindric algebras. The main theorem of this paper is that every diagonal cylindric algebra is representable; this result represents a considerable improvement of some previously known representation theorems. Several interesting corollaries are derived from this result.

1. Introduction. We use the notation of [5] with the following additions. For abbreviational purposes we use standard logical notation: $\rightarrow$ (implies), $\bigvee$ (there exists), and $\bigwedge$ (for all). The identity map on a set A is denoted by δ_A. The function f restricted to the subset A of its domain is denoted by $f \restriction A$. If R is a binary relation and A is a set, then $R^*(A) = \{y \mid \bigvee_{x \in A} \langle xy \rangle \in R\}$. If $\mathfrak{A} = \langle A, +, \cdot, -, c_\kappa, d_{\kappa\lambda} \rangle_{\kappa,\lambda<\alpha}$ is a CA_α, then $\mathfrak{A}_0 = \langle A, +, \cdot, - \rangle$ is the *Boolean part* of $\mathfrak{A}$. *Directed systems* are understood in the sense of [7] p. 65.

We need some notions of general algebra, adapted from [9]. Let $\boldsymbol{K}$ be a class of similar algebras; say all algebras of $\boldsymbol{K}$ are indexed by a nonempty set $N_{\boldsymbol{K}}$, so that if $\mathfrak{A} \in \boldsymbol{K}$ then $\mathfrak{A} = \langle A, O_i \rangle_{i \in N_{\boldsymbol{K}}}$, the O_i being operations on A. We let $\boldsymbol{HK}$ = the class of all homomorphic images of algebras of $\boldsymbol{K}$, $\boldsymbol{PK}$ = the class of all Cartesian products of systems of

Received December 15, 1960. The results of this paper were obtained in part while the author was a National Science Foundation predoctoral fellow and in part while the author was engaged in a research project in the foundations of mathematics directed by Alfred Tarski and supported by the National Science Foundation (Grant No. G-14006) The author wishes to thank Professor Alfred Tarski for the valuable advice he gave during the preparation of this paper. The results of this paper constitute part of the author's doctoral dissertation submitted in May 1961 at the University of California, Berkeley.

algebras of $\boldsymbol{K}$, and $\boldsymbol{SK}=$ the class of all subalgebras of algebras of $\boldsymbol{K}$. If $J \subseteq N_{\boldsymbol{K}}$ and $\mathfrak{A} \in \boldsymbol{K}$, we let $\mathfrak{A}_J = \langle A, O_i\rangle_{i\in J}$; and we let $\boldsymbol{K}_J = \{\mathfrak{A}_J \mid \mathfrak{A} \in \boldsymbol{K}\}$.

To fit cylindric algebras into this scheme of universal algebra, let us make the following agreement. For each ordinal α, let $M_\alpha = \{0, 1, 2, \langle 0, \kappa\rangle, \langle 0, \kappa, \lambda\rangle\}_{\kappa,\lambda<\alpha}$. If $\mathfrak{A} = \langle A, +, \cdot, -, c_\kappa, d_{\kappa\lambda}\rangle_{\kappa,\lambda<\alpha}$ is a CA_α, we let $O_0 = +$, $O_1 = \cdot$, $O_2 = -$, $O_{\langle 0,\kappa\rangle} = c_\kappa$, and $O_{\langle 0,\kappa,\lambda\rangle} = d_{\kappa\lambda}$ for all $\kappa, \lambda < \alpha$; finally we let $\mathfrak{A}^* = \langle A, O_i\rangle_{i\in M_\alpha}$. We let $CA_\alpha^* = \{\mathfrak{A}^* \mid \mathfrak{A} \in CA_\alpha\}$. Thus CA_α^* is a class of similar algebras in the above sense. When no confusion results we shall identify CA_α with CA_α^*.

In several of the proofs below we use a method of construction whose general form is as follows. We are given a class $\boldsymbol{K}$ of similar algebras, a directed system $\mathfrak{D} = \langle D, \geqq\rangle$, and, for each $d \in D$, an element $\mathfrak{A}_d$ of $\boldsymbol{K}$. We let $R = \{\langle f, g\rangle \mid f, g \in \prod_{d\in D} A_d$ and $\bigvee_{d\in D}\bigwedge_{e\in D}(d \leqq e \rightarrow f_e = g_e)\}$. Clearly R is a congruence relation on $\prod_{d\in D}\mathfrak{A}_d$; R is called the *eventually equal congruence of* $\mathfrak{A}$ and $\mathfrak{D}$.[1] In case $\boldsymbol{K} = CA_\alpha^*$ for some α, $\{f \mid \langle f, 0\rangle \in R\}$ is called the *eventually zero ideal of* $\mathfrak{A}$ *and* $\mathfrak{D}$. In case $J \subseteq N_{\boldsymbol{K}}$ and $\mathfrak{B}$ is a subalgebra of $\mathfrak{A}_{dJ}$ for each $d \in D$, we may define *natural isomorphisms* g and f of $\mathfrak{B}$ into $\prod_{d\in D}\mathfrak{A}_{dJ}$ and $[\prod_{d\in D}\mathfrak{A}_d/R]_J$ respectively. For each $b \in B$ and $d \in D$ let $g(b)_d = b$. For each $b \in \boldsymbol{B}$ let $f(b) = [g(b)]$. If $\boldsymbol{K} = CA_\alpha$ and $J = \{0, 1, 2\}$, g and f are called the *natural Boolean isomorphisms* of $\mathfrak{B}$ into $\prod_{d\in D}\mathfrak{A}_d$ and $\prod_{d\in D}\mathfrak{A}_d/R$ respectively.

The essential steps in the proofs of the representation theorems here presented use the following theorem of Henkin and Tarski (see [5] Theorem 2.15).

THEOREM A. *A* CA_α $\mathfrak{A}$ *is representable if and only if for each* $\kappa < \omega$ $\mathfrak{A}$ *can be neatly embedded in some* $CA_{\alpha+\kappa}$.

There now exist purely algebraic proofs of this theorem. Theorem A is to be contrasted with the following theorem:

THEOREM 1. *If* $\delta \geqq \alpha \geqq \omega$, *then every* CA_α *is embeddable in some* CA_δ, *i.e., is a subalgebra of the* α*-reduct of some* CA_δ.[2]

Proof. It suffices to take the case $\delta = \alpha + 1$. For each $\beta < \omega$ we define $\gamma^{(\beta)}$ with domain $\alpha + 1$ by:

$$\gamma_\kappa^{(\beta)} = \begin{cases} \kappa & \text{if } \kappa < \beta\,, \\ \kappa + 1 & \text{if } \beta \leqq \kappa < \omega\,, \\ \kappa & \text{if } \omega \leqq \kappa < \alpha\,, \\ \beta & \text{if } \kappa = \alpha\,, \end{cases}$$

[1] $\Pi_{d\in D}\mathfrak{A}_d/R$ is a reduced product in the sense of Frayne, Scott, and Tarski (Notices Amer. Math. Soc., **5** (1958) 673). In fact, let $J = \{X \mid X \subseteq D$ and $\vee_{d\in D}\wedge_{e\in D}(d \leqq e \rightarrow e \notin X)\}$. Then J is an ideal in the field of all subsets of D, and R is the congruence relation on $\Pi_{d\in D}\mathfrak{A}_d$ determined by J.

[2] This theorem, due to the author, is stated in [5].

for all $\kappa < \alpha + 1$. Thus $\gamma^{(\beta)}$ is one-to-one. Let $\mathfrak{A}$ be a given CA_α, and let $\mathfrak{B}_\beta$ be the $\alpha + 1$, $\gamma^{(\beta)}$-reduct of $\mathfrak{A}$. Let I be the eventually zero ideal of $\mathfrak{B}$ and $\langle \omega, \geqq \rangle$. Let $\mathfrak{C} = \prod_{\beta<\omega} \mathfrak{B}_\beta / I$, and let g, f be the natural Boolean isomorphisms of $\mathfrak{A}$ into $\prod_{\beta>\omega} \mathfrak{B}_\beta$ and $\mathfrak{C}$ respectively. If $0 \leqq \kappa < \beta < \omega$ or $0 \leqq \beta < \omega \leqq \kappa < \alpha$, then $\gamma_\kappa^{(\beta)} = \kappa$, and so, with $a \in A$, $(c_\kappa g(a))_\beta = c_\kappa^{\mathfrak{B}_\beta} g(a)_\beta = c_{\gamma_\kappa^{(\beta)}}^{\mathfrak{A}} a = c_\kappa^{\mathfrak{A}} a = g(c_\kappa a)_\beta$; similarly for diagonal elements. It follows that f is a cylindric isomorphism of $\mathfrak{A}$ into the α-reduct of $\mathfrak{C}$, as required.[3]

Since for each $\alpha \geqq \omega$ there are non-representable CA_α's, Theorems A and 1 indicate the significance of the notion of *neat* embedding.

2. Universal algebra and cylindric algebra.

In [5], Henkin and Tarski state several universal algebraic properties of representable cylindric algebras, indicating that their proofs use in an essential way some metamathematical results. Thus after proving that RCA_α is a universal class, they infer that

(i) a cylindric algebra is representable if and only if every finitely generated subalgebra of it is representable, and

(ii) a cylindric algebra is representable if and only if every finite reduct of it is representable.

Further, after proving that RCA_α is equational they infer that RCA_α is closed under the taking of homomorphic images. For all these algebraic results they raise the question concerning the existence of simple algebraic (as opposed to metamathematical) proofs.

With the essential help of Theorem A, which, as mentioned above, has algebraic proofs, we shall give algebraic proofs of the above results. In addition, we obtain a new proof of the equational character of RCA_α.

THEOREM 2. *A homomorphic image of an RCA_α is an RCA_α.*[4]

Proof. Suppose $\mathfrak{A}$ is an RCA_α and I is a cylindric ideal in $\mathfrak{A}$; we want to show that $\mathfrak{A}/I$ is an RCA_α. Let $\mathfrak{B}$ be a $CA_{\alpha+\kappa}$ such that $\mathfrak{A}$ is neatly embedded in $\mathfrak{B}$ (by Theorem A), where $\kappa < \omega$. Let J be the ideal in $\mathfrak{B}$ generated by I. Clearly $J = \{b \mid b \in B \text{ and } \bigvee_{a \in I}(b \leqq a)\}$, and so $J \cap A = I$. It follows that the natural Boolean homomorphism of $\mathfrak{A}/I$ into $\mathfrak{B}/J$ is a cylindric isomorphism of $\mathfrak{A}/I$ onto an algebra neatly embedded in $\mathfrak{B}/J$, and by Theorem A our theorem follows.

It is easy to see that RCA_α is closed under direct products and subalgebras. Hence by Birkhoff's theorem (Theorem 2.1 of [9]), RCA_α is equational. Thus in particular, RCA_α is a universal class, and the above characterizations (i) and (ii) of RCA_α follow. Recently the author

[3] Theorem 1 can also be easily proved metamathematically. In fact, it was such a proof that first occurred to the author.

[4] [5], Theorem 2.20.

obtained simple algebraic proofs of these two characterizations. Alfred Tarski, upon being informed of these proofs, recalled that in 1955 Saunders MacLane outlined to him a proof of a universal algebraic theorem from which (i) follows; the author's proof of (i) was a specialization of MacLane's proof. Since MacLane's proof has never appeared in print, we shall take this opportunity to present it here. Subsequent to the above work, the author obtained a corresponding algebraic proof of a generalization of (ii).

Of the two corollaries below, the first is a strict specialization of the universal algebraic case, while for the second corollary we apply an additional argument.

THEOREM 3. *Let $\boldsymbol{K}$ be a class of similar algebras such that $\boldsymbol{HK} = \boldsymbol{K}$, $\boldsymbol{PK} = \boldsymbol{K}$, and $\boldsymbol{SK} = \boldsymbol{K}$. Then for every algebra $\mathfrak{A}$, $\mathfrak{A} \in \boldsymbol{K}$ if (and only if) every finitely generated subalgebra of $\mathfrak{A}$ is in $\boldsymbol{K}$.*

Proof. The necessity of the condition is obvious. Now suppose that every finitely generated subalgebra of $\mathfrak{A}$ is in $\boldsymbol{K}$. Let $I = \{F \mid F$ is a finite subset of $A\}$, and for each $F \in I$ let $\mathfrak{B}_F$ be the subalgebra of $\mathfrak{A}$ generated by F. Let R be the eventually equal congruence of $\mathfrak{B}$ and $\langle I, \supseteq \rangle$, and let $\mathfrak{C} = \prod_{F \in I} \mathfrak{B}_F / R$. By hypothesis, $\mathfrak{C} \in \boldsymbol{K}$. Define g with domain A and range included in $\prod_{F \in I} B_F$ by:

$$g(a)_F = \begin{cases} \text{any element of } B_F \text{ if } a \notin B_F\,, \\ a \text{ if } a \in B_F\,, \end{cases}$$

for all $a \in A$ and $F \in I$. It is easy to see that the function f, defined by $f(a) = [g(a)]$ for all $a \in A$, is an isomorphism of $\mathfrak{A}$ into $\mathfrak{C}$. Hence $\mathfrak{A} \in \boldsymbol{K}$.

From Theorems 2 and 3 we obtain:

COROLLARY. *$\mathfrak{A} \in RCA_\alpha$ if (and only if) every finitely generated subalgebra of $\mathfrak{A}$ is representable.*[5]

THEOREM 4. *Let $\boldsymbol{K}$ be a class of similar algebras such that $\boldsymbol{HK} = \boldsymbol{K}$, $\boldsymbol{PK} = \boldsymbol{K}$, and $\boldsymbol{SK} = \boldsymbol{K}$. Then $\mathfrak{A} \in \boldsymbol{K}$ if (and only if) for every finite subset F of $N_{\boldsymbol{K}}$ we have $\mathfrak{A}_F \in \boldsymbol{SK}_F$.*

Proof. The necessity is obvious. Now suppose that the above condition holds. For each finite subset F of $N_{\boldsymbol{K}}$ choose $\mathfrak{B}^{(F)} \in \boldsymbol{K}$ such that $\mathfrak{A}_F \subseteqq \mathfrak{B}_F^{(F)}$. Choose $i_0 \in N_{\boldsymbol{K}}$. Let $I = \{F \mid F$ is a finite subset of $N_{\boldsymbol{K}}$ and $i_0 \in F\}$. Let R be the eventually equal congruence of $\mathfrak{B}$ and $\langle I, \supseteqq \rangle$, and let $\mathfrak{C} = \prod_{F \in I} \mathfrak{B}^{(F)} / R$. Let g and f be the natural isomorphisms of

[5] [5], Theorem 2.13 (i).

$\mathfrak{A}_{\{i_0\}}$ into $\prod_{F\in I}\mathfrak{B}^{(F)}_{\{i_0\}}$ and $\mathfrak{C}_{\{i_0\}}$ respectively. We claim that f is an isomorphism of $\mathfrak{A}$ into $\mathfrak{C}$. For, if $i \in N_K$, say with $O_i^{\mathfrak{A}}$ binary, and if $a, b \in A$, we have for $\{i\} \subseteq F \in I$:

$$\begin{aligned}[g(O_i^{\mathfrak{A}}(a, b))]_F &= O_i^{\mathfrak{A}}(a, b)\\ &= O_i^{\mathfrak{A}F}(a, b)\\ &= O_i^{\mathfrak{B}_F^{(F)}}(a, b)\\ &= O_i^{\mathfrak{B}^{(F)}}([g(a)]_F, [g(b)]_F)\\ &= [O_i(g(a), g(b))]_F\ .\end{aligned}$$

Thus $f(O_i^{\mathfrak{A}}(a, \mathrm{b})) = O_i(f(a), f(b))$. We deduce that $\mathfrak{A} \in \boldsymbol{K}$ by the hypothesis of the theorem.

Again, we have a corollary for cylindric algebras. As mentined previously, this corollary is not quite as immediate as the corollary to Theorem 3; we need the following lemma in order to derive the corollary easily.

LEMMA 1. *Let $\boldsymbol{K}$ be a class of similar algebras such that $\boldsymbol{PK} = \boldsymbol{K}$ and $\boldsymbol{SK} = \boldsymbol{K}$. Suppose that $\mathfrak{A}$ is an algebra such that for all $x, y \in A$ with $x \neq y$ there is a homomorphism f of $\mathfrak{A}$ onto an algebra $\mathfrak{B} \in \boldsymbol{K}$ such that $f(x) \neq f(y)$. Then $\mathfrak{A} \in \boldsymbol{K}$.*

In case additionally $\boldsymbol{K} = CA_\alpha$ it is enough to assume that for all $x \in A$ with $x \neq 0$ there is a homomorphism f of $\mathfrak{A}$ into an algebra $\mathfrak{B} \in \boldsymbol{K}$ such that $f(x) \neq 0$.

The proof of this lemma is simple; it is essentially due to Birkhoff ([1]).

The proof of necessity in the following corollary gives a simple proof of Theorem 2.12 of [5].

COROLLARY. *$\mathfrak{A} \in RCA_\alpha$ if and only if every finite reduct of $\mathfrak{A}$ is representable.*[6]

Proof. Necessity. Suppose $\mathfrak{A} \in RCA_\alpha$, i.e., $\mathfrak{A}$ is isomorphic to a sub-direct product of CSA_α's. Now a reduct of a product of CA_α's is equal to the product of the corresponding reducts. Hence we may assume that $\mathfrak{A}$ is a CSA_α, say with base U. Suppose $\kappa < \omega$ and $\theta \in \alpha^\kappa$ is one-to-one; let $\mathfrak{B}$ be the κ, θ-reduct of $\mathfrak{A}$. Suppose $b \in B$ and $b \neq \mathbf{0}$; choose $f \in b$. For each $g \in U^\kappa$ we define $g^* \in U^\alpha$ by:

$$g^*_\lambda = \begin{cases} g_{\theta^{-1}\lambda} \text{ if } \lambda \in \text{ range } \theta\ , \\ f_\lambda \text{ otherwise.}\end{cases}$$

[6] [5], Theorems 2.12, 2.13 (ii).

Define $F(x) = \{g \in U^\kappa \mid g^* \in x\}$ for each $x \in B$. It is easy to verify that F is a homomorphism of $\mathfrak{B}$ onto a CSA_κ such that $F(b) \neq 0$. Since b is arbitrary, we deduce from Lemma 1 that $\mathfrak{B} \in RCA_\kappa$.

Sufficiency. We now assume that every finite reduct of $\mathfrak{A}$ is representable. Let a finite subset F of M_α be called *regular* if there is a finite subset G of α such that $F = \{0, 1, 2, \langle 0, \kappa\rangle, \langle 0, \kappa, \lambda\rangle\}_{\kappa,\lambda \in G}$. Now it is known, and easy to see, that an RCA_α can be neatly embedded in an RCA_β for each $\beta \geqq \alpha$; if we apply this argument here we see that, by our assumption, $\mathfrak{A}_F^* \subseteqq \mathfrak{B}_F^{(F)*}$ for some $\mathfrak{B}^{(F)} \in RCA_\alpha$, for each regular finite subset F of M_α. If F is any finite subset of M_α, then there is a regular finite subset G such that $F \subseteqq G$, and so $\mathfrak{A}_F^* = (\mathfrak{A}_G^*)_F \subseteqq (\mathfrak{B}_G^{(F)*})_F \in [(RCA_\alpha^*)_G]_F = (RCA_\alpha^*)_F$. Hence by Theorem 4, $\mathfrak{A}^* \in RCA_\alpha^*$, i.e., $\mathfrak{A} \in RCA_\alpha$.

We conclude this section with the following theorem.

THEOREM 2′. *Let α and κ be ordinals. Let $\boldsymbol{K}$ be the class of all CA_α's which can be neatly embedded in a $CA_{\alpha+\kappa}$. Then $\boldsymbol{K}$ is an equational class.*

Proof. Clearly $\boldsymbol{K}$ is closed under direct products and subalgebras. The proof of Theorem 2 may be applied to show that $\boldsymbol{K}$ is closed under homomorphisms. Our theorem is now a consequence of Birkhoff's theorem.

From this theorem we can derive two corollaries similar to the above stated corollaries. This can be done metamathematically, in the obvious way, or mathematically as follows. For the first corollary we can again use Theorem 3, while for the second we can use a direct argument similar to the proof of Theorem 4. (We do not know of any way of using Theorem 4 or something like it to derive the second corollary.)

3. Some representation theorems.

Now we shall prove several new sufficient conditions for the representability of cylindric algebras. The following simple lemma will be found useful in the proofs of the main results.

LEMMA 2. *Let α, β, and γ be ordinals, and suppose that $\tau \in \beta^{\alpha+\gamma}$ is one-to-one. Suppose $\mathfrak{A}$ is a CA_α, $\mathfrak{B}$ is a CA_β, $T \in B^A$, and the following conditions hold:*

(i) *T is a Boolean homomorphism of $\mathfrak{A}$ into $\mathfrak{B}$,*

(ii) *$c_{\tau\kappa}^{\mathfrak{B}} \circ T = T \circ c_\kappa^{\mathfrak{A}}$ for all $\kappa < \alpha$,*

(iii) *$T(d_{\kappa\lambda}^{\mathfrak{A}}) = d_{\tau\kappa,\tau\lambda}^{\mathfrak{B}}$ for all $\kappa, \lambda < \alpha$.*

Then T is a cylindric homomorphism of $\mathfrak{A}$ into the α-reduct of some $CA_{\alpha+\gamma}$. If in addition the following condition holds:

(iv) *$c_{\tau\kappa}^{\mathfrak{B}} \circ T = T$ for $\alpha \leqq \kappa < \alpha + \gamma$,* *then T is a cylindric homomorphism of $\mathfrak{A}$ into an algebra neatly embedded in some $CA_{\alpha+}$.*

Proof. Let $\mathfrak{C}$ be the $\alpha+\gamma$, τ-reduct of $\mathfrak{B}$, and let $\mathfrak{D}$ be the α-reduct of $\mathfrak{C}$. Then T is a cylindric homomorphism of $\mathfrak{A}$ into $\mathfrak{D}$, for $T\circ c_k^{\mathfrak{A}} = c_{\tau\kappa}^{\mathfrak{B}}\circ T$ (by (ii)) $= c_\kappa^{\mathfrak{C}}\circ T = c_k^{\mathfrak{D}}\circ T$ for each $\kappa < \alpha$, and $T(d_{\kappa\lambda}^{\mathfrak{A}}) = d_{\tau\kappa,\tau\lambda}^{\mathfrak{B}} = d_{\kappa\lambda}^{\mathfrak{C}} = d_{\kappa\lambda}^{\mathfrak{D}}$ for all $\kappa, \lambda < \alpha$. If in addition (iv) holds, then for $\alpha \leqq \kappa < \alpha+\gamma$ we have $c_\kappa^{\mathfrak{C}}\circ T = c_{\tau\kappa}^{\mathfrak{B}}\circ T = T$.

As a consequence of Theorem A and Lemma 2 we have the following representation theorem, which was independently obtained by Alfred Tarski.

THEOREM 5. *Assume that $\mathfrak{A}$ is a CA_α, σ is a one-to-one element of α^α such that $\alpha \sim$ range (σ) is infinite, and $\mathfrak{B}$ is the α, σ-reduct of $\mathfrak{A}$. Suppose T is an isomorphism of $\mathfrak{A}$ into $\mathfrak{B}$ such that $c_\kappa^{\mathfrak{A}}T(x) = T(x)$ whenever $x \in A$ and $\kappa \in \alpha \sim$ range (σ). Then $\mathfrak{A}$ is representable.*

Proof. Let τ be a one-to-one element of $\alpha^{\alpha+\omega}$ such that $\tau \restriction \alpha = \sigma$. Then for all $\kappa < \alpha$ we have $c_{\tau\kappa}^{\mathfrak{A}}\circ T = c_{\sigma\kappa}^{\mathfrak{A}}\circ T = c_\kappa^{\mathfrak{B}}\circ T = T\circ c_\kappa^{\mathfrak{A}}$. Moreover, for all $\kappa, \lambda < \alpha$ we have $T(d_{\kappa\lambda}^{\mathfrak{A}}) = d_{\kappa\lambda}^{\mathfrak{B}} = d_{\tau\kappa,\tau\lambda}^{\mathfrak{A}}$. Finally, if $\alpha \leqq \kappa < \alpha+\omega$, then $c_{\tau\kappa}^{\mathfrak{A}}\circ T = T$. Hence by Lemma 2 $\mathfrak{A}$ can be neatly embedded in a $CA_{\alpha+\omega}$, and our theorem follows from Theorem A.

We should mention that recently Tarski obtained a stronger version of Theorem 5, in which the condition "$\alpha \sim$ *range* (σ) *is infinite*" is replaced by the condition "$\alpha \sim$ *range* $(\sigma) \neq 0$".

Theorem 5 leads to an interesting insight into the relationship between cylindric and polyadic algebras, of a different kind from the insight obtained from the relationships established in [2]. A polyadic algebra with equality is, roughly speaking, a cylindric algebra with two additional structures: infinite cylindrification, and substitution (see [3]). If we eliminate only the infinite cylindrification, we arrive at a notion of a *substitution* on a cylindric algebra. A substitution on a CA_α $\mathfrak{A}$ is a function $S \in (A^A)^{\alpha^\alpha}$ which satisfies certain natural conditions (due to Halmos). As a corollary of Theorem 5 we easily see that if $\mathfrak{A}$ is a CA_α with a substitution and if $\alpha \geqq \omega$, then A is representable. Now from [6] it is known that every infinite dimensional polyadic algebra is representable, while there are infinite dimensional polyadic equality algebras which are not representable (with equality corresponding to the functional equality). Here by *representable* we mean as in cylindric algebras—isomorphic to a subdirect product of $\bigcirc$-valued functional polyadic algebras. Our corollary shows that by eliminating infinite cylindrification we recapture representation.

It is natural to ask if the corollary can be strengthened by replacing "substitution" by "finite substitution"—a concept defined like that of substitution, but in which S applies only to those $\tau \in \alpha^\alpha$ for which there is a finite subset F of α such that $\tau \upharpoonright \alpha \sim F = \delta_{\alpha \sim F}$. The answer is no: for each $\alpha \geqq \omega$ there exists a CA_α with a finite substitution which is not representable. The construction of such algebras depends on the results of [8], which in turn depend upon unpublished work of Henkin and Tarski.

We now define a class of cylindric algebras which includes both the class of simple infinite dimensional cylindric algebras and the class of dimensionally complemented cylindric algebras. A CA_α $\mathfrak{A}$ is called a *diagonal cylindric algebra* (GCA_α) provided that for every non-zero $a \in A$ and every finite subset F of α there are distinct $\kappa, \lambda \in \alpha \sim F$ such that $a \cdot d_{\kappa\lambda} \neq 0$. The importance of this concept derives from the following theorem:

THEOREM 6. *Every diagonal cylindric algebra is representable.*[7]

Proof. Let $\mathfrak{A}$ be a GCA_α. We want to apply Lemma 1, with $\boldsymbol{K}$ replaced by the class of all CA_α's which can be neatly embedded in $CA_{\alpha+1}$'s. Hence suppose that $a \in A$ and $a \neq 0$. Since $\mathfrak{A} \in GCA_\alpha$ we can define functions μ, ν with domain ω inductively by letting μ_κ and ν_κ be distinct members of $\alpha \sim \{\mu_\lambda, \nu_\lambda \mid \lambda < \kappa\}$ such that $a \cdot d_{\mu_\kappa \nu_\kappa} \neq 0$.

Now we prepare to apply Lemma 2. It is easy to see that there is a unique $\tau \in \alpha^{\alpha+1}$ such that the following conditions hold:

(1) τ is one-to-one,

(2) τ is the identity on $\alpha \sim \{\mu_\kappa, \nu_\kappa \mid \kappa < \omega\}$,

(3) $\tau\mu_\kappa = \nu_\kappa$ for each $\kappa < \omega$,

(4) $\tau\nu_\kappa = \mu_{\kappa+1}$ for each $\kappa < \omega$,

(5) $\tau\alpha = \mu_0$.

For each $\kappa < \omega$, let $\mathfrak{B}_\kappa = \mathfrak{A}$. Let I be the eventually zero ideal of $\mathfrak{B}$ and $\langle \omega, \geqq \rangle$, and let $\mathfrak{C} = \mathfrak{A}^\omega / I$. For each $x \in A$ and $\kappa < \omega$, define

$$f(x)_\kappa = S^{\mu_0}_{\nu_0} S^{\nu_0}_{\mu_1} \cdots S^{\nu_{\kappa-1}}_{\mu_\kappa} S^{\mu_\kappa}_{\nu_\kappa} x \ ,$$

where $S^\theta_\rho x = c_\theta(d_{\theta\rho} \cdot x)$ for all $\theta, \rho < \alpha$ and $x \in A$. Let $T(x) = [f(x)]$ for all $x \in A$. The following statements may now be verified:

(6) T is a Boolean homomorphism of $\mathfrak{A}$ into C,

(7) $c^{\mathfrak{C}}_{\tau\lambda} \circ T = T \circ c^{\mathfrak{A}}_\lambda$ for all $\lambda < \alpha$,

[7] After reading a preliminary draft of this paper, Henkin obtained a generalization of this theorem, which may be stated as follows. If for every nonzero $x \in A$ and for every finite $\Gamma \subseteqq \alpha$ there is a $\xi \in \alpha \sim \Gamma$ and an endomorphism T of $\mathfrak{A}_0$ such that $c_\xi \circ T = T$, $c_\kappa \circ T = T \circ c_\kappa$ for each $\kappa \in \Gamma$, and $T(x) \neq 0$, then $\mathfrak{A}$ is representable.

(8) $T(d^{\mathfrak{A}}_{\lambda\mu}) = d^{\mathfrak{C}}_{\tau\lambda,\tau\mu}$ for all $\lambda, \mu < \alpha$,

(9) $c^{\mathfrak{C}}_{\tau\alpha} \circ T = T$.

In verifying (7), one can make use of the following easily verified arithmetic law:

(10) $S^{\rho}_{\theta} c_{\rho} S^{\theta}_{\pi} x = c_{\theta} S^{\rho}_{\theta} S^{\theta}_{\pi} x$ for all $x \in A$ and all distinct $\rho, \theta, \pi < \alpha$.

We can now apply Lemma 2, and infer that T is a cylindric homomorphism of $\mathfrak{A}$ into an algebra neatly embedded in a $CA_{\alpha+1}$. Suppose $T(a) = 0$. Choose $k < \omega$ such that $f(a)_{\kappa} = 0$. Applying successively $S^{\nu_0}_{\mu_0}, S^{\mu_1}_{\nu_0}, \cdots, S^{\mu_\kappa}_{\nu_{\kappa-1}}$ we infer that $S^{\mu_\kappa}_{\nu_\kappa} a = 0$, and so $a \cdot d_{\mu_\kappa \nu_\kappa} = 0$, which is a contradiction. Since a is arbitrary, from Lemma 1 we conclude that $\mathfrak{A}$ can be neatly embedded in some $CA_{\alpha+1}$ $\mathfrak{D}$. Let g be an isomorphism of $\mathfrak{A}$ onto an algebra neatly embedded in $\mathfrak{D}$.

Let N be maximal among ideals P such that $g^*(A) \cap P = \{0\}$ (by Zorn's lemma). Let $\mathfrak{E} = \mathfrak{D}/N$, and let pr be the natural homomorphism of $\mathfrak{D}$ onto $\mathfrak{E}$. Clearly $pr \circ g$ is an isomorphism of $\mathfrak{A}$ onto an algebra neatly embedded in $\mathfrak{E}$. Suppose $x \in D$, F is a finite subset of α, and $[x] \leqq [-d_{\kappa\lambda}]$ for all distinct $k, \lambda \in \alpha \sim F$. Suppose that $x \notin N$. Then $N \cup \{x\}$ generates an ideal P such that $P \cap g^*(A) \neq \{0\}$. Choose $y \in A$ such that $g(y) \neq 0$ and $g(y) \in P$. Then there are $\kappa_0. \cdots, \kappa_{\nu-1} \in \alpha + 1$ and $n \in N$ such that $g(y) \leqq n + c_{\kappa_0} \cdots c_{\kappa_{\nu-1}} x$. Let $F' = F \cup \{\kappa_0, \cdots, k_{\nu-1}\}$. Then $[g(y)] \leqq [-d_{\kappa\lambda}]$ for all distinct $\kappa, \lambda \in \alpha \sim F'$; but this contradicts the fact that $\mathfrak{A}$ is a diagonal cylindric algebra.

It follows that $\mathfrak{E}$ is a $GCA_{\alpha+1}$. Hence all the preceding proof can be applied inductively to give, in virtue of Theorem A, the desired result.

We now proceed to derive some consequences of Theorem 6.

THEOREM 7. *Every simple infinite dimensional algebra is a diagonal cylindric algebra, and so is representable.*

Proof. Suppose $\mathfrak{A}$ is a simple CA_α, $\alpha \geqq \omega$, $a \in A$, $a \neq 0$, and F is a finite subset of α. There are $\lambda \in \omega \sim 1$ and $\mu \in \alpha^\lambda$ such that $c_{\mu_0} \cdots c_{\mu_{\lambda-1}} a = 1$. Choose κ, ν distinct in $\alpha \sim (F \cup \{\mu_0, \cdots, \mu_{\lambda-1}\})$. If $a \cdot d_{\kappa\nu} = 0$, then, applying $c_{\mu_0} \cdots c_{\mu_{\lambda-1}}$, we see that $d_{\kappa\nu} = 0$; hence $0 = 1$, contradicting the simplicity of $\mathfrak{A}$.

From Theorem 7 we can infer the following negative theorem which limits the possible extensions of Theorem 1.

THEOREM 8. *If* $1 < \alpha < \omega$, *then it is not the case that every* CA_α *can be embedded* (*in the sense of Theorem* 1) *in a* CA_ω.

Proof. Assume the contrary. Henkin and Tarski have constructed a non-representable CA_α $\mathfrak{A}$, in unpublished work. Clearly we may assume that $\mathfrak{A}$ is simple. Let $\mathfrak{B}$ be a CA_ω such that $\mathfrak{A}$ is a subalgebra of the α-reduct of $\mathfrak{B}$. Let I be a maximal ideal in $\mathfrak{B}$, and let $\mathfrak{C} = \mathfrak{B}/I$. By Theorem 7, $\mathfrak{C}$ is representable. Since $\mathfrak{A}$ is simple, $A \cap I = \{0\}$, and so the natural homomorphism of $\mathfrak{A}$ into $\mathfrak{C}$ is an isomorphism. It follows that $\mathfrak{A}$ is representable; but this is a contradiction.

A CA_α $\mathfrak{A}$ is *weakly dimensionally complemented*, $\mathfrak{A} \in WDCA_\alpha$ if $\alpha \sim \Delta x$ is infinite for every $x \in A$.

THEOREM 9. *Every weakly dimensionally complemented cylindric algebra is a diagonal cylindric algebra, and so is representable.*[8]

Proof. Suppose $\mathfrak{A}$ is a $WDCA_\alpha$, $a \in A$, $a \neq 0$, and F is a finite subset of α. Choose κ, λ distinct in $\alpha \sim F$ such that $c_\kappa c_\lambda a = a$. If $a \cdot d_{\kappa\lambda} = 0$, then $a = 0$, contradiction.

THEOREM 10. *Let α be an infinite ordinal, and let $\mathfrak{A}$ be a CA_α. Suppose there is a finite subset F of $\alpha^2 \sim \delta_\alpha$ such that $\prod_{\langle\kappa,\lambda\rangle \in F} - d_{\kappa\lambda} = 0$. Then $\mathfrak{A}$ is a diagonal cylindric algebra, and so is representable.*[9]

Proof. Suppose $a \in A$, $a \neq 0$, and G is a finite subset of α. Choose $\kappa \in \omega \sim 1$ and $\mu \in \alpha^\kappa$ such that μ maps κ one-to-one onto the field of F, i.e., onto $\{\lambda \mid \bigvee_{\nu<\alpha}(\langle\lambda, \nu\rangle \in F \text{ or } \langle\nu, \lambda\rangle \in F)\}$. Also choose $\nu \in \alpha^\kappa$ such that ν is one-to-one and range $\nu \subseteq \alpha \sim G \sim$ (field of F). Let

$$H = \{\langle \nu\mu^{-1}\kappa, \nu\mu^{-1}\lambda\rangle \mid \langle\kappa, \lambda\rangle \in F\} .$$

Applying $S^{\mu_0}_{\nu_0} \cdots S^{\mu_{\kappa-1}}_{\nu_{\kappa-1}}$ to $\prod_{\langle\kappa,\lambda\rangle \in F} - d_{\kappa\lambda}$, we see that $\prod_{\langle\kappa,\lambda\rangle \in H} - d_{\kappa\lambda} = 0$. Moreover, H is a finite subset of $\alpha^2 \sim \delta_\alpha$ such that (field of $H) \cap G = 0$. Since $a \neq 0$, choose $\langle\kappa, \lambda\rangle \in H$ such that $a \cdot d_{\kappa\lambda} \neq 0$. Thus $\mathfrak{A}$ is a diagonal cylindric algebra.

In conclusion, we would like to make a few remarks about the general theory of diagonal cylindric algebras. In the first place, GCA_α is properly included in RCA_α; the cylindric set algebra formed from all subsets of ω^ω forms an example of an element of $RCA_\alpha \sim GCA_\alpha$; in this algebra the element $\{\delta_\omega\}$ is included in the complement of every non-unity diagonal element. Clearly GCA_α is closed under direct products and subalgebras. But from Theorem 2.19 of [5] it follows that GCA_α is not equational, and so is not closed under homomorphisms. For,

[8] This is a solution of a problem of Henkin and Tarski, who showed that $\mathfrak{A}$ is representable if $\alpha \sim (\Delta x \cup \Delta y)$ is infinite for all $x, y \in A$.

[9] Actually a somewhat stronger theorem holds. In fact, instead of assuming that F is finite, it suffices to assume that $\alpha \sim$ Field (F) is infinite. Then, in general, the product mentioned in Theorem 10 may be an infinite product.

$LCA_\alpha \subseteqq GCA_\alpha \subset RCA_\alpha$, and by the quoted theorem RCA_α is the smallest equational class including LCA_α.

BIBLIOGRAPHY

1. Garrett Birkhoff, *Subdirect unions in universal algebra,* Bull. Amer. Math. Soc., **50** (1944), 764-768.
2. B. A. Galler, *Cylindric and polyadic algebras,* Proc. Amer. Math. Soc., **8** (1952), 176-183.
3. Paul R. Halmos, *Algebraic logic* IV, Trans. Amer. Math. Soc., **86** (1957), 1-27.
4. Leon Henkin and Alfred Tarski, *Cylindrical algebras,* Summaries of talks presented at the Summer Institute of Symbolic Logic v. 3 Cornell 1957, 332-340.
5. ——— and ———, *Cylindric algebras,* Proc. of symposia in pure math. v. 2, Lattice Theory, 1960, Amer. Math. Soc.
6. H. J. Keisler, *The representation of polyadic algebras of infinite degree,* Notices Amer. Math. Soc., **7** (1960), 364 (Abstract 569-48).
7. John L. Kelley, *General Topology,* 1955.
8. Donald Monk, *Non-representable polyadic algebras of finite degree,* Notices Amer. Math. Soc., **7** (1960), 735 (Abstract 574-5).
9. Alfred Tarski, *Contributions to the theory of models* I, II III, Proc. Neder. Ak. van Weten. Ser. A 57: 572-591; 58: 56-64.

UNIVERSITY OF CALIFORNIA, BERKELEY

A NOTE ON GENERALIZATIONS OF SHANNON-MCMILLAN THEOREM

SHU-TEH C. MOY

1. Introduction. This paper is a sequel to an earlier paper [6]. All notations in [6] remain in force. As in [6] we shall consider tw probability measures μ, ν an the infinite product σ-algebra of subsets of the infinite product space $\Omega = \pi X$. ν is assumed to be *stationary* and μ to be *Markovian with stationary transition probabilities*. Extensions to K-Markovian μ are immediate. $\nu_{m,n}$, the contraction of ν to $\mathscr{F}_{m,n}$, is assumed to be *absolutely continuous* with respect to $\mu_{m,n}$, the contraction of μ to $\mathscr{F}_{m,n}$, and $f_{m,n}$ is the Radon-Nikodym derivative. In [6] the following theorem is proved. If $\int \log f_{0,0}\, d\nu < \infty$ and if there is a number M such that

$$(1) \qquad \int (\log f_{0,n} - \log f_{0,n-1}) d\nu \leqq M \text{ for } n = 1, 2, \cdots$$

then $\{n^{-1} \log f_{0,n}\}$ converges in $L_1(\nu)$. (1) is also a necessary condition for the $L_1(\nu)$ convergence of $\{n^{-1} \log f_{0,n}\}$. We consider this theorem as a generalization of the Shannon-McMillan theorem of information theory. In the setting of [6] the Shannon-McMillan theorem may be stated as follows. Let X be a finite set of K points. Let ν be any stationary probability measure of $\mathscr{F}$, and μ the equally distributed independent measure on $\mathscr{F}$. Then $\{n^{-1} \log f_{0,n}\}$ converges in $L_1(\nu)$. In fact, the $P(x_0, x_1, \cdots, x_n)$ of Shannon-McMillan is equal to $K^{(n+1)} f_{0,n}$. The convergence with probability one of $\{n^{-1} \log P(x_0, \cdots, x_n)\}$ for a finite set X was proved by L. Breiman [1] [2]. K.L. Chung then extended Breiman's result to a countable set X. [3]. In this paper we shall prove that the convergence with ν-probability one of $\{n^{-1} \log f_{0,n}\}$ follows from the following condition.

$$(2) \qquad \int \frac{f_{0,n}}{f_{0,n-1}} d\nu \leqq L,\ n = 1, 2, \cdots .$$

(2) is a stronger condition than (1) since by Jensen's inequality

$$\log \int \frac{f_{0,n}}{f_{0,n-1}} d\nu \geqq \int \log \frac{f_{0,n}}{f_{0,n-1}} d\nu .$$

An application to the case of countable X is also discussed.

Received November 28, 1960.

2. The convergence theorem. As was proved in [6], condition (1) implies the $L_1(\nu)$ convergence of $\{\log f_{-k,0} - \log f_{-k,-1}\}$ ([6] Theorem 1, 4). The convergence with ν-probability one is automatically true ([6] Theorem 3). Applying a theorem (with obvious modification for T not necessarily ergodic) of Breiman ([1], Theorem 1) the convergence with ν-probability one of $\{n^{-1}\log f_{0,n}\}$ follows from the condition

$$\int \sup_{k\geqq 1} |\log f_{-k,0} - \log f_{-k,-1}|\, d\nu < \infty . \tag{3}$$

We shall now investigate conditions under which (3) is valid.

Lemma 1. *The following inequality is always true.*

$$\int \sup_{k\geqq 1} \log \frac{f_{-k,-1}}{f_{-k,0}} d\nu < \infty . \tag{4}$$

Proof. Let $\nu'_{-k,0}$ be as in Lemma 1 [6]. Then

$$\nu_{-k,0} \ll \nu'_{-k,0} \ll \mu_{-k,0}$$

and

$$\frac{d\nu_{-k,0}}{d\nu'_{-k,0}} = \frac{f_{-k,0}}{f_{-k,-1}},\ \frac{d\nu'_{-k,0}}{d\mu_{-k,0}} = f_{-k,-1} .$$

Since μ is Markovian, $\nu'_{-k,0}$ are consistent for $k = 1, 2, \cdots$. We shall prove (4) under the assumption that there is a probability measure ν' on $\mathscr{F}_{-\infty,0}$ which is an extension of $\nu'_{-k,0}$ for $k = 1, 2, \cdots$. We shall also prove Lemma 2 under this assumption. If no such ν' exists, the usual procedure of representing Ω into the space of real sequences may be used and the same conclusion follows (cf. the proof of Theorem 4[6]).

Let m be a nonnegative integer and

$$E(m) = [\sup_{k\geqq 1} \log \frac{f_{-k,-1}}{f_{-k,0}} > m] ,$$

$$E_k(m) = [\sup_{1\leqq j<k} \log \frac{f_{-j,-1}}{f_{-j\,0}} \leqq m,\ \log \frac{f_{-k,-1}}{f_{-k,0}} > m] .$$

On $E_k(m)$ we have

$$f_{-k,0} \leqq 2^{-m} f_{-k,-1} .$$

Hence

$$\int_{E_k(m)} f_{-k,0}\, d\mu \leqq 2^{-m} \int_{E_k(m)} f_{-k,-1}\, d\mu$$

so that

$$\nu[E_k(m)] \leqq 2^{-m}\nu'[E_k(m)] .$$

Therefore

$$\nu[E(m)] \leqq 2^{-m}\nu'[E(m)] \leqq 2^{-m}$$

and

$$\int \sup_{k>1} \log \frac{f_{-k,-1}}{f_{-k,0}} d\nu \leqq \sum_{m\geqq 0} \nu[E(m)] \leqq \sum_{m\geqq 0} 2^{-m} < \infty .$$

Note that (4) is proved without assuming the integrability of either $\log f_{-k,0}$ or $\log f_{-k,-1}$ or $\log \dfrac{f_{-k,0}}{f_{-k,-1}}$.

LEMMA 2. *If there is a number L such that*

$$\int \frac{f_{-k,0}}{f_{-k,-1}} d\nu \leqq L \text{ for } k = 1, 2, \cdots \tag{5}$$

then

$$\int \sup_{k\geqq 1} \log \frac{f_{-k,0}}{f_{-k,-1}} d\nu < \infty . \tag{6}$$

Proof. It is clear that

$$\int \frac{f_{-k,0}}{f_{-k,-1}} d\nu = \int \left(\frac{f_{-k,0}}{f_{-k,-1}}\right)^2 d\nu'$$

where ν' is defined in the proof of Lemma 1.

Since $\{f_{-k,0}/f_{-k,-1}, k = 1, 2, \cdots\}$ is a ν'-martingale, $\{(f_{-k,0}/f_{-k,-1})^2, k = 1, 2, \cdots\}$ is a ν'-semi-martingale. Hence (5) implies that

$$\nu_{-\infty,0} \ll \nu', \int \left(\frac{d\nu_{-\infty,0}}{d\nu'}\right)^2 d\nu' < \infty, \left(\frac{f_{-k,0}}{f_{-k,-1}}\right)^2$$

are uniformly ν'-integrable and $\{(f_{-1,0}/f_{-1,-1})^2, (f_{-2,0}/f_{-2,-1})^2 \cdots, (d\nu_{-\infty,0}/d\nu')^2\}$ is a ν'-semi-martingale (Theorem 4.1s, pp. 324[5]).

Hence for any set F defined by $x_0, x_{-1}, \cdots, x_{-k}$

$$\int_F \left(\frac{f_{-k,0}}{f_{-k,-1}}\right)^2 d\nu' \leqq \int_F \left(\frac{f_{-(k+1),0}}{f_{-(k+1),-1}}\right)^2 d\nu' \leqq \int_F \left(\frac{d\nu_{-\infty,0}}{d\nu^1}\right)^2 d\nu'$$

so that

$$\int_F \frac{f_{-k,0}}{f_{-k,-1}} d\nu \leqq \int_F \frac{f_{-(k+1),0}}{f_{-(k+1),-1}} d\nu \leqq \int_F \frac{d\nu_{-\infty,0}}{d\nu^1} d\nu . \tag{7}$$

In fact, we have just proved that

$$\left\{\frac{f_{-1,0}}{f_{-1,-1}}, \frac{f_{-2,0}}{f_{-2,-1}}, \cdots, \frac{d\nu_{-\infty,0}}{d\nu'}\right\}$$

is a ν-semi-martingale. Now let

$$F(m) = [\sup_{k\geqq 1} \log \frac{f_{-k,0}}{f_{-k,-1}} > m]$$

and

$$F_K(m) = [\sup_{1\leqq j<k} \log \frac{f_{-j,0}}{f_{-j,-1}} \leqq m, \log \frac{f_{-k,0}}{f_{-k,-1}} > m] .$$

On $F_k(m)$ we have

$$f_{-k,-1} \leqq 2^{-m} f_{-k,0} .$$

Hence

$$\int_{F_k(m)} f_{-k,-1} \frac{f_{-k,0}}{f_{-k,-1}} d\mu \leqq 2^{-m} \int_{F_k(m)} \left(\frac{f_{-k,0}}{f_{-k,-1}}\right)^2 d\mu$$
$$= 2^{-m} \int_{F_k(m)} \frac{f_{-k,0}}{f_{-k,-1}} d\nu .$$

Applying (7), we obtain

$$\nu[F_k(m)] \leqq 2^{-m} \int_{F_k(m)} \frac{d\nu}{d\nu'} d\nu,$$

therefore,

$$\nu[F(m)] \leqq 2^{-m} \int_{F(m)} \frac{d\nu}{d\nu'} d\nu \leqq 2^{-m} L .$$

Hence

$$\int \sup_{k\geqq 1} \log \frac{f_{-k,0}}{f_{-k,-1}} d\nu \leqq \sum_{m\geqq 0} \nu[F(m)] \leqq \sum_{m\geqq 0} 2^{-m} L < \infty .$$

Combining Lemmas 1, 2 and noting that

$$\int \frac{f_{0,n}}{f_{0,n-1}} d\nu = \int \frac{f_{-n,0}}{f_{-n,-1}} d\nu$$

(cf. Theorem 1, [6]), we obtain the following theorem.

THEOREM 1. *If there is a number L such that*

$$\int \frac{f_{0\,n}}{f_{0,n-1}} d\nu \leqq L \text{ for } n = 1, 2, \cdots \text{ then}$$

$$\int \sup_{k \geqq 1} |\log f_{-k,0} - \log f_{-k,-1}| d\nu < \infty$$

and $\{n^{-1} \log f_{0,n}\}$ *converges with* ν*-probability one.*

Extensions of Lemma 1, Lemma 2 and Theorem 1 to K-Markovian μ are immediate.

3. **The countable case.** Let X be countable with elements denoted by a. Let ν be an arbitrary stationary probability measure on $\mathscr{F}$. Let

$$P(a_0, a_1, \cdots, a_n) = \nu[x_0 = a_0, x_1 = a_1, \cdots, x_n = a_n] .$$

Let

$$H_1 = -\sum_a P(a) \log P(a) = -\int \log P(x_n) d\nu .$$

Carleson showed that

$$H_1 < \infty \tag{8}$$

implies the $L_1(\nu)$ convergence of $\{n^{-1} \log P(x_0, x_1, \cdots, x_n)\}$ [3]. Chung showed that (8) also implies the convergence with ν-probability one of $\{n^{-1} \log P(x_0, x_1, \cdots, x_n)\}$ [4]. Let μ be defined by

$$\mu[x_m = a_0, x_{m+1} = a_1, \cdots, x_n = a_{n-m}] = P(a_0)P(a_1) \cdots P(a_{n-m}) .$$

μ may be called the independent measure obtained from ν. Then $\nu_{m,n} \ll \mu_{m,n}$ with derivative

$$f_{m,n} = \frac{P(x_m, \cdots, x_n)}{P(x_m) \cdots \cdot P(x_n)}$$

and

$$\log \frac{f_{m,n}}{f_{m,n-1}} = \log \frac{P(x_m, \cdots, x_n)}{P(x_m, \cdots, x_{n-1})} - \log P(x_n) . \tag{9}$$

It follows from (9) that

$$\int (\log f_{0,n} - \log f_{0,n-1}) d\nu \leqq \int - \log P(x_n) d\nu = H_1 .$$

Hence (8) implies that (1) is satisfied, therefore $\{n^{-1} \log f_{0,n}\}$ converges in $L_1(\nu)$ by Theorem 5 [6]. Since

$$\log f_{0,n} = \log P(x_0, \cdots, x_n) + \sum_{k=0}^{n} \log P(x_k) ,$$

Carleson's theorem follows immediately. Furthermore, it follows from (9) and Lemma 1 that

$$\int \sup_{k \geqq 1} [\log \frac{P(x_{-k}, \cdots, x_{-1})}{P(x_{-k}, \cdots, x_0)} + \log P(x_0)] d\nu < \infty .$$

Hence (8) implies

$$\int \sup_{k \geqq 1} \log \frac{P(x_{-k}, \cdots, x_{-1})}{P(x_{-k}, \cdots, x_0)} d\nu < \infty$$

and Chung's theorem [4] follows.

By using a similar approach we shall give a sharpend version of Carleson's and Chung's theorems.

Let

$$P(a_0 \mid a_{-l}, \cdots, a_{-1} = \frac{P(a_{-l}, \cdots, a_{-1}, a_0)}{P(a_{-l}, \cdots, a_{-1})}$$

and let

$$\begin{aligned} H_l &= - \sum_{a_{-l}, \cdots, a_{-1}} P(a_{-l}, \cdots, a_0) \log P(a_0 \mid a_{-l}, \cdots, a_{-1}) \\ &= - \int \log P(x_n \mid x_{n-l}, \cdots, x_{n-1}) d\nu . \end{aligned}$$

H_l is nonnegative but may be $+\infty$. It is known that

$$H_1 \geqq H_2 \geqq H_3 \geqq \cdots\cdots$$

Let

$$H = \lim_{l \to \infty} H_l .$$

The limit is taken to be $+\infty$ if all H_l are $+\infty$.

THEOREM 2. *If $H < \infty$ then $\{n^{-1} \log P(x_0, \cdots, x_n)\}$ converges both in $L_1(\nu)$ and with ν-probability one.*

Proof. There is an l such that $H_l < \infty$. We define an l-Markovian measure μ on $\mathscr{F}$ as follows.

$$\mu[x_m = a_0, x_{m+1} = a_1, \cdots, x_n = a_{n-m}] = P(a_0, \cdots, a_{n-m})$$

if $n - m \leqq l$,

$$\begin{aligned} &\mu[x_m = a_0, x_{m+1} = a_1, \cdots, x_n = a_{n-m}] \\ &= P(a_0, \cdots, a_l) P(a_{l+1} \mid a_1, \cdots, a_l) \cdots\cdot P(a_{n-m} \mid a_{n-m-l}, \cdots\cdot, a_{n-m-1}) \end{aligned}$$

if $n - m > l$. It is easy to check that μ is well defined and $\nu_{m,n} \ll \mu_{m,n}$. It is clear that, if $n - m > l$,

$$\log \frac{f_{m,n}}{f_{m,n-1}} = \log \frac{P(x_m, \cdots, x_n)}{P(x_m, \cdots, x_{n-1}} - \log P(x_n \mid x_{n-l}, \cdots, x_{n-1}) .$$

The rest of the proof goes in the same manner as for the case $H_1 < \infty$ since Theorem 5 [6] and Lemma 1 of this paper remain true for l-Markovian μ.

REFERENCES

1. L. Breiman, *The individual ergodic theorem of information theory*, Ann. Math. Stat. **28** (1957), 809-811.
2. ———, *A correction to the individual ergodic theorem of information theory*, Ann. Math. Stat., **31** (1960), 809-810.
3. L. Carleson, *Two remarks on the basic theorems of information theory*, Math. Scand., **6** (1958), 175-180.
4. K. L. Chung, *A note on the ergodic theorem of information theory*, Ann. Math. Stat., **32** (1961), 612-614.
5. J. L. Doob, *Stochastic Processes*, John Wiley and Sons Inc., New York, 1953.
6. Shu-Teh C. Moy, *Generalizations of Shannon-McMillan theorem*, Pacific J. Math., 11 1961), 705-714.

SYRACUSE UNIVERSITY

AN IMBEDDING SPACE FOR SCHWARTZ DISTRIBUTIONS

DONALD E. MYERS

1. Introduction. We consider here a facet of the problem of justifying the methods of the operational calculus and in particular the use of the "Dirac Delta Function". L. Schwartz's "Theorie des Distributions" [6] is the most complete exposition to date on generalized functions but the operational calculus as such is largely omitted. B. Van Der Pol [8] discusses the latter but not in the context of distributions. Ketchum and Aboudi [4] suggested using unilateral Laplace Transforms to construct a link between Schwartz's theory and the operational calculus. This paper will enlarge on the latter suggestion. Two principal results are obtained. An imbedding space is constructed and a comparison between the topologies is made.

Let S denote the strip $\sigma_1 < R(z) < \sigma_2$, in the complex plane. Consider the one parameter family of functions $\{e^{zt}\}$, where the parameter z ranges over S and $-\infty < t < \infty$. This family is not a linear space but each member possesses derivatives of all orders. In a manner analogous to Schwartz we define an L_S-Distribution to be an analytic complex-valued functional on the above family of functions, where by analytic we mean with respect to the parameter z. If α is any complex scalar and F, σ are two such functionals then we require that $F \cdot e^{zt} + \sigma \cdot e^{zt} - (F + \sigma) \cdot \sigma^{zt}$, and $(\alpha F) \cdot e^{zt} = F \cdot (\alpha e^{zt})$. The latter property then allows us to define the derivative in a manner similar to that of Schwartz, that is $F' \cdot e^{zt} = F \cdot (e^{zt})' = F \cdot ze^{zt} = zF \cdot e^{zt}$. It also follows that the Laplace Transform supplies an integral representation of some of the functionals. The other L_S-Distributions define generalized functions for similar integral representations. That is, each function analytic for $z \in S$ has for its values, the values of an L_S-Distribution acting on a function e^{zt} and the L_S-Distribution has an integral representation utilizing the symbolic inverse Laplace Transform of the analytic function. In most of this paper we deal only with analytic functions whose inverse transforms exist but the definitions and theorems will be stated without this restriction where possible. Following a practice used by other authors, we will call the inverse Laplace Transform, symbolic or not, an L_S-Distribution rather than the functional. Because of the relation between the functional and an analytic function we concentrate on the latter and utilize the already known properties of such functions. By emphasizing the integral representations rather than the functionals we utilize the

Received January 3, 1961. Based on the author's thesis at the University of Illinois, 1960 (unpublished).

Riesz Representation Theorem for continuous linear functionals to establish a correspondence to Schwartz Distributions.

As stated above each functional has a representation by an analytic functions, using this we will define convergence in a fashion similar to that of Schwartz. That is, a sequence of L_S-Distributions will converge if the sequence of values, when operating on an arbitrary member of the one-parameter, converges. Because of the parameterization this definition can be stated directly in terms of the representations by the analytic functions.

2. L_S-Distributions.

DEFINITION 1. If an L_S-Distribution is determined by an analytic function $f(z)$, then $f(z)$ is its bilateral Laplace Transform. Denote this L_S-Distribution by $[f(z)]_t$ or f_t. Further, abbreviate L_S-Distribution by L_S-*D*.

DEFINITION 2. The derivative of an L_S-D, $[f(z)]_t = f_t$ is the L_S-D, $[zf(z)]_t = f_t'$. For a fixed S, the set of all L_S-D's is metrized by a Frechet type metric on the transforms. See [7], page 137. For a pair of functions $f(z), g(z)$ analytic in S, denote the metric by $N_S(f, g)$. The following property of this metric could have been used a definition since it is the only property used in this paper.

THEOREM 3. *A sequence of functions, all analytic in S, converges with respect to the metric N_S if and only if the sequence converges uniformly on every compact subset of S.*

DEFINITION 4. $\rho_S(f_t, g_t) = N_S(f, g)$ where f_t, g_t are the L_S-D's whose transforms are $f(z), g(z)$ respectively.

DEFINITION 5. If $f(z)$, analytic in S, is the bilateral Laplace Transform of a point-function $F(t)$, then $F(t)$ is called a Point-Function L_S-D or P.F.L_S-D.

THEOREM 6. *If $F_i(t)$, $i = 0, 1, 2, 3, \cdots$ all possess bilateral Laplace Transforms analytic in a strip S, $\sigma_1 < R(z) < \sigma_2$, and*

$$\int_0^\infty | e^{-\sigma_1 t} F_i(t) |^2 \, dt < \infty \, ,$$

$$\int_{-\infty}^0 | e^{-\sigma_2 t} F_i(t) |^2 \, dt < \infty$$

for all i, then let

$$d(F_k, F_j) = \left[\int_0^\infty e^{-2\sigma_1 t} \mid F_k(t) - F_j(t) \mid^2 dt\right]^{1/2} + \left[\int_{-\infty}^0 e^{-2\sigma_2 t} \mid F_k(t) - F_j(t) \mid^2 dt\right]^{1/2} .$$

If $d(F_i, F_0) \to 0$ as $i \to \infty$ then

$$F_i(t) \to F_0(t) \text{ as P.F.L}_S\text{-D's.}$$

Proof. Write the transform of $F_i(t) - F_0(t)$ as

$$\int_{-\infty}^0 e^{-t(z-\sigma_2)} e^{-\sigma_2 t}[F_i(t) - F_0(t)]dt + \int_0^\infty e^{-t(z-\sigma_1)} e^{-\sigma_1 t}[F_i(t) - F_0(t)]dt .$$

By the Cauchy-Schwartz Inequality

$$\begin{aligned} &| f_i(z) - f_0(z) | \\ &\leqq \left[\frac{1}{2[\sigma_2 - R(z)]}\int_{-\infty}^0 e^{-2\sigma_2 t} \mid F_i(t) - F_0(t) \mid^2 dt\right]^{1/2} \\ &+\left[\frac{1}{2[R(z) - \sigma_1]}\int_\sigma^\infty e^{-2\sigma_1 t} \mid F_i(t) - F_0(t) \mid^2 dt\right]^{1/2} . \end{aligned}$$

If

$$g(z) = \max_{z\in S}\left[\sqrt{\frac{1}{2(R(z) - \sigma_1)}} , \sqrt{\frac{1}{2(\sigma_2 - R(z))}}\right]$$

then

$$| f_i(z) | \leqq g(z) g(F_i, F_0)$$

and hence $f_i(z) \to f_0(z)$ uniformly on each compact subset in S if $d(F_i, F_0) \to 0$ as $i \to \infty$.

An interpretation of Theorem 6 might be that if $\{e^{-\sigma_1 t}F_i(t)\}$ converges in $L_2[0, \infty]$ to $e^{-\sigma_1 t}F_0(t)$ and $\{e^{-\sigma_3 t}F_i(t)\}$ converges in $L_2[\infty, 0]$ to $e^{-\sigma_2 t}F_0(t)$ and each $F_i(t)$ has a bilateral Laplace Transform then the sequence of P.F.L$_S$-D's converges with respect to the metric ρ_S.

THEOREM 7. *Let $f_j(z), j = 0, 1, 2, 3, \cdots$ be an infinite sequence of functions analytic in a strip S, $\sigma_1 < R(z) < \sigma_2$, and further suppose there exists a C such that $| f_j(z) | < Ce^{-\eta_0 |I(z)|}$ for some $\eta_0 > 0$, in all of S. If $N_S(f_j, f_0) \to 0$ as $j \to \infty$ then $F_j(t) \to F_0(t)$ uniformly on every bounded interval in the t-line. $F_j(t)$ denotes the inverse bilateral Laplace Transform of $f_j(z)$.*

Proof. The hypothesis is sufficient to ensure the existence of the inverse transform of each $f_j(z)$, [2]. That is,

$$F_j(t) = \int_{-\infty}^{\infty} \frac{e^{xt}}{2\pi}[e^{ity} f_j(x+iy)]dy$$
$$\text{for } \sigma_1 < x < \sigma_2 .$$

Then

$$\begin{aligned} | F_j(t) - F_0(t) | \leqq \frac{e^{xt}}{2\pi} \left| \int_{-\infty}^{-\rho} e^{ity}[f_j(x+iy) - f_0(x+iy)]dy \right| \\ + \frac{e^{xt}}{2\pi} \left| \int_{-\rho}^{p} e^{ity}[f_j(x+iy) - f_0(x+iy)]dy \right| \\ + \frac{e^{xt}}{2\pi} \left| \int_{\rho}^{\infty} e^{ity}[f(x+iy) - f_0(x+iy)]dy \right| . \end{aligned}$$

For $\varepsilon > 0$, and $a < t < b$, let ρ be such that $\dfrac{2e^{\sigma_2 b}C' e^{-\eta_0 \rho}}{\pi \eta_0} < \varepsilon/2$ and J such and J such that $\left[\dfrac{e^{\sigma_2 b}}{\pi}\right] N_S(f_j, f_0) < \varepsilon/2$ for $j > J$, then

$$| F_j(t) - F_0(t) | < \varepsilon \text{ for } j > \text{ and}$$
$$a \leqq t \leqq b .$$

THEOREM 7.1. *If in Theorem* 7, $\sigma_1 < 0 < \sigma_2$ *then* $F_j(t) \to F_0(t)$ *uniformly for* $-\infty < t < \infty$.

DEFINITION 8. For each L_S-D, f_t, define f_{t+h} to be $[e^{hz} f(z)]_t$.

THEOREM 9. *If* f_t *is an arbitrary* L_S-D *then*

$$\rho\left(\frac{f_{t+h} - f_t}{h}, f_t'\right) \to 0 \; as \; h \to 0 .$$

Proof. By definition

$$f_t' = [z f(z)]_t$$
$$f_{t+h} = [e^{hz} f(z)]_t$$

so that

$$\begin{aligned} &\frac{f_{t+h} - f_t}{h} - f_t' \\ &= \left[\frac{e^{hz} f(z) - f(z) - z f(z)}{h}\right]_t \\ &= \left[f(z)\left[\frac{hz^2}{2!} + \frac{h^2 z^3}{3!} + \cdots\right]\right]_t \end{aligned}$$

and since

$$\frac{hz^2}{2!} + \frac{h^2z^3}{3!} + \cdots \to 0$$

uniformly on each compact set in S as $h \to 0$, the theorem is proved.

DEFINITION 10. An L_S-D f_t is said to have point-values $F(t)$ for $c < t < d$ if there exists a $\sigma(t)$ such that for some $k, f(z) = z^k g(z), g(z)$ being the bilateral Laplace Transform of $\sigma(t)$ and finally that $\sigma^{(k)}(t) = F(t)$ for $(c < t < d)$.

For example $[1]_t$ has zero-point values in every open interval, in the t-line, that does not contain the point $t = 0$. Since

$$H(t) = \begin{cases} 1, & t > 0 \text{ has for} \\ 0, & t > 0 \end{cases}$$

its transform $1/z$ and $[1]_t = [z\frac{1}{z}]_t$ finally $H'(t) = 0$ for all $t \neq 0$. $[1]_t$ is the "Dirac Delta Function".

THEOREM 11. *If* $\{{}_nf_t\}$ *is a sequence of* L_S-D*'s converging to an* L_S-D ${}_0f_t$ *then* $\{{}_nf_t^{(k)}\}$ *converges to* ${}_0f_t^{(k)}$ *for all* $k = 0, 1, 2, \cdots$

Proof. By definition $\{{}_nf_t\}$ converges to

$${}_0f_t \leftrightarrows \max_{z \in K} |\, {}_nf(z) - {}_0f(z) \,| \to 0$$

as $n \to \infty$ for all compact $K \subset S$. Since in the complex plane, a set is compact $\leftrightarrows$ if it is closed and bounded, there exists an M_K for each $K \,|\, z \,| \leqq M_K$ for $z \in k$.

Then $\max_{z \in K} |{}_nf(z) - {}_0f(z)\,| \to 0 \leftrightarrows$

$$|\, M_K \,|^k \,|\, {}_nf(z) - f(z) \,| \to 0 \quad \text{as} \quad n \to \infty \quad \text{for each fixed}$$

positive integer k apply Definition 2.

EXAMPLE. The following will be used as a counter-example in the last section. Consider the Taylor-expansion for

$$e^{-z} = 1 - z + \frac{z^2}{2!} - \frac{z^3}{3!} + \cdots + \frac{(-1)^n z^n}{n!} + \cdots$$

$[e^{-z}]_t$ is the "Delta Dirac Function" translated so that L_S-D has zero point-values for all t except for

$$t = 1. \quad [(-1)^n \frac{z^n}{n!}]_t = \frac{(-1)^n}{n!}[z^n]_t = \frac{(-1)}{n!}[1]_t^{(n)} .$$

The latter L_S-D has zero point-values for all $t \neq 0$. Since the Taylor Series converges for all z and hence uniformly for compact sets the series of L_S-D's converges

3. General L_S-Distribution. The set of all L_S-D's for any fixed S does not contain a subset isomorphic with the set point-functions having pointvalues a.e. For example, the function $F(t) = 1$ does not have a transform even though it is continuous for $-\infty < t < \infty$. However each member of the sequence of functions

$$\begin{aligned} F_i(t) &= 1, && (-i \leqq t \leqq i) \\ &= 0, && (t > i, \quad t < -i) \\ i &= 0, && (0, 1, 2, 3, 4, \cdots) \end{aligned}$$

does possess a transform. Further for each open interval (c, d) only a finite number of the elements of the sequence have different point-values than $F(t)$ in (c, d). The sequence represents $F(t)$.

DEFINITION 12. A sequence $\{{}_nf_t\}$ of L_S-D's is called Fundamental if for each open interval (c, d) there exists an integer N such that for $n > N\, {}_nf_t - {}_{n+p}f_t$, $p = 0, 1, 2, 3, \cdots$ is an L_S-D with zero point-values in (c, d). Fundamental sequence of L_S-D in abbreviated by F.S.S.

DEFINITION 13. Two F.S.S.'s, $\{{}_ng_t\}$ and $\{{}_nf_t\}$ are said to be Similar if for each open interval (c, d) there exists an integer N such that for $n > N\, {}_ng_t - {}_nf_t$ is an L_S-D with zero pointvalues in (c, d).

LEMMA 14. *The Similarity defined in Definition* 13 *for pairs of* F.S.S.'*s is an Equivalence relation and is invariant under addition and differentiation.*

THEOREM 15. *The equivalence classes under the Similarity relation are called* G.L_S-D's *or* General L_S-D's. *They form an Abelian group, closed with respect to scalar multiplication and differentiation.*

The Representation Theorem.

THEOREM 16. *Let A denote the entire complex plane, then there is a subset, D, of the set of all G.* L_A-D'*s that is isomorphic with the set of all Schwartz Distributions. The isomorphism is invariant with respect to addition, scalar multiplication and differentiation.*

(a) By definition, a Schwartz Distribution is a linear functional on the space of infinity differentiable point-functions with compact supports and is continuous when restricted to the set. Each Schwartz Distribution has an integral respresentation when restricted to a bounded closed interval, [3]. This representation has the form

$$D_t(\phi) = (-1)^r \int_a^b F(t)\phi^{(r)}_{(t)} dt$$

where $F(t)$ is continuous on $[a, b]$ and r is an integer dependent on $[a, b]$ and the distribution D_t. $\phi(t)$ is any function with support the closed interval $[a, b]$. Let $[a_n, b_n]$ be a sequence of intervals where $-\infty \leftarrow a_{n+1} \leqq a_n \leqq b_n \leqq b_{n+1} \rightarrow \infty$ as $n \rightarrow \infty$. For each n there is an $F_n(t)$ and an r_n. Let

$$\begin{aligned} G_n(t) &= (-1)^{r_n} F_n(t), && (a_n \leqq t \leqq b_n) \\ &= 0, && (t > b_n,\ t < a_n)\,. \end{aligned}$$

Then let

$$f_n(z) = \int_{a_n}^{b_n} e^{-zt} z^{r_n} G_n(t) dt\,.$$

It remains to be shown that the sequence $\{_n f_t\}$ is an F.S.S. and that the equivalence class is independent of the sequence of covering intervals. The G.L$_S$-D determined is the representative of D_t.

(b) Let I be an arbitrary open interval in the t-line, denoted (c, d). There exists an N then such that for $n > N[a_n, b_n] \supset (c, d)$. Let $F_n(t)$, $F_{n+p}(t)$, r_n, r_{n+p} be the continuous functions and integers given for the representation of the distribution D_t on the intervals $[a_n, b_n]$ and $[a_{n+p}, b_{n+p}]$ respectively. Using Halperin's notation, let $S[a_n, b_n]$ denote the class of testing functions associated with the interval $[a_n, b_n]$ that is, if $\phi \in S[a_n, b_n]$ then $\phi^{(k)}(t)$ is zero for $t \notin [a_n, b_n]$, and $\phi^{(k)}(t)$ exists for all $t \in [a_n, b_n]$ for $k = 0, 1, 2, 3, \cdots$. It is seen that $S[a_n, b_n] \subset S[a_{n+p}, b_{n+p}]$. If $\phi \in S[a_n, b_n] \subset S[a_{n+p}, b_{n+p}]$ then

$$\begin{aligned} D_t(\phi) &= (-1)^{r_n} \int_{a_n}^{b_n} F_n(t)\phi^{(r_n)}(t) dt \\ &= (-1)^{r_{n+p}} \int_{a_{n+p}}^{b_{n+p}} F_{n+p}(t)\phi^{(r_{n+p})}(t) dt \end{aligned}$$

or

$$\int_{a_n}^{b_n} [F_n(t)\phi^{(r_n)}(t) - (-1)^{r_{n+p}-r_n} F_{n+p}(t)\phi^{(r_{n+p})}(t)] dt = 0$$

since

$$\phi^{(k)}(t) = 0 \text{ for } t \notin [a_n, b_n]\,.$$

Let

$$T \cdot F_n(t) = \int_{a_n}^{t} F_n(x) dx$$

$$T^2 \cdot F_n(t) = T \cdot [T \cdot F_n(t)] .$$

Then if $r_{n+p} \geqq r_n$

$$0 = \int_{a_n}^{b_n} [(-1)^{r_{n+p}-r_n} T^{r_{n+p}-r_n} \cdot F_n(t)\phi(t)^{(r_{n+p})} - F_{n+p}(t)\phi^{(r_{n+p})}(t)]dt .$$

It follows then that

$$(-1)^{r_{n+p}-r_n} T^{r_{n+p}-r_n} \cdot F_n(t) - F_{n+p}(t)$$

is a polynomial $Q_m(t)$ of degree $m \leqq r_{n+p} - 1$ for $a_n \leqq t \leqq b_n$. Similar results are obtained if $r_n \geqq r_n + p$.

$$ {}_{n+p}f(z) - {}_nf(z) = \int_{a_{n+p}}^{b_{n+p}} e^{-zt} z^{r_{n+p}} G_{n+p}(t)dt - \int_{a_n}^{b_n} e^{-zt} z^{r_n} G_n(t)dt .$$

Using $Q_m(t)$ we have

$$ {}_{n+p}f(z) - {}_nf(z) = \int_{a_{n+p}}^{a_n} e^{-zt} z^{r_{n+p}} G_{n+p}(t)dt + \int_{a_n}^{b_n} e^{-zt} z^{r_{n+p}} Q_m(t)dt + \int_{b_n}^{b_{n+p}} e^{-zt} z^{r_{n+p}} G_{n+p}(t)dt .$$

The first integral can be considered as the transform of the r_{n+p}th derivative of a function with zero point-values exterior to the interval $[a_{n+p}, a_n]$ and hence interior to the interval (c, d). The second integral can be considered as the transform of the r_{n+p}th derivative of a polynormial of degree less than or equal to $r_{n+p} - 1$. Hence the L_S-D determined has zero pointvalues on the interior and exterior of the interval $[a_n, b_n]$ and hence on the interior of (c, d). This L_S-D may not have zero point-values at $t = a_n$ or $t = b_n$. Finally then the third integral considered as a transform determines an $L_S - D$ with zero pointvalues exterior to the interval $[b_n, b_{n+p}]$ and hence on the interior of (c, d). ${}_{n+p}f_t - {}_nf_t$ is an L_S-D with zero pointvalues on the interior of (c, d), if $n > N$. The sequence constructed in part (a) is an F.S.S.

(c) Suppose $[a_n, b_n]$ and $[c_n, d_n]$ are two expanding sequence of closed intervals covering the real line. Let $\{{}_nf_t\}$ and $\{{}_ng_t\}$ be the F.S.S.'s obtained from the consturuction of part (a) using the former sequences. Let I be an arbitrary open interval in the t-line. Then there exists integers N_I^1, N_I^2 such ${}_{n+p}f_f - {}_nf_t$ for $n < N_I^1$ and ${}_{n+p}g_t - {}_ng_t$ for $n \leqq N_I^2$ L_S-D's with zero point-values for $t \in I$. Further there exists an integer $M \ni [a_n, b_n] \subset [C_N, d_N]$ for $n \geqq M$ and $N = N_I^2$. Consider

$$ {}_nf_J - {}_ng_t = {}_nf_t - Kf_t + Kf_t - Kg_t + Kg_t - {}_ng_t $$

where K is the largest of N_I^1, N_I^2, M. For $n > K$ then the first difference on the right is an L_S-D with zero point-values for $t \in I$ since $\{{}_nf_t\}$ is an F.S.A. The second difference can be shown to be an L_S-D with zero point-values for $t \in I$ by the method of part (b). Finally the third difference is an L_S-D with zero point-values for $t \in I$ since $\{{}_ng_t\}$ is F.S.A. The two F.S.A.'s are similar and hence determine the same G.L_A.-D. The correspondence between the Schwartz Distribution and the G.L_A.-D. is one-to-one. The invariance of this isomorphism with respect to addition, differentiation and scalar multiplication follows from Lemma 14.

4. A Topology for G.L-D's

Definition 16. An F.S.S. $\{{}_nf_t\}$ is said to have point-values $F(t)$ for $t \in (c, d)$, an open interval, if there exist an integer $N_{(c,d)}$ such that for $n > N$, ${}_nf_t$ is an L_S-D possessing pointvalues $F(t)$ for $t \in (c, d)$. A G.L_S-D is said to have pointvalues $F(t)$ for $t \in (c, d)$ if there is an F.S.S. unit equivalence class possessing that property.

Definition 17. Let $\{{}_nf_t\}_1, \cdots, \{{}_nf_t\}_i, \cdots$ be a sequence of F.S.S.'s. Denote the nth element of the jth F.S.S. by $({}_nf_t)_j$. Then sequence is said to converge to the sequence of L_S-D's $\{{}_nf_t\}_0$, if for $\varepsilon > 0$ there exist integers N_ε, J_ε such that $\rho\ [({}_nf_t)_j, ({}_nf_t)_0] < \varepsilon$ when $n > N_\varepsilon, j > J_\varepsilon$.

Definition 18. Let $D_1, D_2, \cdots$ be a sequence of G.L_S-D's. Further suppose $L_1, L_2, \cdots$ is a sequence of F.S.S.'s each having support $[a, b]$ and that for each $j = 1, 2, \cdots L_j$ represents D_j is (a, b). That is, for some F.S.S. in D_j, the difference of L_j and the F.S.S. has zero point-values in (a, b). Then if $L_1, L_2, \cdots$ is convergent in the sense of Definition 17, $D_1, D_2, \cdots$ is said to converge to D_0 where D_0 is the G.L_S-D. determined by L_0.

Theorem 19. *If a sequence of Schwartz Distributions is convergent in an open interval (a, b) in Schwartz's sense then the sequence of* G.L_S-D'*s isomorphic to the respective Schwartz Distributions is convergent in the interior of every closed interval contained in (a, b).*

Proof. Let $D_1, D_2, \cdots$ be a sequence of Schwartz Distributions convergent in Schwartz's sense in (a, b). For any closed interval $[c, d]$ contained in (a, b) there exists a sequence of representation

$$ D_i(\phi) = (-1)^{r_c}\int_c^d F_i(t)\phi(t)^{(r_c)}dt $$

for $\phi \in S[c, d]$. Since the sequence $D_1, D_2, \cdots$ is convergent there exists one integer r_c which can be used in all the representations and also the limit representation.

For each j, construct the F.S.A. $\{_nf_t\}_j$ where $\{_nf_t\}_j = (f_t)_j$ and

$$(f)_j = \left[\int_c^d e^{-zt} G_j(t)_z{}^{r_c} dt\right]_t$$

where

$$\begin{aligned} G_j(t) &= (-1)^{r_c} F_j(t) , & c \leqq t \leqq d \\ &= 0 , & \text{otherwise} . \end{aligned}$$

Since the sequence of Distributions is convergent

$$\lim T^{r_c} F_j(t) = F_0(t) \text{ [uniformly } [c, d]] \; j \to \infty$$

$$\begin{aligned} K &= \left| \int_c^d z^{r_c} e^{-zt} \left[T^{r_c} G_j(t) - G_0(t) \right] dt \right| \\ &\leqq | d - c | e^{-c\sigma_1} | T^{r_c} G_j(t) - G_a(t) || d^{r_0} | . \end{aligned}$$

It follows that

$$K = | (f(z))_j - (f(z))_0 | \leqq M\varepsilon$$
$$\text{for } j > J_\varepsilon, M = | d - c | e^{-c\sigma_1} | \sigma_2{}^{r_c} | .$$

Then $(f(z))_j \to (f(z))_0$ uniformly on every compact set in the strip $\sigma_1 \leqq R(z) \leqq \sigma_2$ and hence in the metric ρ_S. By definition then $(f_t)_j \to (f_t)_0$ and hence

$\{_nf_t\}_I, \cdots, \{_nf_t\}_j, \cdots$ converges to $\{_nf_t\}_0$ in the interior of (c, d). The sequence of G. L_S-D's converges for $t \in (c, d)$.

The example given earlier for a series representation of "Delta" Distribution with a discontinuous at $t = 1$ converges in the sense defined herein but not in Schwartz's sense. The L_S-D $[e^{-z}]_t$ and its series representation furnish a solution to the differential equation

$$f(x-1) = f(x) + f'(x) + \frac{f''(x)}{2!} + \cdots + \frac{f^{(n)}(x)}{n!} + \cdots$$

BIBLIOGRAPHY

1. S. Bochner, *Lectures on Fourier Integrals*, Princeton University Press, Princeton, 1956.
2. G. Doetsch, *Theorie und Andwending der Laplace Transformation*, Juluis Springer, Berlin, 1937.
3. I. Halperin, *Introduction to the Theory of Distributions*, University of Toronto Press, Toronto, Canada, 1952.
4. P. Ketchum, and R. Abondi, *Schwartz Distributions*, Proceedings of the Second Midwest Symposium on Circuit Theory, Michigan State University, 1956.
5. D. Myers, *An Imbedding Space for Schwartz Distributions*, Thesis University of Illinois,

1960.
6. L. Schwartz, *Theorie des Distributions, I et II, Herman et Cie,* Paris, 1950-1951.
7. W. Thron, *Introduction to the Theory of Functions of a Complex Variable,* John Wiley, New York, 1953.
8. B. Van Der Pol, *Operational Calculus,* Cambridge University Press, 1950.

CATEGORY METHODS IN RECURSION THEORY[1,2]

J. MYHILL

The heavy symbolism used in the theory of recursive functions has perhaps succeeded in alienating some mathematicians from this field, and also in making mathematicians who are in this field too embroiled in the details of thier notation to form as clear an overall picture of their work as is desirable.[3] In particular the study of degrees of recursive unsolvability by Kleene, Post, and their successors[4] has suffered greatly from this defect, so that there is considerable uncertainty even in the minds of those whose speciality is recursion theory as to what is superficial and what is deep in this area.[5] In this note we shall examine one particular theorem (namely the Kleene-Post theorem asserting the existence of incomparable degrees[6]) and show that it is a special case of a very easy and well-known theorem of set-theory. Exposition will be such as to require (except in a few footnotes) no preliminary acquaintance with recursive matters. It is to be hoped that some mathematicians in other areas may be stimulated by this exposition to try their hand at some open questions about recursive functions: it is to be hoped also that they will not carry away the impression that all of recursion theory is as trivial as this paper will show the Kleene-Post theorem to be.

First let me describe in an informal way what relative recursiveness is. The only properties of it which we shall need will be apparent from this informal discussion.

Denote by ε the set of all nonnegative integers. A *function* shall mean a number-theoretic function $f: \varepsilon \rightarrow \varepsilon$. A function is called *recursive* if it can be computed in an effective (mechanical) manner: we shall not need the details of the definition.[7] Sometimes two functions f and g are so related that the function f can be calculated in an effective

Received December 27, 1960.

[1] Composition of this paper was supported by NSF grant G-7277.

[2] Category methods have also been used by the author in **[12]**, and form the basis of the entire treatment of degrees in **[3]**.

[3] A related (but much deeper) contribution to the methodology of recursion theory has made by Addison, e.g., in **[1]**.

[4] See, e.g., **[7]**, **[14]**, **[15]**, **[19]**. A sadly neglected paper in the same area which completely avoids these unnecessary complications is Lacombe **[10]**.

[5] The principal result of Spector **[19]** (minimal non-recursive degrees) is probably 'deep' in this sense, as is likewise the Friedberg-Mučnik proof (**[4]**, **[11]**) of the existence of incomparable degrees *of recursively enumerable sets.*

[6] Strictly speaking, the Kleene-Post theorem (**[7]**, p. 390) gives more information than our version, since it gives incomparable degrees $<\mathbf{0}'$. But this result too can be obtained by a category argument, as I shall show in a later publication.

[7] Cf., e.g., Davis **[2]**, p. 41.

(mechanical) way apart from requiring, for the computation of each particular function-value $f(n_0)$, a finite amount of information concerning values of the function g: in this case we say that f is *recursive in* or *relative to* g. The simplest way to envisage this relation is probably in terms of Turing machines.[8] We say that f is recursive in g if there exists a Turing machine with input and output tapes such that if the values $g(0), g(1), g(2), \ldots$ are fed in that order into the input, then for every nonnegative integer n the unique true statement of the form $f(n) = m$ will appear after a finite time on the output tape (and no false statement of that form will ever appear). Another characterization which may also aid the intuition is the following: f is recursive in g if there is a formal system[9] Σ such that every true statement of the form $f(n) = m$, and no false statement of that form, is deducible in Σ from a finite number of true statements of the form $g(x) = y$. The exact definitions of Turing machine and formal system are quite irrelevant for our purposes: all that matters is that

(1) only finitely many values of g are used to compute any value of f and

(2) the total number of Turing machines or formal systems is countable.

In both cases (2) is a consequence of the fact that the process of computation of one function from another can be described by a finite description using only symbols belonging to a finite alphabet fixed in advance; the same will be true if we characterize relative recursiveness in some way other than by Turing machines or formal systems.[10]

To every Turing machine or formal system corresponds uniquely a mapping Φ from functions to functions, called a *partial recursive operator*. It is important to notice that certain such Φ may not be defined for all functions as arguments. It may well be that a certain Turing machine $\boldsymbol{T}$, on being supplied with the values of a certain function g, will print statements of the form $f(m) = n$ on its tape only for certain m. In that case we say that $\boldsymbol{T}$ computes only a partial function from g. We regard the operator Φ as defined on the family of all those g from which $\boldsymbol{T}$ computes a full (everywhere defined) function. For example, suppose we consider the mapping which assigns to every function f the function $< \Phi f >$ such that

$$< \Phi f > (x) = (\mu y)\,(f(y) = 0);^{11}$$

[8] Davis [**2**], Ch. 1-2.

[9] For 'formal system' see Davis [**2**], Ch. 6 and 8, Smullyan [**17**] passim. The first use of formal systems to define partial reqursive functionals seems to date from Myhill-Shepherdson [**13**], p. 315, where we followed a suggestion of Marian Boykan (now Pour-El).

[10] E.g., by systems of recursion equations (Kleene [**5**], pp. 326-327).

[11] $(\mu y)\,(\ldots y \ldots)$ denotes the least y satisfying the condition $\ldots y \ldots$ if such exist, and otherwise is meaningless.

then Φ is a partial recursive operator whose domain of definition is the family of all functions which vanish for at least one value of the argument (and whose range is the family of all constant functions).

We denote by $\mathscr{T}$ the family of all functions, and we topologize it as the product of countably many replicas of the integers each with the discrete topology. This corresponds to the metric

$$\rho(f, g) = \frac{1}{(\mu x)(f(x) \neq g(x)) + 1}$$

or 0 if $f = g$. It is well-known[12] that this is a complete metric space, hence of second category on itself. This is the basic fact that we shall use in what follows.

By a *finite function* we mean a mapping of a finite subset of ε into ε; if f_0 is such a function, we define, $\mathscr{N}(f_0)$ as the family of all (full) functions which extend f_0. We can take as a (countable) basis for $\mathscr{T}$ the collection of all families $\mathscr{N}(f_0)$. $\Phi : \mathscr{F} \rightarrow \mathscr{T}$ with $\mathscr{F} \subseteq \mathscr{T}$ is continuous (in the induced topology on $\mathscr{F}$)[13] just in case

$$f \in \mathscr{F}, < \Phi f > (x) = y \rightarrow (\exists f_0)\ (f \in \mathscr{N}(f_0) \text{ and } (\forall f')$$
$$(f' \in \mathscr{N}(f_0), f' \in \mathscr{F} \rightarrow < \Phi f' > (x) = y)) ,$$

i.e., if and only if any value $< \Phi f >$ is determined by finitely many values of f. In view of what was said above it follows that *all partial recursive operators are continuous*[14] (on their domain). For use later on we observe also that the domain of definition of such an operator is a $G\delta$ set; this too is an immediate consequence of the preceding informal remarks.

We write $f \leqq g$ if f is recursive in $g, f < g$ if $f \leqq g$ but not $g \leqq f$. The relation $f \leqq g$ is a pre-order; hence its symmetrization $f \equiv g$ (i.e., $f \leqq g$ and $g \leqq f$) is an equivalence relation. The equivalence classes into which it divides $\mathscr{T}$ are called *degrees*; we call one degree $\mathscr{D}$ *lower than* another degree $\mathscr{D}^*$ and write $\mathscr{D} < \mathscr{D}^*$ if $f < g$ for all (equivalently, for some) $f \in \mathscr{D}, g \in \mathscr{D}^*$.

Now we can prove the existence of incomparable degrees. Observe first the there are exactly c degrees, since there are c functions and at

12 Sierpinski [**16**], p. 191.

13 A partial recursive operator defined on a dense subset of $\mathscr{T}$ need not have a continuous extension to the whole space (Kleene [**5**], p. 685); and even when it does this extension need not be partial recursive (Lacombe [**10**], p. 155, Theorem XIX). Hence it will not suffice for our purposes to consider only everywhere defined operators.

14 This observation is essentially Kleene's (cf. the proofs of Theorems XXIa and XXVI in [**5**], pp. 339, 348-349); that the property in question amounted to continuity was observed apparently independently by Lacombe (in a series of papers in *Comptes Rendus* going back at least to 1953) and later by Trahtenbrot [**20**]. Davis ([**2**], pp. 164 seqq.) oddly uses the word 'compact' to mean 'continuous'.

most (in fact, exactly, but we shall not need this) $\aleph_0$ functions belonging to any given degree. Observe also that there are at most $\aleph_0$ degrees lower than a given degree. For let $\mathscr{D}^*$ be a degree; then if f belongs to a degree lower than $\mathscr{D}^*$ it must be of the form $< \Phi g >$ where $g \in \mathscr{D}^*$ and Φ is partial recursive. But there are only countably many g's in $\mathscr{D}^*$ and only countably many Φ's; hence there are only countably many functions of degree $< \mathscr{D}^*$ and a fortiori only countably many degrees $< \mathscr{D}^*$.[15] This gives a plausibility argument for the existence of incomparable degrees, for if every two degrees were comparable we would have a simply ordered set of the power of the continuum in which each element had only a (finite or) countable number of predecessors; and this is easily seen[16] to imply the continuum hypothesis.

The continuum hypothesis is equivalent[17] to the assertion that the plane is the union of countable many curves (where a curve is the set of all points $(x, f(x))$ or of all points $(f(x), x)$ for some (not necessarily everywhere defined) real function f). We know also that the plane is not the union of countably many *continuous* curves,[18] since each such curve is nowhere dense and the plane is of second category on itself. These considerations yield at once the existence of incomparable degrees. If every two degrees were comparable the space $\mathscr{T}^2$ would be the union of all curves $\{(f, < \Phi f >)\}$ and $\{(< \Phi f >, f)\}$ with Φ partial recursive. But this is impossible because as we have seen each of these curves is continuous and hence by a classical argument nowhere dense,[19] and because $\mathscr{T}^2$, like $\mathscr{T}$, is a complete metric space and hence of second category on itself, q.e.d.

Now we use the same method to establish a stronger statement which answers a question rather recently raised (and still more recently settled) by Shoenfield.[20] Do there exist uncountably many degrees any two of which are incomparable? We shall obtain an affirmative answer to this question using only the hypotheses that $\mathscr{T}$ is a complete metric space and hence of second category on itself, and that there are only countably many partial recursive operators each of which is continuous

[15] For the lowest degree (that to which recursive functions belong) there are of course *no* degrees lower. There are also degrees than which only a finite nonzero number of degrees are lower (Spector [**19**], Theorem 4).

[16] Sierpinski [**16**], p. 23.

[17] Sierpinski [**16**], p. 11.

[18] Nor of countably many *measurable* curves (i.e., Lebesgue measurable in the plane); this is the foundation of Spector's proof in [**18**] of the existence of incomparable hyperdegrees. (Measure arguments have to replace category arguments in the study of hyperdegrees because hyperarithmetic operators are in general discontinuous.)

[19] The only hypothesis needed is that $\mathscr{T}$ is a Hausdorff space with no isolated points.

[20] Raised in [**15**], settled in [**14**]. More recently Sacks has obtained (unpublished) a *continuum* number of pairwise incomparable degrees and Lacombe and Nerode (unpublished) have obtained a continuum number of *independent* (and minimal non-recursive) degrees (see [**7**], p. 383 for the definition of independence).

in the topology induced on its domain.

Given any basic open set $\mathscr{N}(f_0)$ and any partial recursive operator Φ, it may or may not be the case that $< \Phi f >$ has the same value for all $f \in \mathscr{N}(f_0)$ for which it is defined. If this happens for some $\mathscr{N}(f_0)$ we call $< \Phi f >$ a *singular* function; in symbols

$$g \text{ singular} \leftrightarrow (\exists f_0)(\exists \Phi)(\Phi \text{ partial recursive and } (\forall f \in \mathscr{N}(f_0))(< \Phi f > \text{ defined} \rightarrow < \Phi f > = g)) .$$

A function which is not singular we call *regular*. Clearly there are c regular and at most $\aleph_0$ singular functions.[21]

We wish to exhibit an uncountable collection of pairwise incomparable degrees, or, what comes to the same thing, an uncountable family of functions none of which is recursive in any other. We prove this by establishing successively the following propositions.

A. If f is regular and Φ partial recursive, then $\Phi^{-1}(f)$ is nowhere dense.

B. If f is regular, then the family of all functions of degree $\geqq$ the degree of f is of first category.

C. If f is regular, then the family of all functions of degree comparable with the degree of f is of first category.

D. If $\mathscr{F}$ is a (finite or) countable family of regular functions, then the family of all functions which are either singular or of degree comparable with that of some function belonging to $\mathscr{F}$ is of first category.

E. If $\mathscr{F}$ is a (finite or) countable family of regular functions, there exists a regular function of degree incomparable with the degree of every function in $\mathscr{F}$.

F. There exists an uncountable family of pairwise incomparable degrees.

Clearly $A \rightarrow B \rightarrow C \rightarrow D \rightarrow E \rightarrow F$, so we have only to prove A. Let then f be regular, $\mathscr{N}$ a basic open set, Φ a partial recursive operator. We seek a subneighborhood $\mathscr{N}_0$ of $\mathscr{N}$ such that for all $g \in \mathscr{N}_0$, Φg is undefined or $\neq f$. If $< \Phi g >$ is undefined for all $g \in \mathscr{N}$, take $\mathscr{N}_0 = \mathscr{N}$. If on the other hand $< \Phi g >$ is defined for some $g \in \mathscr{N}$, then there exists (since f is regular) such a g for which $< \Phi g > \neq f$. Let $\mathscr{F}$ be

[21] The singular functions are precisely the functions f for which the relation $f(x) = y$ is hyperarithmetic (see Davis [**2**], p. 192 for the definition of hyperarithmetic). The proof is essentially contained in [**8**].

the domain of Φ. Then $\{g \mid <\Phi g> \neq f\} = \mathscr{N}_1 \cap \mathscr{F}$ for some open $\mathscr{N}_1$. Consequently we can take $\mathscr{N}_0 = \mathscr{N} \cap \mathscr{N}_1$ and $\Phi^{-1}(f)$ is nowhere dense, q.e.d.

It must be stressed that some existence thorems in the literature of degrees apparently cannot be reduced to category arguments, at least not in the topology which we used.[22] Also Shoenfield's proof of the existence of $\aleph_1$ pairwise incomparable degrees is essentially different from the above, and yields the further information that *given any countable family of* non-recursive *functions* (i.e., not of the lowest degree, not effectively calculable) *there is a function of degree incomparable with all of them.* We only obtain the statement (E above) reading 'regular' for 'non-recursive'; and this is weaker as we have seen. If possible we seek a category argument which will yield this stronger result. However we cannot do this without more structure on $\mathscr{T}$. For we can exhibit a countable family of continuous operators

$$\Phi : \mathscr{F} \to \mathscr{T}$$

with the following four properties:

I. They are closed under composition whenever possible.

II. They contain the identity.

III. The domain of each is a $G\delta$.

IV. There exists a minimum in the induced ordering $f \leqq g$

such that it is *false* that given any countable family of functions none of which is minimal in the sense of IV, then there is a functions incomparable with them all.

The following additional assumption however, which is true for partial recursive operators, yields enough additional structure for us to obtain Shoenfield's result by essentially his method.

V. If the domain of Φ is dense on an open set, its intersection with that set contains a minimal (i.e., recursive) point.

It is obviously enough (in view of the earlier part of this paper) to prove that $\Phi^{-1}(f)$ is nowhere dense for each *non-recursive* f. For this, consider such an f and let $\mathscr{N}$ be a basic open set and Φ a partial recursive operator. We seek again a subneighborhood $\mathscr{N}_0$ of $\mathscr{N}$ disjoint from $\Phi^{-1}(f)$. If the domain $\mathscr{F}$ of Φ is not dense on $\mathscr{N}$, this is trivial;

[22] Spector's proof in [**19**] of the existence of minimal non-recursive degrees has been made into a category argument by Lacombe (unpublished); but the topology used is highly artificial.

so assume it is dense. By V, its intersection with $\mathscr{N}$ contains a recursive point g. If $<\Phi g> = f$, f would be recursive, contradicting the hypothesis. Hence $<\Phi g> \neq f$ and as alove we can take $\mathscr{N}_0 = \mathscr{N}_1 \cap \mathscr{N}$ where $\mathscr{N}_1$ is an open set such that $\{g \mid <\Phi g> \neq f\} = \mathscr{N}_1 \cap \mathscr{F}$, q.e.d.

The proof of V however seems to require essential use of (non-topological) properties of recursive *functions* as distinguished from operators, specifically their closure under a certain iterative procedure. We conclude that Shoenfield's result (and a fortiori the results of Sacks and Nerode mentioned in footnote 20) probably do not, like some of the other theorems on degrees mentioned in this note, rest solely on elementary set-theoretic considerations. However, the distinction between those which do and those which do not require more advanced and specialized means (i.e., between those which are truly 'recursive' and those which are merely set-theoretic) seems worth making, if only because it throws some light on aspects of the methodology of the whole domain which the present treatment in the literature leaves almost completely in the dark.

REFERENCES

1. John Addison, *Separation principles in the hierarchies of classical and effective descriptive set-theory*, Fund. Math., **46** (1959), 123-135.
2. Martin Davis, *Computability and unsolvability*, New York, McGraw, Hill, 1958, xxv + 210 pp.
3. Jacob Dekker and John Myhill, *Recursion theory*, to be published by North Holland.
4. Richard Friedberg, *Two recursively enumerable sets of incomparable degrees of unsolvability*, Proc. Nat. Acad. Sci. U.S.A., **43** (1957), 236-238.
5. Stephen Kleene, *Introduction to metamathematics*, New York, Van Nostrand, 1952, x + 550 pp.
6. Stephen Kleene, *Recursive functions and intuitionistic mathematics*, Proc. Int. Cong. Math. (Cambridge, Mass., U.S.A.) **1** (1952), 679-685.
7. Stephen Kleene and Emil Post, *The upper semi-lattice of degrees of recursive unsolvability*, Ann. of Math., ser. 2, **59** (1954), 379-407.
8. A. V. Kuznecov and B. A. Trahtenbrot *Investigation of partial recursive oparators by means of the otheoryf Baire space*, Proc. Acad. Sci. USSR, **105** (1955), 897-900.
9. Daniel Lacombe, *Sur le semi-réseau constituté par les degrés d'indecidabilité récursive*, C. R., **239** (1954), 1108-1109.
10. ———, *Quelques procédés de définition en topologie récursive*, in *Constructivity in Mathematics*, Amsterdam, North Holland, (1958), 129-158.
11. A. A. Mučnik, *Negative anser to the problem of reducibility of the theory of algorithms*, Proc. Acad. Sci. USSR, **18** (1956), 194-197.
12. John Myhill, *Note on degrees of partial functions*, Proc. Amer. Math. Soc., **12** (1961), 519-521.
13. John Myhill and John Shepherdson, *Effective operations on partial recursive functions*, Zeit. für Math. Logik u. Grund. der Math., **1** (1955), 310-317.
14. Joseph Shoenfield *An uncountable set of incomparable degrees*, Proc. Amer. Math. Soc., **11** (1960), 61-62.
15. ———, *On degrees of unsolvability*, Ann. of Math., ser. 2, **69** (1959), 644-653.

16. Waclaw Sierpiński, *Hypothese du continu*, New York, Chelsea, 1956, xvii + 274 pp.
17. Raymond Smullyan, *Theory of formal systems*, Annals of Mathematics Studies, No. 47.
18. Clifford Spector, *Measure-theoretic construction of incomparable hyperdegrees*, J. Symb. Log., **23** (1958), 280-288.
19. Clifford Spector, *On degrees of recursive unsolvability*, Ann. of Math., **64** (1956), 580-592.
20. B. A. Trahtenbrot, *Matrix representation of recursive operators*, Proc. Acad. Sci. USSR, **101** (1955), 417-420.

ON EXTREMAL PROPERTIES FOR ANNULAR RADIAL AND CIRCULAR SLIT MAPPINGS OF BORDERED RIEMANN SURFACES

PAUL A. NICKEL

Introduction. There exist functions which map a planar Riemann surface W of arbitrary conectivity conformally onto plane slit regions. Functionals I, extremized in the class of all conformal mappings of W by only one slit mapping, are known. Such functionals can be represented as limits of functionals I_n, where each I_n is itself extremized by a horizontal or vertical-slit mapping with domain of finite connectivity.

A planar bordered Riemann surface of finite connectivity can be mapped conformally onto a radial or circular-slit annulus with inner and outer boundaries corresponding to any two contours of the surface. In this investigation, extremal properties of such mappings are obtained and extended to surfaces of infinite connectivity. The geometric nature of the extended mappings, called principal analytic functions, is then deduced from the extended extremal properties. In addition, certain combinations of principal analytic functions are investigated from both extremal and geometric points of view.

First, we consider a planar bordered oriented Riemann surface $\bar{W}$, of infinite connectivity. It is assumed that $\bar{W}$ has two compact border components, δ and γ, such that no point of $\delta \cup \gamma$ is a limit point of points of any other boundary components. Such contours are called isolated. $\bar{W}$ is "approximated" by a sequence of compact bordered Riemann surfaces $\{W_n\}$, where each W_n is of finite connectivity. On W_n, annular radial and circular-slit mappings F_{0n} and F_{1n} are constructed. Among all normalized conformal annular mappings F of W_n, F_{0n} maximizes

$$2\pi \log(r(F)) + \mu_n(F)$$

and F_{1n} minimizes

$$2\pi \log(r(F)) - \mu_n(F) .$$

Here, $r(F)$ is the quotient r_γ/r_δ, where r_γ and r_δ represent the radii of the positively oriented $F(\gamma)$ and the negatively oriented $F(\delta)$ respectively, and $\mu_n(F)$ is the complementary area of $\log(F(W_n))$.

It is then shown by the reduction theorem (Sario[4]) that these extremal properties hold in the limit for the limit functions F_0 and F_1.

Received August 22, 1960. The results of this paper are part of the contents of the author's Ph. D. thesis, done under the direction of Professor Leo Sario, to whom I wish to express my most sincere gratitude.

Furthermore, the extremal properties of F_0 and F_1 imply that the former is a radial slit mapping of $\bar{W}$ and that the latter is a circular slit mapping. By establishing a deviation formula, it is seen that the functions F_0 and F_1 are, up to a rotation, the only normalized conformal annular maps of $\bar{W}$ extremizing the limit functionals. As another application of the reduction theorem, we find that the univalent function $\sqrt{F_0 \cdot F_1}$ maximizes $\mu(F)$, the complementary logarithmic area, among all conformal annular mappings of $\bar{W}$.

Next we pose the question: When does $\bar{W}$ have distinct radial and circular-slit mappings. The answer is given in terms of AD-removability, at least when $\bar{W}$ is a plane region bounded by an outer contour γ and an inner contour δ. A point set E of the extended plane is called AD-removable when the only analytic functions with finite Dirichlet integral, defined on the complement of E, are the constant functions. In particular, we find that the principal analytic functions are, up to a rotation, identical, if and only if the plane region bounded by γ and δ minus $\bar{W}$ is AD-removable.

1. We consider $\bar{W}$ an open planar bordered Riemann surface with two compact non-point border components, δ and γ. In order to describe the remaining part of the boundary of $\bar{W}$, we recall that such a surface can be embedded in a Riemann sphere S^2. With respect to this embedding, we assume that $\bar{W}$ and its boundary components satisfy the following conditions:

(1) no point of $\delta \cup \gamma$ is a limit point of points of any other boundary components, and (2) $\bar{W} - (\delta \cup \gamma)$ is open in S^2. Operations in $\bar{W}$ such as interior, boundary, etc., are referred to S^2.

It is possible to exhaust an open Riemann surface by a countable collection of compact approximating regions $\{W_n\}$. In fact, $\bar{W}$ can be countably exhausted in the following modified sense:

1. $\delta \cup \gamma \subset W_n$.
2. $W_n \subset \operatorname{Int} W_{n+1}$.
3. The boundary of W_n consists of a finite number of disjoint analytic Jordan curves.
4. Each component of $\bar{W} - W_n$ is relatively non-compact.
5. $\bar{W} = \cup W_n$.

There is no loss in generality in assuming that each W_n contains a $\zeta \in \bar{W}$, where ζ is arbitrary but fixed in advance.

Evidently δ and γ are two border components of W_n. The remaining border components will be denoted $\beta_1(W_n)$, $\beta_2(W_n)$, $\cdots$, $\beta_{k(n)}(W_n)$. When only one approximating subregion is under consideration, the notation for these remaining border components will be shortened to $\beta_1, \beta_2, \cdots$, $\beta_{k(n)}$. For convenience we define $\bar{\beta}_n$ as

$$\bar{\beta}_n = \cup \beta_i(W_n) . \qquad i = 1, 2, \cdots k(n)$$

I. Extremal Properties of Harmonic Functions Defined on Approximating Regions.

2. We consider, in this and the following section, certain classes of harmonic and analytic functions defined on an approximating region W_n.

DEFINITION. $H_n(h+k)$ is the set of functions p, harmonic on Int $W_n \cup \delta \cup \gamma$ and satisfying

(1) $p(z) = c_2(p) = \text{const.}$ for $z \in \gamma$ with $\int_\gamma dp^* = 2\pi(h+k)$,

(2) $p(\zeta) = 0$,

(3) $p(z) = c_1(p)$ for $z \in \delta$ with $\int_\delta dp^* = -2\pi(h+k)$, and

(4) $\int_{\beta_i} dp^* = 0$ for $i = 1, 2, \cdots, k(n)$.

h and k are real numbers. When the function p is defined only on Int $W_n \cup \delta \cup \gamma$, then the integrals $\int_{\beta_i(W_n)} dp^*$ and $\int_{\beta_i(W_n)} pdp^*$ are understood as $\lim_{k\to\infty}\int_{\beta_i(W'_k)} dp^*$ and $\lim_{k\to\infty}\int_{\beta_i(W'_k)} pdp^*$. Here $\{W'_k\}$ is an exhaustion of the surface Int W_n and each $\beta_i(W'_k)$ is homologous (in W_n) to $\beta_i(W_n)$. An application of Green's formula shows that these limits are independent of the exhaustion $\{W'_k\}$. The class $H_n(1)$ will be denoted H_n.

Principal harmonic functions p_{0n} and p_{1n}, belonging to H_n are obtained as harmonic extensions of functions constructed by use of linear operators on Riemann surfaces (Sario [2]). In fact on each β_i, $i = 1, 2, \cdots, k(n)$, $\partial p_{0n}/\partial n = 0$ and $p_{1n} = \text{const.}$ Hence for arbitrary h and k, the function $p_{hkn} = hp_{0n} + kp_{1n}$ belongs to the class $H_n(h+k)$, which is then not empty.

3. THEOREM 1. *P_{hkn} minimizes the functional* $\int_{\bar{\beta}_n} pdp^* - 2\pi(h-k)c(p)$ *among all* $p \in H_n(h+k)$, *where* $c(p) = c_2(p) - c_1(p)$.

The value of the minimum is $-2\pi[h^2c(p_{0n}) - k^2c(p_{1n})]$.

The deviation of this functional from its minimum is $D_{W_n}(p - p_{hkn})$, *and the minimizing function is unique.*

Proof. Let B be the entire border of W_n. Then by Green's formula, we have

$$D_{W_n}(p - p_{hkn}) = \int_B (p - p_{hkn})d(p - p_{hkn})^* .$$

Since p and $p_{hkn} \in H_n(h+k)$, we conclude at once that $\int_{\delta+\gamma}(p - p_{hkn}) d(p - p_{hkn})^* = 0$. Green's formula becomes

$$D_{W_n}(p - p_{hkn}) = \int_{\bar{\beta}_n} pdp^* + \int_{\bar{\beta}_n} p_{hkn}dp^*_{hkn} - \int_{\bar{\beta}_n} p_{hkn}dp^* + pdp^*_{hkn} .$$

We now expand the last term and find that

$$\int_{\bar{\beta}_n} p_{hkn}dp^* + pdp^*_{hkn} = h\int_{\bar{\beta}_n} p_{0n}dp^* + pdp^*_{0n} + k\int_{\bar{\beta}_n} p_{1n}dp^* + pdp^*_{1n} .$$

But on $\bar{\beta}_n$, p_{0n} has vanishing normal derivative, and p_{1n} is constant. This means that $\int_{\bar{\beta}_n} pdp^*_{0n} = \int_{\bar{\beta}_n} p_{1n}dp^* = 0$ when $p \in H_n(h+k)$. Thus we can infer from Green's formula that

$$\int_{\bar{\beta}_n} p_{hkn}dp^* + pdp^*_{hkn} = h\int_{\delta+\gamma} p_{0n}dp^* - pdp^*_{0n} + k\int_{\delta+\gamma} p_{1n}dp^* - pdp^*_{1n} .$$

A direct application of the conditions (1), (3), (4) of $H_n(h+k)$ now yields the formula

$$\begin{aligned}\int_{\bar{\beta}_n} p_{hkn}dp^* + pdp^*_{hkn} &= 2\pi(h-k)(c_2(p) - c_1(p)) \\ &\quad -2\pi h(h+k)(c_2(p_{0n}) - c_1(p_{0n})) \\ &\quad +2\pi k(h+k)(c_2(p_{1n}) - c_1(p_{1n})) .\end{aligned}$$

We obtain in a similar fashion

$$\begin{aligned}\int_{\bar{\beta}_n} p_{hkn}dp^*_{hkn} &= hk\int_{\bar{\beta}_n} p_{0n}dp^*_{1n} - p_{1n}dp^*_{0n} \\ &= -hk\int_{\delta+\gamma} p_{0n}dp^*_{1n} - p_{1n}dp^*_{0n} \\ &= -2\pi hk[c_2(p_{0n}) - c_1(p_{0n}) - (c_2(p_{1n}) - c_1(p_{1n}))] .\end{aligned}$$

Collecting contributions, we find

$$\begin{aligned}D_{W_n}(p - p_{hkn}) &- 2\pi[h^2(c_2(p_{0n}) - c_1(p_{0n})) - k^2(c_2(p_{1n}) - c_1(p_{1n}))] \\ &= \int_{\bar{\beta}_n} pdp^* - 2\pi(h-k)(c_2(p) - c_1(p)) .\end{aligned}$$

Since the Dirichlet integral is nonnegative, we have that p_{hkn} minimizes the given functional. Clearly, for any $p \in H_n(h+k)$ the deviation of the functional from its minimum is $D_{W_n}(p - p_{hkn})$.

We consider now the uniqueness of the minimizing function. For another minimizing function p', we would have a deviation of the functional from the minimum equal to $D_{W_n}(p' - p_{hkn})$. But p' also minimizes, so $D_{W_n}(p' - p_{hkn}) = 0$. Since $p_{hkn}(\zeta) = p'(\zeta) = 0$, we see that $p_{hkn} = p'$. This completes the proof of Theorem 1.

4. Our interest in Theorem 1 will be with the following special cases which we state as corollaries.

COROLLARY 1. *p_{0n} maximizes the functional $2\pi c(p) - \int_{\bar{\beta}_n} pdp^*$ among all $p \in H_n$.*

COROLLARY 2. *p_{1n} minimizes the functional $2\pi c(p) + \int_{\bar{\beta}_n} pdp^*$ among all $p \in H_n$.*

COROLLARY 3. *$\frac{1}{2}(p_{0n} + p_{1n})$ minimizes the functional $\int_{\bar{\beta}_n} pdp^*$ among all $p \in H_n$.*

COROLLARY 4. *$p_{0n} - p_{1n}$ maximizes the functional $4\pi c(p) - D_{W_n}(p)$ among all $p \in H_n(0)$.*

Each extremizing function is unique.

Corollaries 1, 2, and 3 follow immediately from Theorem 1 for $h + k = 1$. As for Corollary 4, clearly $p_{0n} - p_{1n} \in H_n(0)$. Now for any $p \in H_n(0)$, Green's formula reads $D_{W_n}(p) = \int_{\gamma+\delta+\bar{\beta}_n} pdp^* = \int_{\bar{\beta}_n} pdp^*$, and Corollary 4 follows.

II. Geometric Properties of Analytic Functions Defined on Approximating Regions.

5. DEFINITION. A_n is the class of functions F analytic on Int $W_n \cup \delta \cup \gamma$ such that

(1) $F(\gamma)$ is a circle traced once in the positive direction,

(2) $|F(\zeta)| = 1$,

(3) $F(\delta)$ is a circle traced once in the negative direction,

(4) F is univalent on Int $W_n \cup \delta \cup \gamma$.

In this definition, $F(\gamma)$ and $F(\delta)$ are understood as oriented images of oriented border cycles and the radii of these images are denoted $r_\gamma(F)$ and $r_\delta(F)$.

Some useful relations between the classes A_n and H_n are expressed in the following theorem.

6. THEOREM 2. (a) *For any $F \in A_n$, $\log|F|$ is of class H_n.*

(b) *The following analytic functions are of class A_n:*

$$(1) \quad F_{0n} = \exp(p_{0n} + ip_{0n}^*) , \qquad (2) \quad F_{1n} = \exp(p_{1n} + ip_{1n}^*) .$$

The functions F_{in} are referred to as *principal analytic functions.*

Proof of (a). Evidently $2\pi = \int_\gamma d(argF(z)) = \int_\gamma d(\log|F(z)|)^*$ and

Condition 1 of H_n is verified. Condition 3 is checked just as easily and (2) is apparent. As for (4), let β_i be any component of the border of W_n other than δ or γ. Suppose that $\beta_i' \sim \beta_i$ and that $\int_{\beta_i'} d(\log|F|)^* = 2\pi k$, where k is an integer. There exists a path from δ to γ which does not meet β_i'. But if $k \neq 0$, then every path from $F(\delta)$ to $F(\gamma)$ meets $F(\beta_i')$. But F is univalent, so $k = 0$.

Proof of (*b*). We consider first the function F_{1n} and omit the analogous proof for F_{0n}. First, it is evident that $2\pi = \int_\gamma dp_{1n}^* = \int_\gamma d(\arg F_{1n})$ and $r_\gamma(F) = \exp c_2(p_{1n}) = \text{const}$. Certainly $F_{1n}(\gamma)$ is a circle traced once in the positive direction, and (1) of No. 5 is satisfied. Condition 3 is verified in a similar manner and (2) is trivial.

To verify the Condition 4, we consider the extended version of the argument principle, and reason in a manner analogous to Ahlfors [**1**], p. 203.

7. Definition. The multiple-valued functions P_{in} are defined as $P_{in} = p_{in} + ip_{in}^*$. However $P_{0n} - P_{1n}$ is single-valued, and the principal analytic functions are expressible as $F_{in} = \exp P_{in}$, $i = 0, 1$. We also fix the following terminology: $r(F)$ denotes the ratio $r_\gamma(F)/r_\delta(F)$ and $\mu_n(F)$ denotes the complementary logarithmic area $-\int_{\bar{\beta}_n} \log|F(z)|d(\arg F(z))$, a nonnegative quantity when $F \in A_n$.

Theorem 3. *F_{0n} maximizes $2\pi \log r(F) + \mu_n(F)$ among all $F \in A_n$.*
F_{1n} minimizes $2\pi \log r(F) - \mu_n(F)$ among all $F \in A_n$.
$P_n = \sqrt{F_{0n} \cdot F_{1n}}$ maximizes $\mu_n(F)$ among all $F \in A_n$.
F_{0n}/F_{1n} maximizes $4\pi \log r(F) - D_{W_n}(\log|F|)$ among all quotients of functions in A_n.
$P_{0n} - P_{1n}$ maximizes $4\pi[Re(F(z_2) - F(z_1))] - D_{W_n}(F)$ among all analytic functions on W_n the real part of which is constant on δ, constant on γ, and 0 at ζ. Here z_2 and z_1 are on γ and δ respectively.

Proof. We have $\log|F_{0n}(z)| = p_{0n}(z)$, so it follows from Corollary 1 of Theorem 1 that $\log|F_{0n}|$ maximizes the functional $2\pi c(p) - \int_{\bar{\beta}_n} p dp^*$ among all $p \in H_n$. But according to Theorem 2, when $F \in A_n$, the $\log|F(z)| \in H_n$. Hence F_{0n} maximizes the functional $2\pi \log r(F) + \mu_n(F)$ among all $F \in A_n$. The proof of the second part of this theorem is analogous, and so is the proof of the third part when it is shown that $P_n = \sqrt{F_{0n} \cdot F_{1n}}$ is of class A_n, a fact that is proved in the appendix.

It is easily seen that $\log|F_{0n}/F_{1n}| = p_{0n} - p_{1n}$, hence according to Corollary 4 of Theorem 1, $\log|F_{0n}/F_{1n}|$ maximizes $4\pi c(p) - D_{W_n}(p)$

among all $p \in H_n(0)$. If $F = G/H$, where G and $H \in A_n$, then it follows from Theorem 2 that $\log|G|$ and $\log|H| \in H_n$, and we have $\int_\gamma d(\log|F|)^* = \int_\gamma d(\arg F) = \int_\gamma d(\arg G/H) = 0$. Other similar calculations show that $\log|F(z)| \in H_n(0)$. Thus F_{0n}/F_{1n} maximizes $4\pi \log r(F) - D_{W_n}(\log|F|)$ among all quotients of functions in A_n.

The extremal property of $P_{0n} - P_{1n}$ follows from Corollary 4 as well when it is observed that $Re(P_{0n} - P_{1n}) = p_{0n} - p_{1n}$, and that $ReF \in H_n(0)$ when F is analytic on W_n.

The following corollary of Theorem 3 will be useful when we are considering geometric properties of conformal maps of $\bar{W}$.

COROLLARY. *The functional $r(F)$ is maximized, uniquely up to a rotation, by F_{0n} and minimized, uniquely up to a rotation, by F_{1n} among all $F \in A_n$.*

Proof. It follows from the definition of F_{0n} given in No. 6. that $d(\arg F_{0n}) = dp_{0n}^*$, which is 0 on $\bar{\beta}_n$. Since $\mu_n(F) \geqq 0$, we have $2\pi \log r(F) \leqq 2\pi \log r(F) + \mu_n(F) \leqq 2\pi \log r(F_{0n}) + \mu_n(F_{0n}) = 2\pi \log r(F_{0n})$, that is, $r(F)$ is maximized by F_{0n}.

Analogous reasoning shows that F_{1n} minimizes $r(F)$ among all $F \in A_n$.

In order to establish the uniqueness, we let $r(F) = r(F_{0n})$ for some $F \in A_n$. Then an application of Theorem 3 yields $0 < \mu_n(F) \leqq \mu_n(F_{0n}) \leqq 0$, which means that F also maximizes the functional $2\pi \log r(F) + \mu_n(F)$ among $F \in A_n$. But an application of the deviation formula of Theorem 1 shows that $D_{W_n}(\log|F/F_{0n}|) = 0$, from which it follows that $F = cF_{0n}$ with $|c| = 1$.

III. Extremal Properties of Principal Harmonic Functions.

8. We propose in the present section, to develop for domains of infinite connectivity, extremal theorems which will generalize the results of § 1 for finite connectivity. An essential role is played by the

Reduction Theorem (Sario [4]).

Assume that Z and Z_n are classes of functions with domains W, an arbitrary open Riemann surface, and W_n, an exhausting subregion of W, respectively. In addition, suppose that real-valued functionals m and m_n, defined on Z and Z_n, satisfy the following conditions.

(R1) If $W_m \subset W_n$ and if $f \in Z_n$, then $f|_{W_m} \in Z_m$.

Here W_n may be replaced by W, and Z_n by Z.

(R2) If $\{f_k\}$ is a sequence the elements of which belong to Z_n, and if $\{f_k\}$ converges uniformly to $f \in Z_n$, then $m_n(f_k)$ converges to $m_n(f)$.

(R3) $m(f) = \lim\limits_{n\to\infty} m_n(f)$, for any $f \in Z$.

(R4) There exists a function $f_n \in Z_n$ such that f_n minimizes the functional m_n among all $f \in Z_n$.

(R5) For $k < h$, and $f \in Z_h$, $m_k(f) \leqq m_h(f)$.

(R6) The family $\{f_n; f_n$ minimizes m_n among $f \in Z_n\}$ is a normal family, and the limit functions belong to Z.

Then any limit function $f = \lim_{n\to\infty} f_n$ *minimizes* m *among all* $f \in Z$, *and value of minimum is* $m(f) = \lim_{n\to\infty} m_n(f_n)$.

The proof of the reduction theorem is established by selecting an exhaustion of W, and can be carried out for a bordered surface $\bar{W}$ as well, as soon as an exhaustion is known to exist.

9. Let $\bar{W}$ be an open planar bordered Riemann surface, $\{W_n\}$ an exhausting set, δ and γ separated boundary components, all as described is no. 1.

LEMMA 1. *The families* $\{p_{0n}\}$ *and* $\{p_{1n}\}$ *are normal.*

Proof. If $\{F_{0n}\}$ ($\{F_{1n}\}$) is a normal family, then so is $\{p_{0n}\}$ ($\{p_{1n}\}$). Hence it suffices to show that for every compact set S, there exist a constant M and and integer N such that $|F_{0n}(z)| < M$ ($|F_{1n}(z)| < M$) for all $n > N$ and all $z \in S$. Let S be any compact subset of $\bar{W}$ and choose n sufficiently large so that $S \subset W_n$. For any $z \in S$ and $W_p \subset W_n$, since $F_{0p}(\gamma)$ is the *outer* contour of an image annulus we have $2\pi \log|F_{0p}(z)/r_\delta(F_{0p})| \leqq 2\pi \log(r(F_{0p})) + \mu_n(F_{0p})$. But according to Theorem 3, the right hand side is bounded by $2\pi \log(r(F_{0n}))$. We now recall that $|F_{0p}(\zeta)| = 1$, that is $r_\delta(F_{0p}) < 1$. Hence $|F_{0p}(z)|$ is bounded for all $z \in S$ and for all $p \geqq n$, and the family $\{F_{0n}\}$ is normal.

As for $\{F_{1n}\}$, we have

$$2\pi \log|F_{1n}(z)/r_\delta(F_{1n})| \leqq 2\pi \log(r(F_{1n})) \leqq 2\pi \log(r(F_{0n})) .$$

The second inequality follows from the Corollary of Theorem 3. We conclude that $\{F_{1n}\}$ is bounded on any compact set S and is normal. This completes the proof of Lemma 1.

An immediate consequence of Lemma 1 is that the family $\{p_{hkn}\}$ is normal.

10. LEMMA 2. *If* $n < n'$, *then the inequality*

$$\int_{\bar{\beta}_n} p dp^* \leqq \int_{\bar{\beta}_{n'}} p dp^*$$

holds for all $p \in H_{n'}(h+k)$.

Proof. We apply the first form of Green's formula to the region

$W_{n'} - W_n$ and find

$$\int_{\bar{\beta}_n'} pdp^* - \int_{\bar{\beta}_n} pdp^* = D_{W_{n'}-W_n}(p) \geqq 0 .$$

DEFINITION. $H(h+k)$ is the class of functions p, harmonic on $\bar{W}$, satisfying

(1) $p(z) = c_2(p) = \text{const.}$ for $z \in \gamma$ with $\int_\gamma dp^* = 2\pi(h+k)$,

(2) $p(\zeta) = 0$,

(3) $p(z) = c_1(p) = \text{const.}$ for $z \in \delta$ with $\int_\delta dp^* = -2\pi(h+k)$, and

(4) $\int_\sigma dp^* = 0$ where σ is any cycle which is homeomorphic to a circle and which does not separate δ and γ. A cycle σ is said to *separate* δ and γ if every path from δ to γ intersects σ. Let H denote the class $H(1)$.

DEFINITION. For any $p \in H(h+k)$, $\int_{\bar{\beta}} pdp^*$ is understood to be $\lim_{n\to\infty}\int_{\bar{\beta}_n} pdp^*$. The existence of this limit is guaranteed by the monotonicity condition of Lemma 2.

LEMMA 3. *If the sequence* $\{p_n; p_n \in H_n(h+k)\}$ *converges on compact subsets to* p', *then* $p' \in H(h+k)$.

We recall that a sequence $\{f_n\}$ *converges on compact sets* if for every compact set S, there exists an N such that $\{f_n; n \geqq N\}$ converges uniformly on S.

Proof. The convergence $p_n \to p'$ is uniform on compact sets. The conditions (1), (2), and (3) for $H(h+k)$ can therefore be inferred from those of $H_n(h+k)$. Let σ be any cycle which does not separate δ and γ. Then there exists n such that the compact $\sigma \subset W_n$, and we have

$$\sigma \sim b_1\delta + \sum a_i \beta_i'$$

where the β_i' are homologous to components of the border of W_n (Ahlfors and Sario [1]). We embed W_n in the complex plane with γ as an outer boundary, and fill in the "holes" whose boundaries are the β_i''s. Now $\sigma - b_1\delta = \partial A$, and every path from δ to γ meets σ. This is a contradiction, unless $b_1 = 0$.

Using the uniform convergence of $\{p_n\}$ along with Green's theorem, we obtain

$$\int_\sigma dp'^* = \lim_{n\to\infty}\int_\sigma dp_n^* = \lim_{n\to\infty}\int_{\sum a_i\beta_i'} dp_n^* = \lim_{n\to\infty}\sum a_i \int_{\beta_i'} dp_n^* = 0 .$$

DEFINITION. A harmonic function p_{hk} is defined as the limit of any sequence of the normal family $\{p_{hkn}\}$ which converges on compact sets.

THEOREM 4. *p_{hk} minimizes the functional* $\int_{\bar\beta} pdp^* - 2\pi(h-k)c(p)$ *among all* $p \in H(h+k)$.

The minimum value of this functional is $-2\pi(h^2c(p_0) - k^2c(p_1))$.

The deviation of this functional from its minimum value is $D(p - p_{hk})$ *and the minimizing function is unique.*

11. There exists a subsequence $\{p_{hkn'}\}$ of $\{p_{hkn}\}$ which converges to p_{hk} on compact sets and satisfies $\lim_{n'\to\infty} p_{hkn'} = hp_0 + kp_1$ where $p_i = \lim_{n'\to\infty} p_{in'}$, $i = 0, 1$. The uniqueness of Theorem 4 then allows us to assume that $p_{hk} = hp_0 + kp_1$ for all h and k.

Proof. That p_{hk} minimizes and gives the functional the value $-2\pi(h^2c(p_0) - k^2c(p_1))$ will follow from Theorem 1 if we can verify (R1) – (R6) of the reduction theorem. The functionals m_n and m are taken to be $\int_{\beta_n} pdp^* - 2\pi(h-k)c(p)$ and $\int_{\bar\beta} pdp^* - 2\pi(h-k)c(p)$ respectively, while the classes Z_n and Z are $H_n(h+k)$ and $H(h+k)$.

If $p \in H(h+k)$, then $p|_{W_n}$ satisfies the Conditions 1, 2, and 3 for $H_n(h+k)$. Since no $\beta_i(W_n)$ separates δ and γ, $\int_{\beta_i(W_n)} dp|_{W_n}^* = 0$ and (4) is satisfied. Hence $p|_{W_n} \in H_n(h+k)$ and (R1) is verified. The uniform convergence of f_k to f makes (R2) evident, and the functional $\int_{\bar\beta} pdp^* - 2\pi(h-k)c(p)$ is defined as $\lim_{n\to\infty} \int_{\bar\beta_n} pdp^* - 2\pi(h-k)c(p)$, as required by (R3).

Theorem 1 shows that (R4) is satisfied, and Lemma 2 of no. 10 shows the same for (R5). That the family $\{p_{hkn}\}$ as defined in no. 2 is normal, follows from Lemma 1 of no. 9, and that the limiting functions belong to $H(h+k)$ is then a consequence of Lemma 3 of no. 10. Thus by the reduction theorem, the limit function, p_{hk}, minimizes the limit functional among $p \in H(h+k)$ and the minimum value of the limit functional is the limit of minimum values.

12. In order to establish the deviation formula, we first denote the functional of Theorem 1 by ψ_n and consider its value on the function $p_\epsilon = p_{hk} + \epsilon(p - p_{hk})$. Upon expanding, we find

$$\psi_n(p_\epsilon) = \int_{\bar\beta_n} p_{hk} dp_{hk}^* - 2\pi(h-k)c(p_{hk}) + a_1(h)\varepsilon + \varepsilon^2 \int_{\bar\beta_n} (p - p_{hk}) d(p - p_{hk})^* , \tag{2}$$

where for each n, this is a polynomial in ε, and $a_1(n)$ is the coefficient of the ε term. But the last integral is

$$\int_{\bar{\beta}_n}(p - p_{hk})d(p - p_{hk})^* = \int_{\bar{\beta}_n+\delta+\gamma}(p - p_{hk})d(p - p_{hk})^* = D_{W_n}(p - p_{hk}) .$$

The first equality follows from the fact that p and p_{hk} both belong to $H(h + k)$. Therefore, in the sense of limits, we write

$$\int_{\bar{\beta}}(p - p_{hk})d(p - p_{hk})^* = D(p - p_{hk}) ,$$

where D is the integral over the entire bordered surface $\bar{W}$. In a similar fashion, we find

$$(3) \qquad \int_{\bar{\beta}} p_{hk}dp_{hk}^* - 2\pi(h - k)c(p_{hk}) = D(p_{hk}) - 4\pi hc(p_{hk}) .$$

By an earlier part of this theorem, the left hand side of equation (3) is finite. Thus we have that $D(p_{hk}) < \infty$.

We assume that $D(p - p_k)$ is finite. By the triangle inequlity for the Dirichlet integral (Courant [1]), $D(p)$, and consequently $\int_{\bar{\beta}} pdp^*$ are both finite. Now in equation (2), with $\varepsilon = 1$, consider the limit as $n \to \infty$. The limit of every term, except $a_1(n)$, exists and is finite. Hence the same can be said of $\lim_{n\to\infty} a_1(n)$. But $\psi(p_\varepsilon) = \lim_{n\to\infty}\psi_n(p_\varepsilon)$ has, by part (1) of our theorem, a relative minimum for $\varepsilon = 0$. Therefore, $\lim_{n\to\infty} a_1(n) = 0$, and the deviation formula $\psi(p) = \psi(p_{hk}) + D(p - p_{hk})$ results when $\varepsilon = 1$ is substituted into equation (2) after taking limits.

When $D(p - p_{hk}) = \infty$, this formula holds in the sense that $\psi(p) = \infty$ as well. This completes the proof of Theorem 4.

IV. Extremal and Geometric Properties of Principal Analytic Functions.

Extremal properties for harmonic functions defined on a surface of finite connectivity were used in § 2 to establish extremal properties of analytic functions, also defined on a surface of finite connectivity. In the present section, we exploit the extremal properties of harmonic functions, now defined on a surface of infinite connectivity, for the purpose of establishing both extremal and geometric properties of analytic functions.

13. A competing class of analytic functions is defined as follows.

DEFINITION. A is the class of analytic functions on $\bar{W}$ such that (1) $F(\gamma)$ is a circle traced once in the posititive direction, (2) $|F(\zeta)| =$

1, (3) $F(\delta)$ is a circle traced once in the negative direction, and (4) F is univalent on $\bar{W}$.

THEOREM 5. *For any* $F \in A$, $\log|F| \in H$. *Furthermore* $F_i = \exp(p_i + ip_i^*) \in A$, $i = 0, 1$.

No ambiguity will result in referring also to F_0 and F_1 as *principal analytic functions.*

Proof. For any $F \in A$, consider $\log|F|$, which clearly satisfies (1)-(3) of the definition of H in no. 10. Then let $\int_\sigma d(\log|F|)^* \neq 0$ for σ not separating δ and γ. If $\int_\sigma d(\log|F|)^* = 2\pi k$, k an integer, then $F(\sigma)$ separates $F(\delta)$ and $F(\gamma)$. But F is univalent on $\bar{W}$ and we have the contradiction that σ separates δ and γ. This means that $\log|F| \in H$.

Let $F_i = \exp(p_i + ip_i^*)$, $i = 0, 1$. Conditions 1–3 for A are easily verified. An application of the extended argument principle to any exhausting subregion W_n shows that F_i is univalent on $\delta \cup \gamma$, when univalence is established at interior points. For interior points of $\bar{W}$, F_i can be represented as $\exp(p_i + ip_i^*) = \lim_{n\to\infty}\exp(p_{in} + ip_{in}^*) = \lim_{n\to\infty} F_{in}$. So each F_i is univalent by Theorem 2 and the well-known Hurwitz theorem.

14. The following five theorems are concerned with analytic functions constructed from the harmonic functions p_0 and p_1,w hich are uniquely defined by Theorem 4.

DEFINITION. F is an *annular radial* (*circular*) *slit* mapping of $\bar{W}$ provided that $F(\bar{W})$ is an annulus minus a point set each component of which is a radial (circular) slit or point. Let $\{w; r_\delta(F_i) \leqq |w| \leqq r_\gamma(F_i)\} - F_i(\bar{W})$ be denoted by S_i, $i = 0, 1$.

DEFINITION. For a surface of infinite connectivity, the *complementary logarithmic area* $\mu(F)$ is defined as $\lim_{n\to\infty}\mu_n(F)$ for any $F \in A$. That this limit is defined independently of an exhaustion follows from Theorem 5 and Lemma 2.

THEOREM 6. $F_0 = \exp(p_0 + ip_0^*)$ *maximizes* $2\pi\log(r(F)) + \mu(F)$ *among all* $F \in A$.

The value of the maximum is $2\pi\log(r(F_0))$.

The deviation from the maximum is $D(\log|F/F_0|)$, *and the maximizing function is unique up to a rotation.*

The 2-dimensional Lebesgue measure of the point set S_0 *is* 0.

F_0 *is an annular radial-slit mapping.*

Proof. We apply Theorem 4 with $h = 1, k = 0$ and obtain that $\log|F_0|$ minimizes $\int_{\bar{\beta}} pdp^* - 2\pi c(p)$ among all $p \in H$. According to Theorem 5, we may use Theorem 4 on logarithms of functions in A as well, that is, F_0 maximizes the functional $2\pi \log(r(F)) + \mu(F)$ among all $F \in A$, the maximum value of this functional is $2\pi \log(r(F_0))$, and the deviation from the maximum is $D(\log|F/F_0|)$.

As for the 2-dimensional Lebesgue measure of S_0, consider the annulus

$$\{w; r_\delta(F_0) \leqq |w| \leqq r_\gamma(F_0)\}$$

and set $t = \log w$. The transformation mapping w into $\log w$ is denoted L, and the image of $\{w; r_\delta(F_0) \leqq |w| \leqq r_\gamma(F_0)\}$ under L is called R. Now it is easily seen that

$$LS_0 = \bigcap_{i=1}^{\infty} [C_R(L(F_0(W_n)))] ,$$

where C_R is understood to mean complement with respect to R. $L(F_0(W_n))$ is compact and closed in R, and this means that $C_R[L(F_0(W_n))]$ is open and measurable. Hence, LS_0, a countable intersection of measurable sets, is measurable. Its measure M is then given by

$$M(L(S_0)) = \lim_{n\to\infty} \mu_n(F_0) ,$$

where $\mu_n(F_0)$ is defined in no. 7. But according to an earlier part of this theorem, the term on the right is 0. When we observe that L, defined on the cut annulus, preserves sets of measure zero, we conclude that the 2-dimensional Lebesgue measure of S_0 is zero.

Suppose that the complement, with respect to $\{w; r_\delta(F_0) \leqq |w| \leqq r_\gamma(F_0)\}$, of $F_0(\bar{W})$ is a point set, the components of which are not all radial slits or points. The full annulus

$$\{w; r_\delta(F_0) \leqq |w| \leqq r_\gamma(F_0)\}$$

minus such a component, denoted η, is called W_0. We embed W_0 in the Riemann sphere S^2 and consider the simply connected point set $S^2 - \eta$, which can be mapped conformally onto the complement of a unit disc. Let E be this conformal mapping, and denote by γ'' and δ'' the sets $E(\delta_0)$ and $E(\gamma_0)$, where $\delta_0 = F_0(\delta)$ and $\gamma_0 = F_0(\gamma)$. Now $E(W_0)$ is of finite connectivity, so we can apply Theorem 2 to construct a radial-slit mapping φ of $E(W_0)$ onto an annulus, minus one radial slit, with inner boundary $\varphi(\delta'')$ and outer boundary $\varphi(\gamma'')$. φ is normalized by $|\varphi\circ E\circ F_0(\zeta)| = 1$, and belongs to A_n for $E(W_0)$. We then apply the corollary of Theorem 3 to φ and find that $2\pi\log(r(\varphi)) > 2\pi\log(r(E^{-1})) = 2\pi\log(r(F_0))$. Then the map $\varphi\circ E\circ F_0$, where E and φ are properly restricted, belongs to A. But $2\pi\log(r(\varphi\circ E\circ F_0)) = 2\pi \log(r(\varphi)) > 2\pi\log(r(F_0))$. This is a contradiction, for according to an earlier part of this theorem F_0, up to a rotation, uniquely maximizes the functional $2\pi \log (r(F_0))$ in A. This completes the proof of Theorem 6.

COROLLARY. *The principal analytic function F_0 maximizes the functional $r(F)$ among all $F \in A$.*

Proof. The maximum value of the functional in Theorem 6 is $2\pi \log(r(F_0))$, that is $\mu(F_0) = 0$. The proof is complete when we observe that $\mu(F)$ is nonnegative for all $F \in A$.

THEOREM 7. *$F_1 = \exp(p_1 + ip_1^*)$ minimizes $2\pi \log(r(F)) - \mu(F)$ among all $F \in A$.*

The value of the minimum is $2\pi \log (r(F_1))$.

The deviation from the minimum is $D(\log|F/F_1|)$, and the minimizing function is unique up to a rotation.

The 2-dimensional Lebesgue measure of the point set S_1 is zero.

F_1 is an annular circular-slit mapping.

The proof is analogous to that of Theorem 6 and uses $h = 0$, $k = 1$.

COROLLARY. *The principal analytic function F_1 minimizes the functional $r(F)$ among all $F \in A$.*

THEOREM 8. *$P = \sqrt{F_0 \cdot F_1}$ maximizes $\mu(F)$ among all $F \in A$.*

The value of the maximum is $\mu(P)$.

The deviation from the maximum is $D(\log|F/P|)$, and the maximizing function is unique up to a rotation.

The proof uses $h = 1/2$, $k = 1/2$.

THEOREM 9. *$Q = F_0/F_1$ maximizes $4\pi \log (r(F)) - D(\log|F|)$ among all quotients of functions in A.*

The value of the maximum is $2\pi \log (r(Q))$.

The deviation from the maximum is $D(\log|F/Q|)$.

Proof. When the condition $h = 1$, $k = -1$ is substituted into Theorem 4, it is easily seen that the technique of Theorem 3 will establish Theorem 9.

Consider the multiple-valued functions $P_0 = p_0 + ip_0^*$ and $P_1 = p_1 + ip_1^*$. The difference of these functions has zero flux around any cycle of $\bar{W}$ and is single-valued.

THEOREM 10. *$P_0 - P_1$ maximizes*

$$4\pi[Re(F(z_2) - F(z_1))] - D(F)$$

among all analytic functions on $\bar{W}$ the real part of which is constant on δ, constant on γ, and 0 at ζ. Here z_1, and z_2 are on γ and δ respectively.

The value of the maximum is $-2\pi Re[(P_0 - P_1)(z_2) - (P_0 - P_1)(z_1)]$.
The deviation from the maximum is $D(F - (P_0 - P_1))$.

The proof again applies Theorem 4, with $h = 1$ and $k = -1$, as well as the observation that $Re(P_0 - P_1) = p_0 - p_1$ and $ReF \in H$ when F is analytic on $\bar{W}$.

V. The Existence of Distinct Principal Analytic Functions.

15. We consider the problem of determining conditions under which there exist two different principal analytic functions on the planar bordered Riemann surface $\bar{W}$ of no. 1. The principal analytic functions under consideration are defined in no. 13, and have properties described in Theorems 5, 6, and 7 of no. 14. The following concepts are dealt with in Ahlfors and Sario [1].

DEFINITION. Two compact sets in the plane, each with connected complement, are said to be *equivalent* if their complements are conformally equivalent.

For the remainder of this chapter, we let E be a compact plane set with connected complement.

THEOREM (Ahlfors and Sario [1]). *The complement of E is of class 0_{AD} if and only if every set which is equivalent to E has 2-dimensional Lebesgue measure 0.*

DEFINITION. Let U be any open set which contains E, and suppose that a function F is analytic on $U - E$. E is said to be a *removable singularity for F* if there exists analytic extension of F to U.

THEOREM (Ahlfors and Sario [1]). *E is a removable singularity for all functions of class AD in a neighborhood of E if and only if the complement of E (with respect to the Riemann sphere) is of class 0_{AD}.*

16. DEFINITION. A planar bordered Riemann surface $\bar{W}$ as described in no. 1 is said to have rigid radius when $r(F)$ is constant for every F in the class A of no. 13.

THEOREM 11. *Let F_0 and F_1 be the principal analytic functions* *belonging to A. The surface $\bar{W}$ has rigid radius if and only if $F_0 = cF_1$, where $|c| = 1$.*

Proof. If $\bar{W}$ has rigid radius, then according to Theorems 6 and 7, both F_0 and F_1 minimize the same functional. Hence $F_0 = cF_1$, with $|c| = 1$. On the other hand, if $F_0 = cF_1$, we conclude from the corollaries of Theorems 6 and 7 that F_0 maximizes, and F_1 minimizes the functional $r(F)$ among all $F \in A$. Because $|c| = 1$, we have that the radius is rigid.

6. AD-Removability.

17. Our next condition for distinguishing F_0 from F_1 is most naturally stated if we take the bordered Riemann surface $\bar{W}$ to be a plane region, with γ and δ as outer and inner boundaries respectively. In addition, we let $\bar{W}_1$ denote the plane point set bounded by γ and δ, with E the difference $\bar{W}_1 - \bar{W}$.

THEOREM 12. *Let F_0 and F_1 be the principal analytic functions of no. 13. Then $F_0 = cF_1$, with $|c| = 1$, if and only if $S^2 - E \in 0_{AD}$.*

Sufficiency. F_0 and F_1 map a neighborhood U of E onto an open set of finite area and are of class AD in this neighborhood of E. Then according to no. 15, the principal analytic functions may be extended to all of $\bar{W}$. If the extension $\tilde{F}_i$ of F_i satisfies $\tilde{F}_i(z_0) = w_0$ for some w_0 with $r_\delta(\tilde{F}_i) < |w_0| < r_\gamma(\tilde{F}_i)$, then

$$(2\pi i)^{-1}\int_{\delta+\gamma} \frac{d\tilde{F}_i}{\tilde{F}_i - w_0} = (2\pi i)^{-1}\int_{\delta+\gamma} \frac{dF_i}{F_i - w_0} = 1 \,, \qquad i = 0, 1 \,.$$

Since $F_i \in A$, the second integral is 1 and the extensions are univalent. This means that $\tilde{F}_1 \circ \tilde{F}_0^{-1}$ is a conformal mapping of a full closed annulus, and in fact that $r(F_0)$ is equal to $r(F_1)$. We have $F_0 = cF_1$, with $|c| = 1$, as a consequence of Theorem 11.

Necessity. If $S^2 - E$ is not of class 0_{AD}, then, according to no. 15, there exists a one to one conformal mapping with positive complementary area. Such a mapping will have positive complimentary logarithmic area as well. Therefore, according to Theorem 8, $\mu(\sqrt{F_0 \cdot F_1})$ is positive, and Theorem 6 guarantees that $F_0 \neq cF_1$.

APPENDIX

An argument of Ahlfors and Beurling [1] (p. 111), which will be referred to and not repeated, is crucial in the proof of:

18. THEOREM 11. *The analytic function $P_n = \sqrt{F_{0n} \cdot F_{1n}}$ is of* *class* A_n.

Proof. Verification of the Conditions 1, 2, and 3 for A_n of no. 5 is immediate. Only (4) remains to be checked. If $\log F_{1n}$ and $\log F_{0n}$ are considered in the roles of q and p of Ahlfors and Beurling [**1**], p. 111, then $\log \sqrt{F_{0n} \cdot F_{1n}}$ may be considered in the role of $\frac{1}{2}(q + p)$. We observe that $d(\log F_{1n})/d(\log F_{0n})$ is well defined on the approximating W_n. Hence, by the technique of Ahlfors and Beurling already cited, we may conclude that $Re(d \log F_{1n}/d \log F_{0n})$ is of constant sign with no zeros in W_n. This implies that the image of each contour β_i is a convex curve, and each image is traced once as each β_i is traced once. This also implies that each of the curves $F(\beta_i)$ is traced in the same direction, and this direction will be determined now for one $F(\beta_i)$.

We observe that for each i, $P_n(\beta_i)$ is a compact set, and we may then choose w_i and w_i' so that w_i is that point of $P_n(\beta_i)$ which is closest to $P_n(\gamma)$ and w_i' is that point of $P_n(\gamma)$ which is closest to $P_n(\beta_i)$. We now assume that the β_i are indexed so that $\min\{d(w_i, w_i'); i = 1, 2, \cdots, k(n)\}$ is $d(w_1, w_1')$ where $d(w, w')$ is the usual Euclidean distance from w to w'. That is to say, $P_n(\beta_1)$ is as close to $P_n(\gamma)$ as any of $P_n(\beta_2), \cdots, P_n(\beta_{k(n)})$. The line segment Γ joining w_1 to w_1' is a univalence path for P_n in the sense that each point of Γ is taken exactly once by a point of W_n. Clearly P_n is one to one on $P_n^{-1}\Gamma$, and we may conclude that β_1 and $P_n(\beta_1)$ are similarly oriented. The reasoning in the paragraph above then establishes that each $P_n(\beta_i)$ is oriented as is $P_n(\beta_1)$, and in fact, for each i we have that the winding number for points inside $P_n(\beta_i)$ is -1.

An application of the argument principle is now all that is needed to show that P_n is univalent on Int $W_n \cup \delta \cup \gamma$.

BIBLIOGRAPHY

1. L. Ahlfors, [**1**] *Complex analysis*, McGraw-Hill Book Company, Inc., (1952).
2. L. Ahlfors, and A. Beurling, [**1**] *Conformal invariants and function-theoretic null-sets*, Acta Math., **83** (1950), 101-129.
3. L. Ahlfors, and L. Sario, [**1**]*Riemann surfacees*, Princeton University Press, (1960).
4. R. Courant, [**1**] *Dirichlet's principle, conformal mapping, and minimal surfaces*, Interscience Publishers (1950), 13-14.
5. R. de Possel, [**1**] *Sur quelpues propriétés de la representation conforme des domains multiplement connexes, en relation avec le théorème des fentes parallèles*, Math. Ann., **107** (1932), 496-504.
6. L. Sario, [**1**] *Ueber Riemannsche Flächen mit hebbarem Rand*, Ann. Acad. Sci. Fenn. Ser. A. I. no. **50** (1948), 79 pp.
7. ———, [**2**] *A linear operator method on arbitrary Riemann surfaces*, Trans. Amer. Math. Soc., **72** (1952), 281-295.
8. ———, [**3**] *Strong and weak boundary components*, J. Analyse Math. **5** (1956/57), 389-398.
9. M. Schiffer, [**1**] *Some recent developments in the theory of conformal mapping, Appendix to R. Courant, Dirichlet's principle, conformal mapping, and minimal surfaces*, Interscience Publishers (1950), 249-265.

PRIMAL CLUSTERS OF TWO-ELEMENT ALGEBRAS

EDWARD S. O'KEEFE

1. Introduction. The development of a structure theory for universal algebras which subsumes the familiar structure theory of Boolean and Post algebras and p-rings (Foster, [1]-[4]) has focused attention on certain classes of functionally complete universal algebras, called primal clusters. A primal cluster is a set of primal algebras in which every finite subset is strictly independent (see definitions § 2, below). Each such cluster determines a unique subdirect factorization for each algebra satisfying all the identities common to some finite subset of the cluster. In other words, every function over a direct product of strictly independent primal algebras, expressible in terms of the algebras' operations, has a decomposition and reconstruction analogous both to the Boolean theory and the Fourier transform theory. In order to broaden the domain of application of the generalized theory, we must find strictly independent sets of primal algebras.

The purpose of this paper is to present the theory of independence of primal algebras in a new dimension. Simple necessary and sufficient conditions for strict independence of primal algebras of one primitive operation, regardless of the number of elements, have been obtained [5]. We now give necessary and sufficient conditions for strict independence of certain two-element primal algebras of the same species, regardless of the number of primitive operations.

2. Basic notions, the ϕ-condititition. The following definitions are stated for easy reference.

Let $\mathfrak{A} = (A, o_1, o_2, \cdots)$ be a universal algebra.

2.1. The *species* $Sp = \{n_1, n_2, \cdots\}$ of $\mathfrak{A}$ is the sequence of ranks of the primitive operations o_i of $\mathfrak{A}$, where n_i is the rank of o_i.

2.2. An *expression* $\phi(\xi_1, \cdots, \xi_n)$ of species Sp is a finite set of one or more indeterminate symbols ξ_i, composed by operation-symbols of Sp.

2.3. *A strict $\mathfrak{A}$-function* is an expression interpreted in algebra $\mathfrak{A}$. The notation $\phi = \chi(\mathfrak{A})$ means that the strict function represented by ϕ in algebra $\mathfrak{A}$ is the same as that for χ.

2.4. $\mathfrak{A}$ is a *primal* algebra if every transformation of $A \times A \times \cdots \times A$ into A can be represented by a strict $\mathfrak{A}$-function, and Sp is denumerable.

Received April 25, 1960

2.5. A finite set of algebras $\{\mathfrak{A}_1, \mathfrak{A}_2, \cdots, \mathfrak{A}_p\}$, all of the same species Sp, is *strictly independent* if each given set of strict functions ϕ_i has a single expression ψ which reduces to the given function ϕ_i in the algebra $\mathfrak{A}_i$; i.e., $\psi = \phi_i(\mathfrak{A}_i)$.

2.6. $\widetilde{\mathfrak{A}}$ is a primal cluster if $\widetilde{\mathfrak{A}}$ is a set of primal algebras and every finite subset of $\widetilde{\mathfrak{A}}$ is strictly independent. The totality of pairwise non-isomorphic primal algebras of species $[s]$ constitutes a primal cluster [5]. Various other categories of primal clusters are known, largely of species [2, 1].

The ϕ-condition is analogous to the factorization of functions of real numbers. It is simply that any strict function may be represented by any expression operating on some set of strict functions.

2.7. *The ϕ-condition.* For every strict $\mathfrak{A}$-function, $\theta(\xi, \eta, \cdots, \zeta)$, and every strict $\mathfrak{A}$-function, $\kappa(\xi_1, \cdots, \xi_m)$, provided that no variable ξ_i occurs twice in κ, there exist strict $\mathfrak{A}$-functions, $\psi_1(\xi, \eta, \cdots, \zeta)$, $\cdots$, $\psi_m(\xi, \eta, \cdots, \zeta)$, such that

$$\kappa(\psi_1(\xi, \eta, \cdots, \zeta), \cdots, \psi_m(\xi, \eta, \cdots, \zeta)) = \theta(\xi, \eta, \cdots, \zeta) . \tag{2.1}$$

Formerly primal algebras were defined to be finite. However, this property is now derived from the denumerability of Sp.

THEOREM 2.8. *Every primal algebra is finite.*

Proof. Let $\mathfrak{A} = \langle A, o_1, \cdots, o_n, \cdots \rangle$ be a primal algebra. The two-valued functions on any infinite set have a larger cardinal number than the set of expressions made of a denumerable set of operations. Therefore, the fact that the functions on $A \times A$ to A are represented by expressions in the operations of A means that A is not infinite.

From [5], we require the following basic results.

THEOREM 2.9. *In any primal algebra in which the primitive operations are onto transformations, the ϕ-condition holds.*

THEOREM 2.10. *Let $\mathfrak{A} = (A, o, \cdots)$ and $\mathfrak{A} = (B, o, \cdots)$ be two non-isomorphic primal algebras of the same species, Sp. Then there exists a set of unary expressions $\{\phi_i\} = \{\phi_1, \cdots, \phi_p\}$ of species Sp such that*

$$\phi_1 = \phi_2 = \cdots = \phi_p(\mathfrak{A}) \tag{2.2}$$

and such that every unary $\mathfrak{B}$-function is equivalent modulo $\mathfrak{B}$ to one of the $\phi_1, \cdots \phi_p$.

THEOREM 2.11. *Let* $\{\mathfrak{A}_1, \cdots, \mathfrak{A}_n\}$ *be a set of universal algebras of species* Sp, *in which every pair of algebras is strictly independent. If the* ϕ-*condition holds in each algebra, the set is strictly independent.*

3. The two-element independence theorem.

Our main result is

THEOREM 3.1. *Every set of primal algebras is a primal cluster if*:

(i) *every algebra in the set has exactly two elements,*

(ii) *no two algebras are isomorphic,*

(iii) *no primitive operation is constant,*

(iv) *all algebras in the set are in the same species,* $Sp = [n_1, \cdots, n_m]$.

The proof of Theorem 3.1 is preceded by three lemmas.

LEMMA 3.2. *Let* $\mathfrak{B} = (\{\beta_1, \beta_2\}, o, \cdots)$ *be a two-element primal algebra with no constant primitive operations. Every expression* $\phi(\xi_1, \cdots, \xi_n)$, *in which no variable occurs twice may be changed, modulo algebra* $\mathfrak{B}$, *to any given function* $\chi(\xi)$ *by replacing some variable by a properly chosen strict* $\mathfrak{B}$-*function* $\psi(\xi)$, *and all others by constant strict* $\mathfrak{B}$-*functions.*

Proof. If the expression $\phi(\xi_1, \cdots, \xi_n)$ has but one operation-symbol o_i, then, since no operation-symbol represents a constant, there are constants δ_i and γ_i such that

$$
\begin{aligned}
o_i(\delta_1, \cdots, \delta_{n_i}) &= \chi(\beta_1) \\
o_i(\gamma_1, \cdots, \gamma_{n_i}) &= \chi(\beta_2) \ .
\end{aligned}
\tag{3.1}
$$

We alter δ_1 to γ_1, δ_2 to γ_2, etc. until the function changes value. Some jth argument must give the change from $\chi(\beta_1)$ to $\chi(\beta_2)$. We choose the expression $\psi(\xi)$ so that

$$
\begin{aligned}
&\psi(\beta_1) = \delta_j \\
&\psi(\beta_2) = \gamma_j \\
&o_i(\gamma_1, \cdots, \gamma_{j-1}, \psi(\beta_k), \delta_{j+1}, \cdots, \delta_{ni}) = \chi(\beta_k) \ , \qquad (k = 1, 2) \ .
\end{aligned}
\tag{3.2}
$$

Since there are but two elements in the algebra $\mathfrak{B}$, χ is now completely represented

$$
o_i(\gamma_1, \cdots, \gamma_{j-1}, \psi(\xi), \delta_{j+1}, \cdots, \delta_{n_i}) = \chi(\xi) \ . \tag{3.3}
$$

On the other hand, let $\phi(\xi_1, \cdots, \xi_n)$ be composed of m operation-symbols. Assume that the theorem holds for all expressions with fewer than m operation-symbols. ϕ is a set of expressions $\phi_1, \cdots, \phi_{n_i}$ composed by primitive operations o_j: $\phi = o_j(\phi_1, \cdots, \phi_{n_i})$. $\phi_1, \cdots, \phi_{n_i}$ have fewer than m operation-symbols, so by assumption,

$$\begin{cases} \phi_k = \gamma_k, \text{ for } k = 1, \cdots, j-1 \\ \phi_j = \psi(\xi) \\ \phi_{j+1} = \delta_{j+1}, \cdots, \phi_{n_i} = \delta_{n_i}, \end{cases} \tag{3.4}$$

where all variables but one have been replaced by constants. But, obviously, in ϕ_k, $k \neq j$, the last variable may also be replaced by a constant, since a constant result is desired and is given by either value of the variable. This leaves only one variable in ϕ_j; but with these replacements

$$\phi(\phi_1, \cdots, \phi_{n_i}) = \chi(\xi) \tag{3.5}$$

and the proof is complete.

LEMMA 3.3. *If in two primal algebras $\mathfrak{A}$ and $\mathfrak{B}$, $\mathfrak{A}$ satisfies the ϕ-condition, then, for every $\beta \in B$ and every $a(\xi)$, there is an expression $\Pi(\xi)$ such that*

$$\Pi(\xi) = \begin{cases} a(\xi)(\mathfrak{A}) \\ \beta\ (\mathfrak{B}). \end{cases} \tag{3.6}$$

Proof. Modulo $\mathfrak{B}$, there must exist expressions for constants in B. Therefore, letting $\kappa_\beta(\xi) = \beta(\mathfrak{B})$, replace each occurrence of ξ in κ by a variable from the set $\xi_1, \cdots, \xi_p$, so that in $\kappa_\beta(\xi_1, \cdots, \xi_p)$, no variable occurs more than once. Applying the ϕ-condition to $\kappa_\beta(\xi_1, \cdots, \xi_p)$ with respect to A, there $\psi_1, \cdots, \psi_p$ such that

$$\kappa_\beta(\psi_1, \cdots, \psi_p) = a(\xi)(\mathfrak{A}). \tag{3.7}$$

By Theorem 2.10, there exists a set of expressions $\{\phi_i\}$ with

$$\phi_i = \psi_i(\mathfrak{A}) \text{ and } \phi_1 = \phi_2 = \cdots = \phi_p(\mathfrak{B}). \tag{3.8}$$

Then

$$\kappa_\beta(\phi_1, \cdots, \phi_p) = a(\xi)(\mathfrak{A}), \tag{3.9}$$

by (3.7), but

$$\kappa_\beta(\phi_1, \cdots, \phi_p) = \kappa_\beta(\phi_1(\xi))(\mathfrak{B}) = \beta(\mathfrak{B}). \tag{3.10}$$

LEMMA 3.4. *Let $\mathfrak{A} = (A, o_1, \cdots, o_m)$ be a primal algebra of species Sp in which every primitive operation o_i is a transformation onto A. Let $\mathfrak{B}$ be a two-element primal algebra of the some species Sp, with no constant primitive operations. Then if $\mathfrak{A}$ and $\mathfrak{B}$ are not isomorphic, they are strictly independent.*

Proof. The operations of $\mathfrak{B}$ are transformations onto B, since they are non-constant and B has only two elements. Moreover, $\mathfrak{B}$ is primal. Therefore Theorem 2.9 applies; the ϕ-condition holds in algebra $\mathfrak{B}$. $\mathfrak{A}$ is also primal with the same kind of primitive operations; hence, by Theorem 2.9, the ϕ-condition holds for $\mathfrak{A}$ too.

Since $\mathfrak{A}$ is primal, there exists an expression, $\Sigma(\xi, \zeta)$, and an element $o \in A$ such that

$$\begin{aligned} \Sigma(\xi, o) &= \xi , \\ \Sigma(o, \zeta) &= \zeta . \end{aligned} \tag{3.11}$$

Let p be the number of occurrences of ξ in Σ and q the number of occurrences of ζ. Replace each occurrence of ξ or ζ by a different variable from the set $(\zeta_1, \cdots, \xi_p)$ or $(\zeta_1, \cdots, \zeta_q)$ respectively. Let the resulting expression be denoted $\Sigma(\xi_1, \cdots, \xi_p, \zeta_1, \cdots, \zeta_q)$. By Lemma 3.2, there exist a strict $\mathfrak{B}$-function $\psi_j(\zeta)$ and constant $\mathfrak{B}$-functions such that

$$\Sigma(\gamma_1, \cdots, \gamma_{j-1}, \psi_j(\zeta), \beta_{j+1}, \cdots, \beta_{p+q}) = \zeta(\mathfrak{B}) . \tag{3.12}$$

Suppose $j \leqq p$, then by Theorem 2.10, there are $\phi_i(\zeta)$ such that

$$\phi_i = \begin{cases} 0(\mathfrak{A}) & (i = 1, \cdots, p) \\ \gamma_i(\mathfrak{B}) & (i = 1, \cdots, j-1) \\ \psi_j(\xi) & (\mathfrak{B}) \ i = j \\ \beta_i(\mathfrak{B}) & (i = j+1, \cdots, p) \end{cases} \tag{3.13}$$

and by Lemma 3.3, $\phi_i(\xi)$ such that

$$\phi_i = \begin{cases} \xi(\mathfrak{A}) & (i = p+1, \cdots, p+q) \\ \beta_i(\mathfrak{B}) & (i = p+1, \cdots, p+q) . \end{cases} \tag{3.14}$$

Thus,

$$\Sigma(\phi_1, \cdots, \phi_{p+q}) = \begin{cases} \Sigma(0, \xi) = \xi(\mathfrak{A}) \\ \Sigma(\gamma_1, \cdots, \gamma_{j-1}, \psi_j(\zeta), \beta_{j+1}, \cdots, \beta p_+ q) = \zeta(\mathfrak{B}) . \end{cases} \tag{3.15}$$

An exactly similar argument shows the construction if $p < j$. Therefore, it is always possible to find an expression χ such that

$$\chi(\xi, \zeta) = \begin{cases} \xi(\mathfrak{A}) \\ \zeta(\mathfrak{B}) , \end{cases} \tag{3.16}$$

and the two algebras are strictly independent by Definition 2.5.

We now return to the proof of Theorem 3.1.

Proof. Each algebra is primal, and every primitive operation is an onto transformation because none is constant and each algebra has but two elements. Therefore, by Theorem 2.9, the ϕ-condition holds in each algebra. Moreover, by Lemma 3.4, each pair of algebras is independent. Therefore, by Theorem 2.11, every finite subset of $\{\mathfrak{A}_1, \cdots\}$ is independent, and $\{\mathfrak{A}_1, \cdots, \mathfrak{A}_n, \cdots\}$ is a primal cluster.

BIBLIOGRAPHY

1. A. L. Foster, *Generalized "Boolean" theory of universal algebras*, Part 1: *Subdirect sums and Normal represention theorem*, Math. Zeit., **58** (1953), 306-336.

2. ———, *Generalized "Boolean" theory of universal algebras*, Part II: *Identities and subdirect sums of functionally complete algebras*, Math. Zeit., **59** (1953), 191-199.

3. ———, *The identities of—and unique subdirect factorization within—Classes of universal algebras*, Math. Zeit., **62** (1955), 171-188.

4. ———, *Ideals and their structure in classes of operational Algebras*, Math. Zeit., **65** (1956), 70-75.

5. E. S. O'Keefe, *On the independence of primal algebras*, Math. Zeit., **73** (1960), 79-94.

6. E. L. Post, *The two-valued iterative systems of mathematical logic*, Annals of Math Studies, Princeton, 1941.

APPLICATIONS OF THE TOPOLOGICAL METHOD OF WAZEWSKI TO CERTAIN PROBLEMS OF ASYMPTOTIC BEHAVIOR IN ORDINARY DIFFERENTIAL EQUATIONS

NELSON ONUCHIC

Introduction. The main objective of this paper is to present some results concerning the asymptotic behavior of the integrals of some systems of ordinary differential equations.

As Wazewski's theorem, used in our work, is not very well known, we state it here, giving first some definitions and notations.

HYPOTHESIS H. (a) *The real-valued functions* $f_i(t, x_1, \cdots, x_n)$, $i = 1, \cdots, n$, *of the real variables* $t, x_1, \cdots, x_n$, *are continuous in an open set* $\Omega \subset R^{n+1}$.

(b) *Through every point of* Ω *passes only one integral of the system*

$$\dot{x} = f(t, x) \quad \left(\cdot = \frac{d}{dt}\right) \quad \textit{where}$$

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad f(t, x) = \begin{pmatrix} f_1(t, x_1, \cdots, x_n) \\ \vdots \cdots\cdots\cdots\cdots \\ f_n(t, x_1, \cdots, x_n) \end{pmatrix} \qquad \textit{and } (t, x) \in \Omega .$$

Let ω be an open set of R^{n+1}, $\omega \subset \Omega$ and let us denote by $B(\omega, \Omega)$ the boundary of ω in Ω.

Let P_0: $(t_0, x_0) \in \Omega$. We write $I(t, P_0) = (t, x(t, P_0))$, where $x(t, P_0)$ is the integral of the system $\dot{x} = f(t, x)$ passing through the point P_0.

Let $(\alpha(P_0), \beta(P_0))$ be the maximal open interval in which the integral passing through P_0 exists. We write

$$I(\Delta, P_0) = \{(\mathrm{t}, x(t, P_0)) \mid t \in \Delta\}$$

for every set Δ contained in $(\alpha(P_0), \beta(P_0))$.

We say that the point P_0: $(t_0, x_0) \in B(\omega, \Omega)$ is a *point of egress* from ω (with respect to the system $\dot{x} = f(t, x)$ and the set Ω) if there exists a positive number δ such that $I([t_0 - \delta, t_0), P_0) \subset \omega$; P_0 is a *point of strict egress* from ω if P_0 is a point of egress and if there exists a positive number δ such that $I((t_0, t_0 + \delta], P_0) \subset \Omega - \overline{\omega}$. The set of all points of egress (strict egress) is denoted by $S(S^*)$.

If $A \subset B$ are any two sets of a topological space and $K: B \to A$ is

Received November 28, 1960.

a continuous mapping from B onto A such that $K(P) = P$ for every $P \in A$, then K is said to be a *retraction* from B into A and A a *retract* of B.

THEOREM OF WAŻEWSKI. *Suppose that the system* $\dot{x} = f(t, x)$ *and the open sets* $\omega \subset \Omega \subset R^{n+1}$ *satisfy the following hypotheses*:

(1) *Hypothesis H.*

(2) $S = S^*$.

(3) *There exists a set* $Z \subset \omega \cup S$ *such that* $Z \cap S$ *is a retract of* S *but is not a retract of* Z.

Then there is at least one point P_0: $(t_0, x_0) \in Z - S$ *such that* $I(t, P_0) \subset \omega$ *for every* $t_0 \leqq t < \beta(P_0)$.

The theorem of Ważewski [6, Théorème 1, p. 299] is actually more general than the one stated above.

If $f_i(t, x_1, \cdots, x_n)$, $i = 1, \cdots, n$, are complex-valued functions of the real variable t and of the complex variables $x_1, \cdots, x_n$, the n-dimensional complex system $\dot{x} = f(t, x)$ can be considered as a $2n$-dimensional real system, so that the theorem of Ważewski is also extensible, in a natural way, to complex systems [5, p. 19. § 1 and p. 21, § 2].

The most difficult part in the applications of the method of Ważewski is, in general, to verify that $S = S^*$. To accomplish this Ważewski introduced the concept of a regular polyfacial set [6, § 14 p. 307 and § 15, p. 309]. However the distinction established by Ważewski between positive and negative faces has certain inconveniences. In some applications of the method of Ważewski there appear sets ω such that $S = S^*$ but whose faces are only "almost positive" and "almost negative". We thus have to work sometimes with sets ω that are similar, in some sense, to the regular polyfacial sets and that satisfy the condition $S = S^*$.

In the first part of our work we give a generalization of polyfacial regular sets eliminating the distinction between positive and negative faces and such that the main theorem concerning the polyfacial regular sets [6, Théorème 5, p. 310] remains valid. We observe that the sets ω considered in Z. Szmydtówna's paper [5, § 4, Théorème 1, p. 24][1], in our Theorem II-1 and in Barbălat's paper [1, Théorème 1, p. 303; Théorème 2, p. 305] are generalized regular polyfacial sets, in our sense, but are not regular polyfacial sets.

Szmydtówna [5, Corollaire 1-Remarque 2, p. 30] proves a theorem

[1] Szmydtówna's Theorem 1 is false. We observed that the proof is wrong because the statement: "La frontière de ω touchant celle de Ω exclusivement sur le plan $t = \infty$ $\cdots$" [5, p. 28] is false.

J. Lewowics [3], developing a counter-example suggested by J. L. Massera, has shown that the theorem is actually false. Nevertheless, Theorems 2 and 3 deduced from Theorem 1 are correct because, in the particular case of linear systems $\dot{x} = A(t)x$, with $A(t)$ defined for $T \leqq t < \infty$, the solutions are defined for all $T \leqq t < \infty$.

which generalizes a theorem of Perron. In part II of our work (Theorem II-1) we obtain the same conclusion but starting from hypotheses different from those of Szmydtówna.

Note[2]. Our Theorem II-1 improves a result of N. I. Gavrilov. I. M. Rapoport in his book "On some asymptotic methods in the theory of differential equations", Kiev (1954) has also studied problems of this type. For some reference to their work to see "Forty years of Soviet Mathematics", Moscow (1959), Vol. i., pp. 520–521.

Our Theorem II-2 follows the same line of ideas.

Theorem II-3, due to Professor J. L. Massera, shows that in the case $n = 2$ the asymptotic behavior can be described more completely.

Consider two systems

$$\dot{y} = A(t)y \tag{1}$$

$$\dot{x} = A(t)x + g(t, x) \tag{2}$$

where $A(t)$ is a continuous matrix for $t \geqq T$ and $g(t, x)$ a continuous vector-function in $\Omega = [T, \infty) \times R^{2n}$.

Suppose that $g(t, x)$ satisfies some condition ensuring the uniqueness of the solution through each point $P_0 \in \Omega$ and that all solutions are defined for $T \leqq t < \infty$. We say that (1) and (2) are asymptotically equivalent if there exists a homeomorphism ϕ from the plane $t = T$ onto itself such that if $Q_0 = \phi(P_0)$ then $\lim_{t\to\infty} [x(t, P_0) - y(t, Q_0)] = 0$ [4, Cap. IX, § 4, p. 634].

In part III of our work the main result is the establishment of a condition that implies the asymptotic equivalence between two linear systems (Theorem III-3).

The author is deeply indebted to Professor J. L. Massera for his constant guidance and invaluable help during the preparatation of this paper, the result of work done at the Instituto de Matemática y Estadística, Montevideo, Uruguay

PART I

Let the real-valued functions

$$f_i(t, x_1, \cdots, x_n) \,, \qquad i = 1, \cdots, n \,,$$

of real variables $t, x_1 \cdots, x_n$ belong to C^p, $p \geqq 1$, on an open set $\Omega \subset R^{n+1}$, i.e., all partial derivatives

$$\frac{\partial^k_{f_i}}{\partial t^{p_0} \partial x_1^{p_1} \cdots \partial x_n^{p_n}} \qquad (p_0 + p_1 + \cdots + p_n = k \leqq p)$$

[2] The information given in this Note is due to the referee. We have not had access to the above works. We are indebted to him for this.

exist and are continuous on Ω.

Consider the differential system

$$\text{(I)} \qquad \dot{x} = f(t, x)$$

where

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ and } f(t, x) = \begin{pmatrix} f_1(t, x_1, \cdots, x_n) \\ \vdots \cdots\cdots\cdots\cdots \\ f_n(t, x_1, \cdots, x_n) \end{pmatrix}$$

with $(t, x) \in \Omega$.

Let $g(t, x)$ be a real-valued function belonging to C^{p+1} on Ω, let P_0: $(t_0, x_0) \in \Omega$ and let $x(t)$ be the integral of system (I) passing through the point P_0. We set $\varphi(t) = g(t, x(t))$; since $f(t, x) \in C^p$ and $g(t, x) \in C^{p+1}$ it follows $\varphi(t) \in C^{p+1}$ on $(\alpha(P_0), \beta(P_0))$.

The qth derivative, $q \leqq p + 1$, of $g(t, x)$ at the point P_0: (t_0, x_0) with respect to the system (I), is by definition

$$\left[\frac{d^q}{dt^q}\varphi(t)\right]_{t_0} \text{ and is denoted by } [D^q_{(I)}g(P)]_{P_0}\,.$$

Let $H_i(P) = H_i(t, x)$, $i = 1, \cdots, m$, be functions $\in C^{p+1}$ on the open set $\Omega \subset R^{n+1}$.

Let

$$\omega = \{P \in \Omega \mid H_i(P) < 0, i = 1, \cdots, m\}$$
$$\Gamma_i = \{P \in \Omega \mid H_i(P) = 0, H_j(P) \leqq 0, j = 1, \cdots, m\}$$

The Γ_i are called *faces* of ω.

Such a set ω will be called a *generalized regular polyfacial set* relative to (I) if, for each $i = 1, \cdots, m$ and each P_0: $(t_0, x_0) \in \Gamma_i$, the following alternative holds:

(1) The smallest index $q \leqq p + 1$ such that $[D^q_{(I)}H_i(P)]_{P_0} \neq 0$ is odd and the corresponding derivative is positive;

(2) P_0 is not a point of egress.

Let L_i, M_i be the corresponding sets of points. Useful criteria to verify $P_0 \in M_i$ are:

(*a*) the smallest index $q \leqq p + 1$ such that $[D^q_{(I)}H_i(p)]_{P_0} \neq 0$ is either odd with a negative value of the derivative or even with a positive value of the derivative;

(b) There exists $[a, b] \subset (\alpha(P_0), \beta(P_0))$ such that $a < t_0 \leqq b$ and $I([a, b], P_0) \subset \Gamma_i$.

LEMMA 1. *If ω is a generalized regular polyfacial set relative to* (I),

$$S = S^* = \bigcup_{i=1}^{m} L_i - \bigcup_{i=1}^{m} M_i .$$

Proof. Since $\Gamma_i = L_i \cup M_i$, $B(\omega, \Omega) \subset \bigcup_{i=1}^{m} \Gamma_i$,

$$S^* \subset S \subset \bigcup_{i=1}^{m} L_i - \bigcup_{i=1}^{m} M_i ,$$

it is enough to show that any point P_0 belonging to this last set is a point of strict egress. For such a P_0, $J = \{j \mid P_0 \in L_j\} \neq \phi$. If $j \in J$, $H_j(P_0) = 0$ and there exists a $\delta > 0$ such that $H_j(t, I(t, P_0)) < 0$ in $[t_0 - \delta, t_0)$ and $H_j(t, I(t, P_0)) > 0$ in $(t_0, t_0 + \delta]$. If $j \notin J$, $P_0 \notin \Gamma_j$ whence $H_j(P_0) < 0$ and there exists a $\delta > 0$ such that $H_j(t, I(t, P_0)) < 0$ in $[t_0 - \delta, t_0)$. There exists therefore a $\delta > 0$ such that $H_j(t, I(t, P_0)) < 0, j = 1, \cdots, m$, $t \in [t_0 - \delta, t_0)$, and, for at least one $j(\varepsilon J)$, $H_j(t, P_0)) > 0, t \in (t_0, t_0 + \delta]$, so that $P_0 \in S^*$.

PART II

Consider the linear differential system

$$\dot{y}_i = f_i(t)y_i + \sum_{j=1}^{n} g_{ij}(t)y_j , \qquad i = 1, \cdots, n$$

where the coefficients $f_i, g_{ij}, T \leqq t < \infty$, are continuous functions (in general complex-valued) of the real variable t.

By using Ważewski's method Z. Szmydtówna proved that if

$$R(f_k - f_{k+1}) > 0 , \quad \int_T^\infty R(f_k - f_{k+1})dt = \infty , \qquad k = 1, \cdots, n-1 ,$$

and

$$\lim_{t\to\infty} \frac{g_{ij}}{R(f_k - f_{k+1})} = 0 , \qquad i, j = 1, \cdots, n , \quad k = 1, \cdots, n-1 ,$$

then there is a system of n linearly independent solutions $(y_{1k}, \cdots, y_{nk})$, $k = 1, \cdots, n$, with $\lim_{t\to\infty} y_{ik}/y_{kk} = 0$ for $i \neq k$ [5, Corollaire 1, Remarque 2, p. 30]. This theorem generalizes a theorem of Perron who obtains the same result requiring the existence of a constant $c > 0$ such that $R(f_k) > R(f_{k+1}) + c$, $k = 1, \cdots, n-1$, and $\lim_{t\to\infty} g_{ij} = 0$.

We notice that Szmydtówna allows the f_i, $i = 1, \cdots, n$, to be large and the g_{ij} to be small in some sense. In the following theorem we obtain the same result allowing also the f_i to be large and the g_{ij} to be small but in a sense completely different from Szmydtówna's.

THEOREM II-1. *Suppose that the system*

$$\text{(II)} \quad \dot{x}_i = f_i(t)x_i + \sum_{j=1}^{n} g_{ij}(t)x_j \,, \qquad i = 1, \cdots, n \,,$$

satisfies the following hypotheses:

(1) *The coefficients* $f_i, g_{ij}, T \leqq t < \infty$, *are continuous functions (in general complex-valued) of the real variable* t.

(2) *There exists a real-valued continuous function* $h(t), T \leqq t < \infty$, *such that for all* $i \neq j$ *we have*

$$| R(f_i - f_j) | \leqq h(t) \,,$$
$$\int_T^\infty | g_{ij}(t) | e^{H(t)} dt < \infty$$

and

$$\int_T^\infty | R(g_{ii} - g_{jj}) | e^{H(t)} dt < \infty \,,$$

where $H(t) = \int_T^t h(s)ds$

Then there is a system of n *linearly independent solutions*

$$(x_1(t), \cdots, x_n(t)) = \begin{pmatrix} x_{11}(t), \cdots, x_{1n}(t) \\ \cdots\cdots\cdots\cdots\cdots \\ x_{n1}(t), \cdots, x_{nn}(t) \end{pmatrix}$$

with $\lim_{t\to\infty} x_{ik}/x_{kk} = 0$ *for all* $i \neq k$.

Proof.

For every fixed integer p, $0 < p \leqq n$, we set

$$\omega_p = \{P: (t, x) \mid | x_i |^2 - | x_p |^2 \varphi^2(t) < 0, i \neq p, t > t_0 \geqq T\}$$

where $\varphi(t)$ and t_0 will be conveniently chosen so that, for every $t \geqq t_0$, $\varphi(t) > 0$, φ is differentiable, $\lim_{t\to\infty}\varphi(t) = 0$ and ω_p is a generalized regular polyfacial set.

Let

$$H_i(P) = | x_i |^2 - | x_p | \varphi^2(t) \,, \qquad i \neq p \,,$$
$$H_p(P) = t_0 - t \,,$$

it follows that $\omega_p = \{P \mid H_i(P) < 0, i = 1, \cdots, n\}$.

Set, for $q \neq p$,

$$\begin{aligned} \tilde{\Gamma}_p &= \Gamma_p - \{Q: (t, x) \mid x = 0\} \\ &= \{P \mid | x_q | = | x_p | \varphi(t), | x_i | \leqq | x_p | \varphi(t) \text{ for } i \neq p, t \geqq t_0, x_p \neq 0\} \,. \end{aligned}$$

An easy computation shows that

$$\frac{1}{2}[D_{(\mathrm{II})}H_q(P)]P \in \tilde{\Gamma}_q \geqq |x_p|^2\varphi^2(t)[R(f_q - f_p + g_{qq} - g_{pp})]$$
$$- |x_p|^2\varphi(t)\dot{\varphi}(t) - |x_p|^2\varphi^2(t)\sum_{j\neq p}|g_{pj}|\frac{|x_j|}{|x_p|}$$
$$- |x_p|^2\sum_{j\neq q}|g_{qj}|\frac{|x_j|}{|x_q|}\cdot\frac{|x_q|}{|x_p|} .$$

Since $|x_q| = |x_p|\varphi(t) \geqq |x_j|$ for $j \neq p$ it follows that $|x_j|/|x_p| \leqq \varphi(t)$. As we want $\varphi(t) > 0$ and $\lim_{t\to\infty}\varphi(t) = 0$ we can take t_0 such that $\varphi(t) < 1$ for $t \geqq t_0$. Then

$$\frac{1}{2}[D_{(\mathrm{II})}H_q(P)]P \in \tilde{\Gamma} \geqq |x_p|^2\varphi^2(t)R(f_q - f_p + g_{qq} - g_{pp})$$
$$- |x_p|^2\varphi(t)\dot{\varphi}(t) - |x_p|^2\varphi^2(t)\sum_{j\neq p}|g_{pj}| - |x_p|^2\varphi(t)\sum_{j\neq q}|g_{qj}| .$$

since

$$\varphi(t)R(f_q - f_p + g_{qq} - g_{pp}) - \dot{\varphi}(t) - \varphi(t)\sum_{j\neq p}|g_{pj}| - \sum_{j\neq q}|g_{qj}| >$$
$$- \dot{\varphi}(t) - \varphi(t)h(t) - g(t) ,$$

where

$$g(t) = \{\sum_{i\neq j}|R(g_{ii} - g_{jj})| + |g_{ij}|\} + e^{-H(t)-t} ,$$

in order to have, for $q \neq p$, $[D_{(\mathrm{II})}H_q(P) \in \tilde{\Gamma}_q > 0$, it is sufficient to choose $\varphi(t)$ such that

(A) $\dot{\varphi}(t) + \varphi(t)h(t) + g(t) = 0$.

$\varphi(t) = e^{-H(t)}\int_t^\infty g(s)e^{H(s)}\,ds$ is indeed a solution of (A) satisfying the conditions $\varphi(t) > 0$, φ differentiable and $\lim_{t\to\infty}\varphi(t) = 0$.

If ω_p is defined in this way, taking into account that $[D_{(\mathrm{II})}H_p(P)]_{P\in\Gamma_p} = -1$ and that the set $\{P \in \Gamma_q \mid x_p = 0\} \subset M_q$, for $q \neq p$, it follows that ω_p is a generalized regular polyfacial set.

For $i \neq p$ we have

$$L_i = \tilde{\Gamma}_i \text{ and } L_p = \phi ,$$
$$M_i = M = \{P\colon (t, x) \mid t \geqq t_0,\ x = 0\} \text{ and } M_p = \Gamma_p .$$

By Lemma 1

$$S = S = \bigcup_{i\neq p}\tilde{\Gamma}_i - \Gamma_p - M .$$

We choose

$$Z_p = \{P\colon (t, x) \mid t = \tau > t_0,\ x_p = x_p^0 \neq 0,\ |x_i| \leqq |x_p^0|\varphi(\tau),\ i \neq p\} = \prod_{j\neq p} B_j^2 ,$$

where B_j^2 is a solid sphere in R^2. We have

$$Z_p \cap S = Z_p \cap [\bigcup_{i \neq p} \tilde{\Gamma}_i - \Gamma_p - M] = \bigcup_{i \neq p} z_p \cap [\tilde{\Gamma}_i - \Gamma_p - M] \,.$$

For $i \neq p$

$$Z_p \cap [\tilde{\Gamma}_i - \Gamma_p - M] = \{P\colon (t, x) \mid t = \tau, x_p = x_p^0, \mid x_i \mid = \mid x_p^0 \mid \varphi(\tau), \\ \mid x_j \mid \leqq \mid x_p^0 \mid \varphi(\tau), j \neq p\} = B_1^2 \times \cdots \times B_{i-1}^2 \times S_i^1 \times B_{i+1}^2 \times \cdots \times B_n^2$$

(in the cartesian product above B_p^2 is exclued) where S_i^1 is the boundary of B_i^1 in R^2.

Modulo homeomorphisms we have therefore $Z_p = B^{2n-2}$ (solid sphere in R^{2n-2}) and $Z_p \cap S = S^{2n-3} =$ Boundary of B^{2n-2} in R^{2n-2}, so that $Z_p \cap S$ is not a retract of Z_p.

There is however a retraction $\phi\colon S \rightarrow Z_p \cap S$ given by $\phi(P) = P^*$, with $t^* = \tau$, $x_p^* = x_p^0$, $x_i^* = \varphi(\tau)/\varphi(t) \mid x_p^0 \mid / \mid x_p \mid \cdot x_i$, $i \neq p$. The verification is trivial.

By using the theorem of Ważewski we can conclude the existence of at least one point $P_0\colon (\tau, x_0) \in Z_p - S$ with $I(t, P_0) \subset \omega_p$ for every $t \geqq \tau$. This means that the solution $x_p(t) = (x_{1p}(t), \cdots, x_{np}(t))$ of (II) passing through P_0 satisfies

$$\frac{\mid x_{ip}(t) \mid}{\mid x_{pp}(t) \mid} < \varphi(t) \text{ for } t \geqq \tau \text{ and } i \neq p \,.$$

Letting $p = 1, \cdots, n$ we find n solutions $(x_1(t), \cdots, x_n(t))$ with the required property. Let us show that these solutions can be taken linearly independent.

By choosing Z_p with sufficiently large τ and $x_{pp}^0 = 1$ the absolute values of the coordinates x_{ip}, $i \neq p$, of the points of Z_p can be made arbitrarily small. We then have

$$\begin{pmatrix} x_1(\tau) \\ x_2(\tau) \\ \vdots \\ x_n(\tau) \end{pmatrix} = \begin{pmatrix} \varepsilon_{11} \cdots \varepsilon_{1n} \\ \varepsilon_{21} \cdots \varepsilon_{2n} \\ \cdots\cdots\cdots \\ \varepsilon_{n1} \cdots \varepsilon_{nn} \end{pmatrix}$$

where $\varepsilon_{ii} = 1$ and the $\mid \varepsilon_{ij} \mid$ are smaller than any given positive number for all $i \neq j$. This completes the proof

In the following theorem we will look for linearly independent solutions of (II) with similar properties to those of Theorem II-I but not necessarily requiring that they form a fundamental set of solutions of (II).

THEOREM II-2. *Suppose that the system (II) satisfies the following hypotheses:*

(1) *The coefficients* f_i, g_{ij}, $T \leqq t < \infty$, *are continuous functions (in general complex-valued) for the real variable* t.

(2) *There exists a natural number* $r \leqq n$ *such that* $R(f_r) = \cdots = R(f_n)$, $R(f_i) \geqq R(f_r)$ *for all* $i < r$, $\int_T^\infty | g_{ij}(t) | dt < \infty$ *for all* $i \neq j$ *and* $\int_T^\infty | R(f_{ii} - g_{jj}) | dt < \infty$.

Then there exists $s + 1$ $(r + s = n)$ *linearly independent solutions*

$$(x_r(t), \cdots, x_n(t)) = \begin{pmatrix} x_{1r}(t) \cdots x_{1n}(t) \\ \cdots\cdots\cdots\cdots \\ x_{nr}(t) \cdots x_{nn}(t) \end{pmatrix}$$

such that $\lim_{t\to\infty} x_{ik}/x_{kk} = 0$ for all $i \neq k$, $k = r, \cdots, n$.

Proof. Given an integer p, $r \leqq p \leqq n$, we prooceed exactly as in Theorem II-1 up to the point where we got the expression:

$$\varphi(t) R(f_q - f_p + g_{qq} - g_{pp}) - \dot\varphi(t) - \varphi(t) \sum_{j\neq p} | g_{pj} | - \sum_{j\neq q} | g_{qj} |$$

which we denote by B_q.

As we have $R(f_q - f_p) \geqq 0$ for all q, $0 < q \leqq n$, it follows $(\varphi(t) < 1)$

$$B_q \geqq - \dot\varphi(t) - g(t) \text{ where } \int_{t0}^\infty g(t) dt < \infty .$$

Making $\varphi(t) = \int_t^\infty [g(s) + e^{-s}] ds$ it follows that $\varphi(t) > 0$, φ is differentiable, $\lim_{t\to\infty} \varphi(t) = 0$ and $B_q > 0$.

Proceeding as in Theorem II-1 we find a set of $(s + 1)$ solutions $(x_r(t), \cdots, x_n(t))$. Still by a similar reasoning we may show that these solutions can be so chosen that for a sufficiently large τ we have

$$(x_r(\tau), \cdots, x(\tau)) = \begin{pmatrix} \varepsilon_{1r}, \cdots, \varepsilon_{1n} \\ \cdots\cdots\cdots\cdots \\ \varepsilon_{rr}, \cdots, \varepsilon_{rn} \end{pmatrix}$$

with $\varepsilon_{ii} = 1$ and the $| \varepsilon_{ij} |$, $i \neq j$, smaller than any given positive number, so that, they are linearly independent.

If $n = 2$ Theorem II-2, with some supplementary hypotheses, leads us to a deeper result. As already mentioned in the Introduction the following theorem is due to Professor J. L. Massera with whose permission it is reproduced here.

THEOREM II-3. *Suppose that the system*

$$\dot x = f_1(t) x + g_{11}(t) x + g_{12}(t) y$$

$$\dot{y} = f_2(t)y + g_{21}(t)x + g_{22}(t)y$$

satisfies the following hypotheses:

(1) *The coefficients* $f_i, g_{ij}, T \leqq t < \infty$, *are continuous real-valued functions of the real variable* t.

(2) $f_1(t) \geqq f_2(t)$, $\int_T^\infty (f_1(t) - f_2(t))dt = \infty$, $\int_T^\infty | g_{ij}(t) | dt < \infty$ *for* $i \neq j$ *and* $\int_T^\infty | g_{11}(t) - g_{22}(t) | dt < \infty$.

Then there exists a solution $(x_1(t), y_1(t))$ *satisfying* $\lim_{t\to\infty} x_1(t)/y_1(t) = 0$ *and, for any other solution* $(x(t), y(t))$ *which is not proportional to* $(x_1(t), y_1(t))$, *we have* $\lim_{t\to\infty} y(t)/x(t) = 0$.

Proof. The existence of a solution $(x_1(t), y_1(t))$ with the required property follows from Theorem II-2.

Without loss of generality we may assume $g_{11} = g_{22} = 0$. Choose $t_0 \geqq T$ so large that $\int_{t0}^\infty (| g_{12} | + | g_{21} |)dt < \pi/4$. Let $(x_2(t), y_2(t))$ be the solution which satisfies $x_2(t_0) = 1$, $y_2(t_0) = 0$. Setting $\theta(t) = \arg (x_2(t), y_2(t))$, we claim, in the first place, that $| \theta(t) | < \pi/4$ for $t \geqq t_0$. Assume that this were not the case. It then follows that there exists an interval (t_1, t_2), $t_1 \geqq t_0$, such that $\theta(t_1) = 0$, $| \theta(t_2) | = \pi/4$, $0 < | \theta(t) | < \pi/4$ for $t_1 < t < t_2$, say, $\theta(t_2) = \pi/4$, $0 < \theta(t) < \pi/4$ for $t_1 < t < t_2$, whence $x_2(t) \cdot y_2(t) > 0$ in (t_1, t_2). Since

$$\dot{\theta} = \frac{\dot{y}_2 x_2 - \dot{x}_2 y_2}{x_2^2 + y_2^2} = \frac{g_{21}x_2^2 - g_{12}y_2^2 + (f_2 - f_1)x_2 y_2}{x_2^2 + y_2^2} ,$$

we arrive to the contradiction $\pi/4 = \theta(t_2) - \theta(t_1) = \int_{t_1}^{t_2} \dot{\theta} dt < \pi/4$.

We next prove that $\lim_{t\to\infty} y_2(t)/x_2(t) = 0$, or equivalently $\lim_{t\to\infty} \theta(t) = 0$. There exsists a sequence $t_n \to \infty$ with $\theta(t_n) \to 0$, otherwise $\theta(t) > \theta_0 > 0$, say, which leads to the contradiction

$$\theta(t) - \theta(t_0) \leqq - \int_{t_0}^t (f_1(t) - f_2(t)) \sin \theta(t) \cos \theta(t) dt + \int_{t_0}^t (| g_{12} | + | g_{21} |)dt \to -\infty .$$

Now, given $\varepsilon > 0$, choose t_n such that $| \theta(t_n) | < \varepsilon/2$, $\int_{tn}^\infty (| g_{12} | + | g_{21} |)dt < \varepsilon/2$. An argument similar to the one used to prove $| \theta(t) | < \pi/4$ then shows that $| \theta(t) | < \varepsilon$ for $t \geqq t_n$.

Assume t_0 large enough so that $| x_1(t) |/| y_1(t) | < 1$, $| y_2(t) |/| x_2(t) | < 1$ for $t \geqq t_0$ and, say, $y_1(t) > 0$, $x_2(t) > 0$; then

$$\dot{y}_1(t) \leqq (f_2(t) + | g_{21}(t) |). y_1(t) ,$$

$$\dot{x}_2(t) \geqq (f_1(t) - |g_{12}(t)|).\ x_2(t)\ ,$$

whence

$$y_1(t) \leqq y_1(t_0).\exp\left(\int_{t_0}^{t}(f_2(t) + |g_{21}(t)|)dt\right),$$

$$x_2(t) \geqq x_2(t_0).\exp\left(\int_{t_0}^{t}(f_1(t) - |g_{12}(t)|)dt\right)$$

and

$$\frac{y_1(t)}{x_2(t)} \leqq \frac{y_1(t_0)}{x_2(t_0)}.\exp\left(\int_{t_0}^{t}(f_2(t) - f_1(t) + |g_{12}(t)| + |g_{21}(t)|)dt\right) \to 0\ .$$

Finally, any solution $(x(t), y(t))$ which is not proportional to $(x_1(t), y_1(t))$ satisfies, for a certain constant value k,

$$\frac{y(t)}{x(t)} = \frac{y_2(t) + ky_1(t)}{x_2(t) + ky_2(t)} = \frac{(y_2(t)/x_2(t)) + k(y_1(t)/x_2(t))}{1 + k(x_1(t)/y_1(t))(y_1(t)/x_2(t))} \to 0\ .$$

PART III

Consider the linear differential systems

(III) $$\dot{x} = A(t)x + B(t)x$$

(III′) $$\dot{y} = A(t)y$$

where $A(t)$, $B(t)$, $T \leqq t < \infty$, are continuous complex matrix functions.

Conti [2, Theorem I, p. 589] proved that: *if* $\int^{\infty}|B(t)|\,dt < \infty$ *where* $B(t) = (b_j^i(t))$ *and* $|B(t)| = \sum_{i,j}|b_j^i(t)|$ *and if (III′) is uniformly stable, then the system (III) and (III′) are asymptotically equivalent*[3].

The theorem of Wintner [7, 7-i, p. 423] stating that:

If $B(t) = (b_j^i(t))$, $T \leqq t < \infty$, $i, j = 1, \cdots, n$, *is a matrix of* n^2 *continuous functions satisfying* $\int^{\infty}|B(t)|\,dt < \infty$, *then every solution of* $\dot{x} = B(t)x$ *tends to a finite limit as* $t \to \infty$, is a particular case of Conti's result $(A(t) = 0)$.

Our Theorem III-3, is also a generalization of Wintner's theorem but different from that of Conti.

Theorems III-1 and III-2, which are preliminary to Theorem III-3, give us some information, though less than asymptotic equivalence, concerning the behavior of two systems, one of which not necessarily linear.

THEOREM III-1. *Suppose that the systems*

[3] The theorem of Conti is actually more general. We have considered the theorem applied to linear systems only.

$$\text{(III-1)} \qquad \dot{x}_i = f_i(t)x_i + g_i(t, x) \ ,$$

$$\text{(III-2)} \qquad \dot{y}_i = f_i(t)y_i \ ,$$

$$i = 1, \cdots, n \ , \qquad g(t, x) = (g_i(t, x))$$

satisfy the following hypotheses:

(1) *$f_i(t)$, $T \leqq t < \infty$, are continuous functions (in general complex-valued) of the real variable t; $g_i(t, x)$ are functions (in general complex-valued) continuous in*

$$\Omega = \{(t, x) \mid t > T, \mid x \mid = \sum_{i=1}^{n} \mid x_i \mid < \infty\}$$

and satisfy some condition which implies the existence of only one integral passing through each point of Ω.

(2) *$\mid g(t, x) \mid \leqq \mid x \mid F(t)$ on Ω.*

(3) *There exists a negative constant K such that*

$$K \leqq \int_t^v R(f_i(\tau))d\tau$$

for all $v \geqq t > T$ and

$$\int^{\infty} F(t) \exp\left[\int_T^t R(f_i(s))ds\right]dt < \infty$$

for all $i = 1, \cdots, n$.

Then for every solution $y(t)$ of III-2 there is a solution $x(t)$ of III-1 such that $\lim_{t\to\infty}[x(t) - y(t)] = 0$.

Proof. We define $\omega = \{P \in \Omega \mid \mid x_i - y_i(t) \mid < \varphi_i(t), t > t_0 \geqq T\}$ where the $\varphi_i(t)$ and t_0 will be adequately chosen so that for all $t \geqq t_0$, $i = 1, \cdots, n$, we have: $\varphi_i(t) > 0$, φ_i differentiable, $\lim_{t\to\infty}\varphi_i(t) = 0$ and ω a generalized regular polyfacial set.

If we put

$$H_i(P) = \mid x_i - y_i(t) \mid^2 - \varphi_i^2(t) \ , \qquad i = 1, \cdots, n$$
$$H_{n+1}(P) = t_0 - t$$

it follows that $\omega = \{P \mid H_i(P) < 0 \ , i = 1, \cdots, n + 1\}$.

For all $i, 1 \leqq i \leqq n$,

$$\Gamma_i = \{P \in \Omega \mid\mid x_i - y_i(t) \mid = \varphi_i(t), \mid x_j - y_j(t) \mid \leqq \varphi_j(t), j = 1, \cdots, n, t \geqq t_0\} \ .$$

An easy computation shows that

$$\frac{1}{2}[D_{(\mathrm{III}-1)}H_i(P)]P \in \Gamma_i \geqq \varphi_i^2(t)R[f_i(t)] - \varphi_i(t)\dot{\varphi}_i(t)$$

$$- \varphi_i(t)F(t)\Big[\sum_{k=1}^{n} | x_i - y_k | + | y_k |\Big] .$$

As we want $\varphi_i(t) > 0$ and $\lim_{t\to\infty}\varphi_i(t) = 0$, we can take t_0 such that $\sum_{i=1}^{n}\varphi_i(t) < 1$ for all $t \geqq t_0$. Then,

$$\frac{1}{2}\Big[D_{(\mathrm{III}-1)}H_i(P)\Big]P \in \Gamma_i \geqq \varphi_i^2(t)R[f_i(t)] - \varphi_i(t)\dot{\varphi}_i(t) - \varphi_i(t)F(t)$$

$$- \varphi_i(t)F(t) \sum_{k=1}^{n} | c_k | \exp \int_{t_0}^{t} R[f_k(s)]ds \geqq \varphi_i^2(t)R[f_i(t)] - \varphi_i(t)\dot{\varphi}_i(t)$$

$$- \varphi_i(t)F(t) \sum_{k=1}^{n} | d_k | \exp \int_{t_0}^{t} R[f_k(s)]ds \geqq \varphi_i^2(t)R[f_i(t)] - \varphi_i(t)\dot{\varphi}_i(t)$$

$$- \varphi_i(t)h(t)$$

where we can assume $\int^{\infty} h(t)dt < \infty$ and, without loss of generality, $h(t) > 0$ for all $t \geqq t_0$.

In order to have, for all $i = 1, \cdots, n$, $[D_{(\mathrm{III}-1)}H_i(P)]P \in \Gamma_i > 0$ it is sufficient to choose $\varphi(t)$ such that

$$- \dot{\varphi}_i(t) + R[f_i(t)]\varphi_i(t) - h(t) > 0 .$$

The problem is then to look for a solution $z(t)$ of $\dot{z} < \sigma(t)z - \gamma(t)$ satisfying $z(t) > 0$ for all $t \geqq t_0$, $\lim_{t\to\infty} z(t) = 0$, knowing that $\gamma(t) > 0$ for all $t \geqq t_0$, $\int^{\infty}\gamma(t)dt < \infty$ and $\int_t^v \sigma(s)ds \geqq K$ for some constant K and all $v \geqq t \geqq t_0$. If $W(t)$ is a solution of $\dot{W} = \sigma(t)W - \gamma(t)$ it follows that $z(t) = 2W(t)$ is a solution of $\dot{z} < \sigma(t)z - \gamma(t)$. It is then sufficient to find a solution $W(t)$ satisfying $W(t) > 0$ for all $t \geqq t_0$ and $\lim_{t\to\infty} W(t) = 0$. The solution $W(t) = \exp\left(\int_{t_0}^{t}\sigma(s)ds\right)$. $\int_t^{\infty}\gamma(v)\exp\left(-\int_{t_0}^{v}\sigma(s)ds\right)dv$ exists and indeed $W(t) \to 0$ as $t \to \infty$ because

$$W(t) = \int_t^{\infty}\gamma(v)\exp\left(-\int_t^{v}\sigma(s)ds\right)dv \leqq e^{-K}\int_t^{\infty}\gamma(v)dv .$$

Since $[D_{(\mathrm{III}-1)}H_{n+1}(P)] = -1$ it follows that ω is a generalized regular polyfacial set and $S = S^* = \bigcup_{i=1}^{n}\Gamma_i - \Gamma_{n+1}$.

If we choose

$$Z = \{(t, x) \mid t = \tau > t_0, | x_j - y_j(\tau) | \leqq \varphi_j(\tau), j = 1, \cdots, n\}$$

it follows that $S \cap Z = \bigcup_{i=1}^{n} \Gamma_i \cap Z - \Gamma_{n+1}$

$$\Gamma_i \cap Z = \{t, x) \mid t = \tau, | x_i - y_i(\tau) | = \varphi_i(\tau), | x_j - y_j(\tau) | \leqq \varphi_j(\tau), j = 1, \cdots, n\} .$$

Then $Z = \prod_{j=1}^{n} B_j^2$

$$Z \cap S = \bigcup_{j=1}^{n} B_1^2 \times \cdots \times B_{j-1}^2 \times S_j^1 \times B_{j+1}^2 \times \cdots \times B_n^2$$

and, modulo homomorphisms, $Z = B^{2n}$, $Z \cap S = S^{2n-1}$ so that $Z \cap S$ is not a retract of Z. However, it is easily seen that $\phi : S \to S \cap Z$ given by $\phi(P) = P^*$, with $t^* = \tau$, $x_i^* = y_i(\tau) + [x_i - y_i(t)]\varphi_i(\tau)/\varphi_i(t)$, is a retraction.

Using the theorem of Ważewski we can conclude the existence of at least one point $P_0 : (\tau, x_0) \in Z - S$ such that $(t, x(t, P_0)) = I(t, P_0) \subset \omega$ for all $t \geqq t_0$.

Since $x(t, P_0)$ is defined in the future, i.e., $\beta(P_0) = \infty$ (because $\beta(P_0) < \infty$ implies $\{I(t, P_0) \mid t_0 \leqq t < \beta(P_0)\}$ bounded, which is not possible), it follows that $\lim_{t\to\infty}[x(t, P_0) - y(t)] = 0$.

COROLLARY 1. *Suppose that the systems*

$$\text{(III-1')} \qquad \dot{x}_i = f_i(t)x_i + \sum_{j=1}^{n} g_{ij}(t)x_j$$

$$\text{(III-2')} \qquad \dot{y}_i = f_i(t)y_i$$

$$i = 1, \cdots, n , \quad g(t) = (g_{ij}(t))$$

satisfy the following hypotheses:

(1) *The coefficients f_i, g_{ij}, $T \leqq t < \infty$, are continuous functions (in general complex-valued) of the real variable t.*

(2) *There exists a constant K such that*

$$K \leqq \int_t^v R[f_i(s)]ds \text{ for all } v \geqq t \geqq T \text{ and}$$

$$\int^{\infty} | g(t) | \exp\left\{\int_T^t R[f_i(s)]ds\right\}dt < \infty , \qquad i = 1, \cdots, n .$$

Then for every solution $y(t)$ of (III-2') there exists a solution $x(t)$ of (III-1') such that $\lim_{t\to\infty}[x(t) - y(t)] = 0$.

The theorem of Wintner mentioned before follows a once from Corollary 1.

THEOREM III-2. *Suppose that the systems*

$$\text{(III-A)} \qquad \dot{x}_i = \sum_{j=1}^{n} f_{ij}(t)x_j + g_i(t, x)$$

$$\text{(III-B)} \qquad \dot{y}_i = \sum_{j=1}^{n} f_{ij}(t)y_j$$

$$i, j = 1, \cdots, n , \quad g(t, x) = (g_j(t, x))$$

satisfy the following hypotheses:

(1) *$f_{ij}(t)$, $T \leqq t < \infty$, are continuous functions (in general complex-valued) of real variable t; $g_i(t, x)$ are functions (in general complex-valued) continuous in*

$$\Omega = \{(t, x) \mid t < T, \mid x \mid < \infty\}$$

and satisfy some condition which implies the existence of only one integral passing through each point of Ω.

(2) $\mid g(t, x) \mid \leqq \mid x \mid F(t)$ *in Ω.*

(3) *There exists a constant K such that*

$$K \leqq \int_t^v R[f_{ii}(s)]ds \text{ for all } v \geqq t \geqq T \text{ and}$$

$$\int^\infty F(t) \exp\left\{\int_T^t R[f_{ii}(s)]ds\right\}dt < \infty\ , \qquad i = 1, \cdots, n$$

$$\int^\infty \mid f_{ij}(t) \mid \exp\left\{\int_T^t R[f_{kk}(s)]ds\right\}dt < \infty\ , \qquad k = 1, \cdots, n, i \neq j\ .$$

Then for every solution $y(t)$ of (III-B) there is a solution $x(t)$ of (III-A) such that $\lim_{t\to\infty}[x(t) - y(t)] = 0$

Proof. Consider the systems

(III-A) $\quad \dot{x}_i = f_{ii}(t)x_i + \tilde{g}_i(t, x)$ where $\tilde{g}_i(t, x) = g_i(t, x) + \sum_{j\neq i} f_{ij}(t)x_j$

(III-C) $\quad \dot{z}_i = f_{ii}(t)z_i\ .$

These systems satisfy the condition of Theorem III-1. Hence for every solution $z(t)$ of (III-C) there is a solution $x(t)$ of

(III-A) $\qquad$ such that $\lim_{t\to\infty}[z(t) - x(t)] = 0$

Consider now the systems

(III-B) $$\dot{y}_i = \sum_{j=1}^n f_{ij}(t)y_j$$

(III-C) $$\dot{z}_i = f_{ii}(t)z_i\ .$$

It is easy to see that they also satisfy the hypotheses of Theorem III-1. Hence for every solution $z(t)$ of (III-C) there is a solution $y(t)$ of (III-B) such that $\lim_{t\to\infty}[y(t) - z(t)] = 0$. But we can also prove that for every solution $y(t)$ of (III-B) there is a solution $z(t)$ of (III-C) such that $y(t) - z(t) \to 0$ as $t \to \infty$. For that purpose it is enough to show that there is a fundamental set $z^1(t), \cdots, z^n(t)$ of solutions of (III-C) such that the solutions $y^1(t), \cdots, y^n(t)$ satisfying $y^i(t) - z^i(t) \to 0$ as $t \to \infty$, for all $i = 1, \cdots, n$, are a fundamental set of solutions of (III B).

Let us take $z^i(t) = \begin{pmatrix} z_1^i(t) \\ \vdots \\ z_n^i(t) \end{pmatrix}$ such that $z_j^i(t) = 0$ for all $j \neq i$ and $z_j^i(t) = \exp\left(\int_T^t f_{ii}(s)ds\right)$ for all $i = 1, \cdots, n$.

The corresponding $y^i(t)$, $i = 1, \cdots, n$, satisfy $\lim_{t\to\infty} y_j^i(t) = 0$ if $j \neq i$ and $\lim_{t\to\infty} | y_1^i(t) - \exp\left[\int_T^t f_{ii}(s)ds \right| = 0$. Hence, there exists t_0 such that $t \geqq t_0$ implies

$$| y_i^i(t) - \exp\left[\int_T^t f_{ii}(s)ds\right] | < \frac{1}{2}e^K .$$

Whence

$$| y_i^i(t) | > \exp\left\{\int_T^t R[f_{ii}(s)]ds\right\} - \frac{1}{2}e^K \geqq \frac{1}{2}e^K .$$

Therefore, for any $\varepsilon > 0$ there is a $t(\varepsilon)$ such that $t \geqq t(\varepsilon)$ implies $| y_i^i(t) | > 1/2\, e^K$, $i = 1, \cdots, n$, and $| y_j^i(t) | < \varepsilon$ for all $i \neq j$. This implies the existence of a $\tau \geqq T$ with $\det(y^1(\tau), \cdots, y^n(\tau)) \neq 0$ and $(y^1(t), \cdots, y^n(t))$ is a fundamental set of solutions of (III-B).

From the results concerning the systems (III-A), (III-C) and (III-B), (III-C) we conclude that for every solution $y(t)$ of (III-B) there is a solution $x(t)$ of (III-A) such that $\lim_{t\to\infty}[x(t) - y(t)] = 0$.

THEOREM III-3. *Suppose that the systems*

$$\text{(III-}\alpha\text{)} \qquad \dot{x}_i = \sum_{j=1}^n f_{ij}(t)x_j(t)x_j + \sum_{j=1}^n g_{ij}(t)x_j$$

$$\text{(III-}\beta\text{)} \qquad \dot{y}_i = \sum_{j=1}^n f_{ij}(t)y_j \qquad i, j = 1, \cdots, n$$

satisfy the following hypotheses:

(1) *The coefficients* f_{ij}, g_{ij}, $T \leqq t < \infty$, *are continuous functions (in general complex-valued) of the real variable* t.

(2) *There exists constant* K *such that* $K \leqq \int_t^v R[f_{ii}(s)]ds$ *for all* $v \geqq t \geqq T$, $i = 1, \cdots, n$, *and*

$$\int^\infty | g_{ij}(t) | \exp\left\{\int_T^t R[f_{kk}(s)]ds\right\}dt < \infty , \qquad i, j, k = 1, \cdots, n$$

$$\int^\infty | f_{ij}(t) | \exp\left\{\int_T^t R[f_{kk}(s)]ds\right\}dt < \infty , \qquad i, j, k = 1, \cdots, n, i \neq j .$$

Then the systems (III-α*) and (III-*β*) are asymptotically equivalent.*

Proof. By Theorem III-2 for every solution $y(t)$ of (III-β) there is a solution $x(t)$ of (III-α) such that $\lim_{t\to\infty}[x(t) - y(t)] = 0$.

Let us show that given a fundamental set $(y^1(t), \cdots, y^n(t))$ of solutions of (III-β) the corresponding solutions $(x^1(t), \cdots, x^n(t))$ of (III-α) satisfying $\lim_{t\to\infty}[x^i(t) - y^i(t)] = 0, i = 1, \cdots, n$, also form a fundamental set of solutions.

Consider the auxiliary system

$$\text{(III-}\gamma) \qquad \dot{z}_i = f_{ii}(t)z_i \, , \qquad i = 1, \cdots, n \, .$$

Applying the argument used in Theorem III-2 to the systems (III-β), (III-γ) we conclude that there exists a fundamental set $(y^1(t), \cdots, y^n(t))$ of solutions of (III-β) and a t_0 such that $t \geqq t_0$ implies

$$| y^i_i(t) | \geqq \frac{1}{2}e^K \text{ and } y^i_j(t) \to 0 \text{ as } t \to \infty \text{ for all } i \neq j \, .$$

Let $(x^1(t), \cdots, x^n(t))$ be the solutions of (III-α) such that $\lim_{t\to\infty} [x^i(t) - y^i(t)] = 0$ (the existence of which follows from Theorem III-2). Then $\lim_{t\to\infty}x^i_j(t) = 0$ for all $i \neq j$ and there exists $\tau \geqq t_0$ such that $t \geqq \tau$ implies $| x^i_i(t) | > 1/4e^K$.

For sufficiently large t we have therefore

$$\det (x^1(t), \cdots, x^n(t)) \neq 0$$

and this means that $(x^1(t), \cdots, x^n(t))$ is a fundamental set of solutions of (III-α).

The systems (III-α) and (III-β) being linear this implies that they are asymptotically equivalent.

BIBLIOGRAPHY

1. I. Barbălat, *Applications du principe topologique de T. Wazewski aux équations différentielles du second ordre*, Ann. Polon. Math., **3** (1958-1959), 303-317.

2. R. Conti, *Sistemi differenziali asintoticamente equivalenti*, Rendiconti dell'Accademia Nazionale dei Lincei. ser. VIII, vol. XXII, fasc. 5, (1957), 588-592.

3. J. Lewowics, *Sobre um teorema de Szmdtówna*. Publicaciones del Instituto de Matemática y Estadística de la Facultad de Ingenieria y Agrimensura, Montevideo, Uruguay, III, (1960), 125-134.

4. G. R. Sansone, e Conti, *Equazioni differenziali non lineari*, Consiglio Nazionale delle Ricerche, Monografie Matematiche, 3 Edizioni Cremonese, Roma, (1956).

5. Z. Szmydtówna, *Sur l'allure asymptotique des intégrales des éqations différentielles ordinaires*, Ann, Soc. Polon. Math. Tath. Tome 24, Fasc. II, (1951), 17-34.

6. T. Wazewski, *Sur un principe topologique de l'examen de l'allure asymptotique des intégrales des équations différentielles ordinaires*, Ann. Soc. Polon. Math. Tome 20, (1947), 279-313.

7. A. Wintner, *Small pertubations*, Amer. J. Math., **67**, (1945), 417-430.

FACULDADE DE FILOSOFIA, CIÊNCIAS E LETRAS
RIO CLARO—SÃO PAULO—BRASIL
RIAS

A THEOREM ON REGULAR MATRICES

PETER PERKINS

In this paper it will be proved that if any nonnegative, square matrix P of order r is such that $P^m > 0$ for some positive integer m, then $P^{r^2-2r+2} > 0$. This result has already appeared in the literature, [2], but the following is a complete and elementary proof given in detail except for one theorem of I. Schur in [1] which is stated without proof. The term regular is taken from Markov chain theory[1] in which a regular chain is one whose transition matrix has the above property.

A *graph* G_P associated with any nonnegative, square matrix P of order r is a collection of r distinct points $S = \{s_1, s_2, \cdots, s_r\}$, some or all of which are connected by directed lines. There is a directed line (indicated pictorially by an arrow) from s_i to s_j in the graph G_P if and only if $p_{ij} > 0$ in the matrix $P = (p_{ij})$. A *path sequence* or *path* in G_P is any finite sequence of points of S (not necessarily distinct) such that there is a directed line in G_P from every point in the sequence to its immediate successor. The *length* of a path is one less than the number of occurrences of points in its sequence. A *cycle* is any path that begins and ends with the same point and a *simple cycle* is a cycle in which no point occurs twice except, of course, for the first (and last). Two cycles are *distinct* if their sequences are not cyclic permutations of each other. A nonnegative, square matrix P is *regular* if $P^m > 0$ for some positive integer m. Likewise, a graph G_P associated with a nonnegative square matrix P is *regular* if there exists a positive integer m such that an infinite set of paths $A_0, A_1, \cdots, A_n, \cdots$ can be found, the length of each path being $L_n = m + n$, $n = 0, 1, 2, \cdots$. The usual notation $p_{ij}^{(m)}$ is used to denote the ijth entry of the matrix P^m. In all that follows we shall consider only regular matrices P and their associated graphs G_P.

Some immediate consequences of these definitions and the definition of matrix multiplication are the following:

(1) There is a path $s_{k_1} \cdots s_{k_{m+1}}$ in G_P if and only if $p_{k_1 k_{m+1}}^{(m)} > 0$ in P^m.
(2) P is regular if and only if G_P is regular.
(3) There exists some path from any point in G_P to any point in G_P.
(4) For any given i and j there exists some m such that $p_{ij}^{(m)} > 0$.
(5) If $P^m > 0$ then $P^{m+n} > 0$, $n = 0, 1, 2, \cdots$.

Let $C = \{C_1, C_2, \cdots, C_t\}$ be all the distinct simple cycles of G_P and $\{c_1, c_2, \cdots, c_t\}$ be the corresponding lengths.

Received November 21, 1960. I wish to thank Professor R. Z. Norman for his suggestions in the writing of this paper.

[1] This is as treated by Kemeny and Snell in [3].

LEMMA 1. *The length of any cycle C^* is always of the form $c^* = \sum_{i=1}^{t} a_i c_i$, where a_i is some nonnegative integer.*

Proof. Let any cycle $C^* = s_{k_1}, s_{k_2}, \cdots, s_{k_m}$ be given ($k_1 = k_m$). Let $C^* = C_1^*$ and form C_{i+1}^* in the following manner from C_i^*: Wherever simple cycle C_i occurs in cycle C_i^* delete it except for its last point, thus forming the new cycle C_{i+1}^*. It is clear that after the tth step there will remain only a single point of the original C^*, which has of course zero length. If we let a_i be the number of times simple cycle C_i occurred in cycle C_i^* then the lemma follows.

THEOREM 1. *If G_P is any regular graph then it must contain a set of simple cycles whose lengths are relatively prime.*

Proof. By the regularity assumption and (1) there exists a positive integer m such that cycles of lengths $L_n = m + n$, $n = 0, 1, 2, \cdots$ can be found in G_P. Also, from Lemma 1, $L_n = \sum_{i=1}^{t} a_i c_i$ for $n = 0, 1, 2, \cdots$, and suitable a_i. Let d be the common factor of the simple cycle lengths c_i. Then

$$\sum_{1=1}^{t} a_i c_i = d \sum_{i=1}^{t} a_i c_i'$$

which could never equal $m + n$, $n = 0, 1, 2, \cdots$ unless $d = 1$.

We would like to find a *least* integer M such that for arbitrary points s_i and s_j there are paths beginning at s_i and ending at s_j and whose lengths are $L_n = M + n$, $n = 0, 1, 2, \cdots$. If we can do this, then, by (1), we shall have also found a least integer M such that $P^M > 0$ where P is the regular matrix associated with G_P.

Let us say that a path *touches* a given set of points if there is some point belonging to both the path and the set. Then we have

LEMMA 2. *Let G_P be a regular graph with r points, let S be a subset containing r_k distinct points of the graph, and let g be any point of G_P. Then there always exists a path from g which touches S whose length is less than or equal to $r - r_k$.*

Proof. If $g \in S$ then the lemma is trivial. Suppose $g \notin S$. By (3) there is at least one path which starts at g and touches the set S. Let $p = g_0, g_1, \cdots, s$ be such a path of shortest length. Obviously no point of S can precede the final point s in this path sequence p. Furthermore, there can be no repeated points in p, for the deletion of any cycle (except for its last point) would produce a path from g to S shorter than path p, contrary to the choice of p. Therefore, p can have at most $r - r_k$ points.

We shall say that a *minimal set* of relatively prime integers is a set of relatively prime integers such that if one of the integers is deleted the remaining integers are no longer relatively prime. A *step* along a path in G_P is a pair of consecutive points of the path sequence.

THEOREM 2. *If $R = \{R_1, R_2, \cdots, R_k\}$ is a set of simple cycles of graph G_P whose lengths $\{r_1, r_2, \cdots, r_k\}$ form a minimal set of relatively prime integers and if s_i and s_j are arbitrary points of G_P, then there is always a path which starts at s_i, ends at s_j, touches each cycle of R and whose length $L \leqq (k+1)r - \sum_{i=1}^{k} r_i - 1$.*

Proof. Note that the set of distinct points belonging to a simple cycle contains a number of points exactly equal to the length of the cycle. Hence, by Lemma 2 there is a path from an arbitrary point s_i which touches a particular cycle R_p and whose length is less than or equal to $r - r_p$. Thus, we have the following:

from		*to*		*greatest number of steps needed*
arb. pt.	s_i	cycle	R_1	$r - r_1$
cycle	R_1	〃	R_2	$r - r_2$
.		.		.
.		.		.
.		.		.
cycle	R_{k-1}	cycle	R_k	$r - r_k$
〃	R_k	arb. pt.	s_j	$r - 1$
		TOTAL		$L \leqq (k+1)r - \sum_{i=1}^{k} r_i - 1.$

We shall now state without proof I. Schur's theorem cited above and use it in our final theorem.

THEOREM 3. (Schur) *If $\{a_1, a_2, \cdots, a_n\}$ is a set of relatively prime integers with a_1 the least and a_n the greatest, then $B = \sum_{i=1}^{n} x_i a_i$ has solutions in nonnegative integers x_i for any $B \geqq (a_1 - 1)(a_n - 1)$. This is a best bound for $n = 2$.*

THEOREM 4. *If M is the least integer such that paths between any two points of G_P can be found whose lengths are $L_n = M + n$, $n = 0, 1, 2, \cdots$, then $M \leqq r^2 - 2r + 2$.*

Proof. Given any two points s_i and s_j of G_P we know by Theorem 2 that there is a path from s_i to s_j touching each of the cycles $\{R_1, R_2, \cdots, R_k\}$ and whose length is

$$L \leqq (k+1)r - \sum_{i=1}^{k} r_i - 1 .$$

We can, then, interject into this path the simple cycles $\{R_1, R_2, \cdots, R_k\}$ at the touching points, interjecting cycle R_i say x_i times. The length L of the original path has now been increased to $L + \sum_{i=1}^{k} x_i r_i = L + B$, the second part of which, by Schur's theorem, can be made to take on any integral value B where $B \geqq (r_s - 1)(r_g - 1)$, and $r_s = \min(r_1, r_2, \cdots, r_k)$, $r_g = \max(r_1, r_2, \cdots, r_k)$. Therefore, we have:

$$M \leqq L + B = (k+1)r - \sum_{i=1}^{k} r_i - r_s - r_g + r_s r_g \tag{7}$$

Case I. Suppose $k = 2$. Then $M \leqq 3r - (r_s + r_g) - r_s - r_g + r_s r_g = 3r - 2r_s - 2r_g + r_s r_g = 3r + (r_g - 2)(r_s - 2) - 4$. The right side of this inequality is obviously maximum when r_s and r_g are as large as possible. Recall that $r_g \leqq r$ and $r_s \leqq r - 1$. Therefore we have:

$$M \leqq 3r + (r-2)(r-3) - 4 = r^2 - 2r + 2\ . \tag{8}$$

Case II. Suppose $k \geqq 3$. The reader may wish to skip the following formidable looking, though straightforward calculations. They result in a proof that the integer M with the desired property is in fact smaller when the arbitrary graph contains a larger set of these cycles.

Since the lengths of these cycles are a minimal set of relatively prime integers, it is certainly true that

$$\begin{aligned}\sum_{i=1}^{k} r_i &\geqq r_s + [r_s + 2] + [r_s + 4] + \cdots + [r_s + 2(k-2)] + r_g \\ &= (k-1)r_s + (k-1)(k-2) + r_g\ .\end{aligned}$$

Thus, with (7) we have:

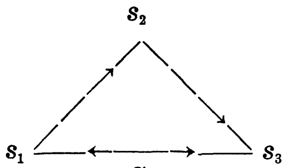
$$\begin{aligned}M &\leqq (k+1)r - [(k-1)r_s + (k-1)(k-2) + r_g] - r_s - r_g + r_s r_g \\ &= (k+1)r - kr_s - 2r_g + r_s r_g - (k-1)(k-2) \\ &= (k+1)r + (r_s - 2)(r_g - k) - 2k - (k-1)(k-2)\ .\end{aligned}$$

Since r_g must be larger than k, the right side again is maximum when r_g and r_s are as large as possible. But $r_g \leqq$ and $r_s \leqq r - k + 2$. So

$$\begin{aligned}M &\leqq (k+1)r + (r-k)(r-k) - k^2 + k - 2 \\ &= r^2 + (1-k)r + k - 2\ .\end{aligned}$$

This is easily seen to be less than $r^2 - 2r + 2$ of Case I, if $r > 1$. So in any case $M \leqq r^2 - 2r + 2$.

To see that $r^2 - 2r + 2$ is the least value for an arbitrary graph of r points and thus for an arbitrary matrix of order r, we need only consider the following example in which $r = 3$ and $M = 5$.



G_P

$$\begin{array}{c} \\ s_1 \\ s_2 \\ s_3 \end{array}\!\!\begin{array}{c} \begin{array}{ccc} s_1 & s_2 & s_3 \end{array} \\ \begin{pmatrix} 0 & \times & \times \\ 0 & 0 & \times \\ \times & 0 & 0 \end{pmatrix} \end{array} \quad \begin{array}{c} \\ s_1 \\ s_2 \\ s_3 \end{array}\!\!\begin{array}{c} \begin{array}{ccc} s_1 & s_2 & s_3 \end{array} \\ \begin{pmatrix} \times & 0 & \times \\ \times & 0 & 0 \\ 0 & \times & \times \end{pmatrix} \end{array} \quad \begin{array}{c} \\ s_1 \\ s_2 \\ s_3 \end{array}\!\!\begin{array}{c} \begin{array}{ccc} s_1 & s_2 & s_3 \end{array} \\ \begin{pmatrix} \times & \times & \times \\ \times & 0 & \times \\ \times & \times & \times \end{pmatrix} \end{array}$$

$$P \qquad\qquad\qquad P^2 \qquad\qquad\qquad P^4$$

As a matter of fact it can be shown for *any* regular matrix P of order r whose graph G_P contains only two cycles, one of length r and one of length $r - 1$, that P^{r^2-2r+1} is not positive. We have, therefore, established the claim of the paper as stated in the opening paragraph.

BIBLIOGRAPHY

1. Alfred Brauer, *On a problem of partitions*, Amer. J. Math. **LXIV**, (1942), 299-312.

2. J. C. Holladay, and R. S. Varga, *On powers of nonnegative matrices*, Proc. Amer. Math. Soc. Vol. IX, no. 4, (August, 1958), 631.

3. John G. Kemeny and J. Laurie Snell, *Finite Markov Chains*, New York, D. Van Nostrand Company, Incorporated, 1960.

4. David Rosenblatt, *On the Graphs and Asymptotic Forms of Finite Boolean Relation Matrices and Stochastic Matrices*, Naval Research Logistics Quarterly, Vol. IV, no. 2, (June, 1957), 151-167.

DARTMOUTH COLLEGE

CENTROID SURFACES

C. M. PETTY

1. Introduction. Let $M_1, \cdots, M_{n-1}$ denote $(n-1)$ bounded closed sets in E_n. Busemann [1] has established the expression

$$(1.1)\quad |M_1| \cdots |M_{n-1}| = \frac{(n-1)!}{2}\int_{\Omega_n}\left(\int_{M_1(u)} \cdots \int_{M_{n-1}(u)} T(z, p_1, \cdots, p_{n-1}) dV_{p_1}^{n-1} \cdots dV_{p_{n-1}}^{n-1}\right) d\omega_u^n$$

where $|M_i|$ is the n-dimensional Lebesgue measure or volume of M_i. On the righthand side $M_i(u)$ is the cross-section of M_i with the hyperplane through z normal to the unit vector u, the point p_i varies in $M_i(u)$ and the differential $dV_{p_i}^{n-1}$ is the $(n-1)$-dimensional volume element of $M_i(u)$ at p_i. The final integration is extended over the surface Ω_n of the solid-unit sphere U_n and $d\omega_u^n$ is the area element of Ω_n at point u. By $T(z, p_1, \cdots, p_r)$ we will denote the r-dimensional volume of the simplex (possibly degenerate) with vertices $z, p_1, \cdots, p_r$.

Let

$$(1.2)\qquad \pi_r = \frac{\pi^{r/2}}{\Gamma(r/2+1)}\,.$$

For $n \geqq 3$, Busemann also shows by Steiner's symmetrization that

$$(1.3)\quad |M_1| \cdots |M_{n-1}| \geqq \frac{1}{n}\frac{\pi_n^{n-2}}{\pi_{n-1}^n}\int_{\Omega_n} |M_1(u)|^{n/(n-1)} \cdots |M_{n-1}(u)|^{n/(n-1)} d\omega_u^n$$

for nondegenerate convex bodies M_i where the equality sign holds only when the M_i are homothetic solid ellipsoids with center z. Here $|M_i(u)|$, of course, denotes the $(n-1)$-dimensional volume of $M_i(u)$. In this regard we will also, as a matter of convenience, not index lower dimensional mixed discriminates and mixed volumes since the dimension will be evident from the number of components.

The primary purpose of this note is to reinterpret (1.1) as an integration of the type (1.3) retaining the equality sign. This is given in § 3 by (3.20). In addition other integral expressions and inequalities are derived which are geometrically of the same type as those considered above.

2. Fenchel's momental ellipsoid. Let M be a bounded closed set with positive volume. The centroid s of M is defined by its rectangular coordinates

Received January 3, 1961.

$$s_i = \frac{1}{|M|}\int_M x_i dV_x^n . \tag{2.1}$$

If L_ν is a ν-flat through the origin z, then the second moment of M with respect to $L_\nu (0 \leqq \nu \leqq n-1)$ is defined by

$$I(M, L_\nu) = \int_M r^2 \sin^2 \varphi dV_x^n \tag{2.2}$$

where the distance zx is r and φ is the angle between the ray zx and L_ν (for $\nu = 0$, we define $\varphi = \pi/2$). By the same type of integration technique in [1, pp. 5–6], the reader may verify that

$$I(U_n, L_\nu) = \frac{n-\nu}{n+2}\pi_n \tag{2.3}$$

where U_n has center z; a calculation which will be used later.

The matrix A_M given by

$$A_M = \left[\frac{1}{|M|}\int_M x_i x_j dV_x^n\right] \tag{2.4}$$

is positive definite since

$$y^T A_M y = \frac{1}{|M|}\int_M (\Sigma x_i y_i)^2 dV_x^n ,$$

where y is a column vector and y^T is its transpose. The ellipsoid with surface $x^T A_M x = 1$ will be called Fenchel's momental ellipsoid and its polar reciprocal with respect to Ω_n given by $x^T A_M^{-1} x = 1$ will be called simply Fenchel's ellipsoid. This name is chosen since *W.* Fenchel first observed the affine character of this polar reciprocal (unpublished):

(2.5) Let M be transformed into $\bar{M}$ by a central affinity with matrix B. If F and $\bar{F}$ are the Fenchel ellipsoids of M and $\bar{M}$ respectively, then this central affinity also carries F into $\bar{F}$.

To see this, it may be observed from (2.4) that $A_{\bar{M}} = BA_M B^T$ or $A_{\bar{M}}^{-1} = (B^{-1})^T A_M^{-1} B^{-1}$ which completes the proof.

If $|F|$ is the volume of the Fenchel ellipsoid F of M, then

$$|F|^2 = \pi_n^2 \det(A_M) . \tag{2.6}$$

The result (2.5) enables one to prove readily that

$$\pi_n^{-2}|F|^2 = \det(A_M) \geqq (n+2)^{-n}\pi_n^{-2}|M|^2 \tag{2.7}$$

with equality only if, except for a set of measure zero, M is a solid ellipsoid with center z. For if we transform M into $\bar{M}$ by a unimodular central affinity so that $\bar{F}$ is a sphere, then

$$\det(A_M) = \left[\frac{1}{n\,|M|}\int_{\bar{M}} r^2 d V_x^n\right]^n .$$

Comparison of $\int_{\bar{M}} r^2 d V_x^n$ with that for a sphere with center z and volume $|M|$ proves (2.7).

We will adopt the same notation for mixed discriminates as in [2, pp. 51–57] where the reader will find an exposition of their properties. Consider the r quadratic forms $q_i = x^T A_i x$, $i = 1, \cdots, r$, where $A_k = [a_{ij}^{(k)}]$ is a real symmetric matrix. For any real $\lambda_1, \cdots, \lambda_r$, set $q = \lambda_1 q_1 + \cdots + \lambda_r q_r = x^T A x$ where $A = \sum_{k=1}^{r} \lambda_k A_k$. The discriminant $D(q) = \det(A)$ can be written

$$D(q) = \sum_{i_1=1}^{r} \cdots \sum_{i_n=1}^{r} \lambda_{i_1} \cdots \lambda_{i_n} D(q_{i_1}, \cdots, q_{i_n})$$

where $D(q_{i_1}, \cdots, q_{i_n})$ is independent of the order of the q_{i_k} and is called the mixed discriminant of $q_{i_1}, \cdots, q_{i_n}$. For n forms q_i we have

$$D(q_1, \cdots, q_n) = \frac{1}{n!}\sum_{(i_1 \cdots i_n)} \begin{vmatrix} a_{11}^{(i_1)} & \cdots & a_{1n}^{(i_n)} \\ \vdots & & \vdots \\ a_{n1}^{(i_1)} & \cdots & a_{nn}^{(i_n)} \end{vmatrix} \tag{2.8}$$

where $(i_1 \cdots i_n)$ is a permutation of $(1 \cdots n)$.

Now consider n closed and bounded sets M_i with positive volume and let $q_i = x^T A_{M_i} x$ be the quadratic form associated with the Fenchel momental ellipsoid of M_i. By (2.4) and (2.8) we have

$$\begin{aligned} D(q_1, \cdots, q_n) &= \frac{1}{n!\,|M_1| \cdots |M_n|} \\ &\sum_{(i_1 \cdots i_n)} \int_{M_{i_1}} \cdots \int_{M_{i_n}} x_1^{(i_1)} \cdots x_n^{(i_n)} \begin{vmatrix} x_1^{(i_1)} & \cdots & x_1^{(i_n)} \\ \vdots & & \vdots \\ x_n^{(i_1)} & \cdots & x_n^{(i_n)} \end{vmatrix} d V^n_{x^{(i_1)}} \cdots d V^n_{x^{(i_n)}} \\ &= \frac{1}{n!\,|M_1| \cdots |M_n|} \int_{M_1} \cdots \int_{M_n} \begin{vmatrix} x_1^{(1)} & \cdots & x_1^{(n)} \\ \vdots & & \vdots \\ x_n^{(1)} & \cdots & x_n^{(n)} \end{vmatrix}^2 d V^n_{x^{(1)}} \cdots d V^n_{x^{(n)}} . \end{aligned} \tag{2.9}$$

Since $T(z, x^{(1)}, \cdots, x^{(n)}) = \pm (1/n!) \det(x_i^{(j)})$ we then have

$$\begin{aligned} &D(q_1, \cdots, q_n) = \\ &\frac{n!}{|M_1| \cdots |M_n|} \int_{M_1} \cdots \int_{M_n} T^2(z, p_1, \cdots, p_n) d V^n_{p_1} \cdots d V^n_{p_n} . \end{aligned} \tag{2.10}$$

The fundamental inequality for mixed discriminants (see [2, p. 53]) is:
(2.11) If the forms $q_1, \cdots, q_{n-1}$ are positive definite and Q is any sym-

metric form, then

$$D^2(q_1, \cdots, q_{n-1}, Q) \geqq D(q_1, \cdots, q_{n-1}, q_{n-1}D(q_1, \cdots, q_{n-2}, Q, Q)$$

where the equality sign holds only if $Q = \lambda q_{n-1}$.

If we set

$$D_p(q, Q) = D(q_1, \cdots, q_{n-p}, \underbrace{Q, \cdots, Q}_{p}) , \tag{2.12}$$

then for n positive definite forms q_i, (2.11) generalizes to

$$D^r(q_1, \cdots, q_n) \geqq \prod_{k=0}^{r-1} D_r(q, q_{n-k}),\ r = 2, 3, \cdots, n \tag{2.13}$$

with equality only if $q_{n-k} = \lambda_{n-k}q_n$ for $k = 0, \cdots, r-1$.

The proof of (2.13) and the condition for equality proceed by induction from the case $r = 2$. The proof is analogous to Alexandrov's generalization [2, p. 50] of a corresponding inequality for mixed volumes and consequently will be omitted here.

If we now set

$$W(M_1, \cdots, M_n, z) = \int_{M_1} \cdots \int_{M_n} T^2(z, p_1, \cdots, p_n)d\,V^n_{p_1} \cdots d\,V^n_{p_n} \tag{2.14}$$

and

$$W_p(M, M_k, z) = W(M_1, \cdots, M_{n-p}, \underbrace{M_k, \cdots M_k}_{p}, z) , \tag{2.15}$$

then by (2.13) and (2.10) we have

$$W^r(M_1, \cdots, M_n, z) \geqq \prod_{k=0}^{r-1} W_r(M, M_{n-k}, z),\ r = 2, \cdots, n \tag{2.16}$$

with the equality sign only if the Fenchel ellipsoids of M_{n-k} are homothetic for $k = 0, \cdots, r-1$. Applying (2.16) to the case $r = n$ and using (2.10) and (2.7), we have

$$[\,|M_1| \cdots |M_n|\,]^{(n+2)/n} \leqq n!\pi_n^2(n+2)^n W(M_1, \cdots, M_n, z) \tag{2.17}$$

with equality only if (except for a set of measure zero) the M_i are homothetic ellipsoids with center z.

The reader will find other inequalities of the above type in [3, pp. 70–71].

3. Centroid surfaces. As before, M is a bounded closed set with positive volume. An oriented hyperplane $L(u)$ through z normal to the direction $u\,(u \neq 0)$ bounds a closed half-space lying on its positive side.

The intersection of this halfspace with M will be denoted by $C(u)$.

Consider the function

$$(3.1) \qquad H(u) = \frac{1}{|M|}\int_M |u \cdot x|\, dV_x^n,\; u \cdot x = \sum_{i=1}^n u_i x_i\ .$$

Since

(a) $H(0) = 0$,
(b) $H(\mu u) = \mu H(u)$ for $\mu > 0$,
(c) $H(u + v) \leqq H(u) + H(v)$,

$H(u)$ is the supporting function (s.f.) of a convex body K^* (see [4, p. 26]), which is nondegenerate and has center z. Let P_0 be the supporting plane (s.p.) to K^* in the direction $u^{(0)}$, the supporting function of $K^* \cap P_0$ is given by the directional derivative

$$(3.2) \qquad \begin{aligned} H'(u^{(0)}; u) &= \lim_{h \to 0+} \frac{H(u^{(0)} + hu) - H(u^{(0)})}{h} \\ &= \frac{1}{|M|}\int_{C(u^{(0)})} u \cdot x dV_x^n - \frac{1}{|M|}\int_{C(-u^{(0)})} u \cdot x dV_x^n\ . \end{aligned}$$

Since $H'(u^{(0)}; u)$ is a linear function of the u_i, P_0 touches K^* in a single point and thus every s.p. of K^* is regular and K^* is strictly convex. (See [4, pp. 25–26].) The derivatives $\partial H / \partial u_i$ are continuous, homogeneous of degree 0, and if y is the point of contact of the s.p. to K^* in the direction u, then

$$(3.3) \qquad y_i = \frac{\partial H}{\partial u_i} = \frac{1}{|M|}\int_{C(u)} x_i dV_x^n - \frac{1}{|M|}\int_{C(-u)} x_i dV_x^n\ .$$

We will call K^* the centroid body of M (with respect to z) and the surface of K^* will be called the centroid surface of M. One may observe that if M happens to have center z, then the centroid surface of M is precisely the set of all controids of $C(u)$ for $u \in \Omega_n$. In general, let $s^{(1)}$ and $s^{(2)}$ be the centroids of $C(u)$ and $C(-u)$ respectively, the y is the center of mass of the two points $s^{(1)}$ and $-s^{(2)}$ provided with mass $|C(u)|/|M|$ and $|C(-u)|/|M|$ respectively. If $|C(u)| = 0$, we will define the centroid of $C(u)$ to be the point z.

It is evident that if M is transformed into $\bar{M}$ by a central affinity, then this transformation also carries the centroid surface of M into the centroid surface of $\bar{M}$.

We now wish to impose additional restrictions on M such that $H(u)$ has continuous second partial derivatives and the surface of K^* has positive Gauss curvature. The following two conditions are sufficient for this purpose:

(a) The set $M(u)$ has positive $(n-1)$-dimensional measure for all $u \in \Omega_n$.

(b) For any $u^{(0)} \in \Omega_n$ and any sequence $u^{(i)} \to u^{(0)}$, the $\lim_{u^{(i)} \to u^{(0)}} M(u^{(i)})$ coincides with $M(u^{(0)})$ except for a possible set of zero $(n-1)$-dimensional measure.

To simplify the calculation of the second partial derivatives at a point $u^{(0)}$, we introduce what Busemann [2, p. 57] calls "standard coordinates." With the same origin and orientation, the x_n axis is chosen such that $u_1^{(0)} = \cdots = u_{n-1}^{(0)} = 0$ and $u_n^{(0)} > 0$. It then follows from (3.3) that

$$\frac{\partial^2 H(u^{(0)})}{\partial u_k \partial u_n} = \frac{\partial^2 H(u^{(0)})}{\partial u_n \partial u_k} = 0 \, . \tag{3.4}$$

Although standard coordinates vary from point to point, the end result (3.9) is expressed geometrically and therefore independent of the coordinate system.

For $j < n$, let $u = (0, \cdots, 0, u_j, 0, \cdots, 0, u_n^{(0)})$ and set

$$N_1 = C(u) \cap C(u^{(0)}),\ N_1^* = C(-u) \cap C(-u^{(0)}),\ N_2 = C(u^{(0)}) - N_1,$$
$$N_2^* = C(-u) - N_1^*,\quad N_3 = C(u) - N_1, \qquad N_3^* = C(-u^{(0)}) - N_1^*.$$

Except for a set of zero n-dimensional measure, $N_2 = N_2^*$ and $N_3 = N_3^*$. By (3.3) we have for $i, j < n$

$$\frac{\left(\frac{\partial H}{\partial u_i}\right)_u - \left(\frac{\partial H}{\partial u_i}\right)_{u^{(0)}}}{u_j} = \frac{2}{u_j \,|\, M \,|}\left(\int_{N_3} x_i d V_x^n - \int_{N_2} x_i d V_x^n\right) . \tag{3.5}$$

We will calculate the limit of (3.5) as either $u_j \to 0+$ or $u_j \to 0-$. In either case for $x \in N_3$, $x_j u_j \geqq 0$, $x_n \leqq 0$ and for $x \in N_2$, $x_j u_j \geqq 0$, $x_n \geqq 0$. For $-\pi/2 < v_n < \pi/2$, let the hyperplane $x_n = (\tan v_n) x_j$ intersect M in $M^+(v_n)$ for $x_n \geqq 0$ and in $M^-(v_n)$ for $x_n \leqq 0$. Also the volume element $d V_x^{n-1}$ of this hyperplane is

$$d V_x^{n-1} = dx_1 \cdots dx_{n-1} \sec v_n \, . \tag{3.6}$$

We introduce new coordinates $v_1, \cdots, v_n$ by $x_i = v_i$ for $i = 1, \cdots, n-1$ and $x_n = v_j \tan v_n$ which uniquely define the v_i with $-\pi/2 < v_n < \pi/2$ for all x for which $x_j \neq 0$. The Jacobian J of this transformation is $J = v_j \sec^2 v_n$. Also define α, $0 \leqq \alpha < \pi/2$, by $u_n^{(0)} \tan \alpha = | u_j |$. Then $| J | / u_j = \pm v_j \sec^2 v_n / u_n^{(0)} \tan \alpha$ with the plus sign for $x \in N_3$ and the minus sign for $x \in N_2$. The difference quotient (3.5) is, consequently, given by

$$\frac{2}{u_n^{(0)} \,|\, M \,| \tan \alpha}\left[\int_0^{\alpha} \sec v_n \left(\int_{M^-(v_n)} v_i v_j d V_v^{n-1}\right) d v_n\right.$$

$$+ \int_0^\alpha \sec v_n \Big(\int_{M^+(v_n)} v_i v_j d V_v^{n-1} \Big) d v_n \Big]$$

and since the integrands are continuous functions of v_n by assumption (b) we have

$$(3.7) \qquad H_{ij}(u^{(0)}) = \frac{\partial^2 H(u^{(0)})}{\partial u_i \partial u_j} = \frac{2}{u_n^{(0)} \mid M \mid} \int_{M(u^{(0)})} x_i x_j d V_x^{n-1}, \ (i, j < n) .$$

Now let $H^{(i)}(u)$ be the supporting function (3.1) for the set M_i, $i = 1, \cdots, n-1$. Set $H = \lambda_1 H^{(1)} + \cdots + \lambda_n H^{(n-1)}$, then $D_{n-1}(H)$ is defined as the sum of all principal $(n-1)$ rowed minors of the matrix H_{ij} (with components evaluated for a unit vector) and is a homogeneous polynomial of degree $(n-1)$ in the λ_i. (See [4, p. 59] or [2, pp. 45–46].) The quantity $D(H^{(1)}, \cdots, H^{(n-1)})$ denotes the factor of $\lambda_1 \cdots \lambda_{n-1}$ in $D_{n-1}(H)$ divided by $(n-1)!$. If we calculate (3.7) for each of the $H^{(i)}$ using the same standard coordinates we have, because of (3.4),

$$(3.8) \qquad D(H^{(1)}, \cdots, H^{(n-1)}) = \frac{1}{(n-1)!} \sum_{(i_1 \cdots i_{n-1})} \begin{vmatrix} H_{11}^{(i_1)} & \cdots & H_{n(n-1)}^{(i_{n-1})} \\ \vdots & & \vdots \\ H_{(n-1)1}^{(i_1)} & \cdots & H_{(n-1)(n-1)}^{(i_{n-1})} \end{vmatrix} .$$

In the same way as we derived (2.10), we find for any $u \in \Omega_n$,

$$(3.9) \qquad D(H^{(1)}, \cdots, H^{(n-1)}) = \frac{(n-1)! 2^{n-1}}{\mid M_1 \mid \cdots \mid M_{n-1} \mid} \int_{M_1(u)} \cdots \int_{M_{n-1}(u)} T^2(z, p_1, \cdots, p_{n-1}) d V_{p_1}^{n-1} \cdots d V_{p_{n-1}}^{n-1} .$$

By comparison with (2.10) we observe that

$$(3.10) \qquad D(H^{(1)}, \cdots, H^{(n-1)}) = \frac{2^{n-1} \mid M_1(u) \mid \cdots \mid M_{n-1}(u) \mid}{\mid M_1 \mid \cdots \mid M_{n-1} \mid} D(q_1, \cdots, q_{n-1})$$

where q_i is the quadratic from associated with the Fenchel momental ellipsoid of $M_i(u)$ in the $(n-1)$-dimensional space $L(u)$.

From (3.9), we may give an integral interpretation of an elementary symmetric function $\{R_1 \cdots R_m\}$ of the principal radii of curvature of the centroid surface of M. With H given by (3.1) we have for $m = 1, \cdots, n-1$ (see [4, p. 63]),

$$(3.11) \qquad \{R_1 \cdots R_m\} = \binom{\mathrm{n}-1}{\mathrm{m}} D(\underbrace{\mid u \mid, \cdots, \mid u \mid}_{n-m-1}, \underbrace{H, \cdots, H}_{m}) = D_m(H) .$$

Set $M = M_1 = \cdots = M_m$ and $U_n = M_{m+1} = \cdots = M_{n-1}$. Since

$$(3.12) \qquad \frac{1}{\mid U_n \mid} \int_{U_n} \mid u \cdot x \mid d V_x = \frac{2\pi_{n-1}}{(n+1)\pi_n} \mid u \mid ,$$

we obtain

$$\left[\frac{2\pi_{n-1}}{(n+1)\pi_n}\right]^{n-m-1} D_m(H) = \binom{n-1}{m} \frac{(n-1)!2^{n-1}}{|M|^m \pi_n^{n-m-1}}$$
$$\cdot \underbrace{\int_{M(u)} \int_{M(u)}}_{m} \underbrace{\int_{U_n(u)} \int_{U_n(u)}}_{n-m-1} T^2(z, p_1, \cdots, p_{n-1}) dV_{p_1}^{n-1} \cdots dV_{p_{n-1}}^{n-1} .$$

By integrating successively over the $U_n(u)$ and using (2.3) applied to the appropriate dimensions we obtain

$$\text{(3.13)} \quad \{R_1 \cdots R_m\} = \frac{m!2^m}{|M|^m} \int_{M(u)} \cdots \int_{M(u)} T^2(z, p_1, \cdots, p_m) dV_{p_1}^{n-1} \cdots dV_{p_m}^{n-1}$$

for $m = 1, \cdots, n-1$.

We may also give an interpretation of each individual principal radius of curvature. First we show:

(3.14) The Dupin indicatrix of the centroid surface of $M(wrtz)$ at the point of contact y of the tangent plane in the direction u is homothetic to the Fenchel ellipsoid ($wrt\ z$) of $M(u)$ in the space $L(u)$.

A central affinity sends homothetic figures in parallel hyperplanes into homothetic figures. Due to the affine nature of Fenchel ellipsoids and centroid surfaces, we need only show that if the Fenchel ellipsoid of $M(u)$ is a sphere, then the Dupin indicatrix at y is a sphere. However, this follows at once from (2.4) and the representation (3.7) in standard coordinates since the principal radii of curvature R_i must satisfy

$$\begin{vmatrix} H_{11}-R & \cdots & H_{1n} \\ \vdots & & \vdots \\ H_{n1} & \cdots & H_{nn}-R \end{vmatrix} = 0$$

where H_{ij} are evaluated for a unit vector. (See [4, p. 61].)

Now, let the line through z, parallel to the ith principal direction of the centroid surface at y, be normal to the $(n-2)$ space L_{n-2} through z in $L(u)$. Then R_i is given by

$$\text{(3.14)} \qquad R_i = \frac{2}{|M|} I(M(u), L_{n-2})$$

where $I(M(u), L_{n-2})$ is the second moment, in $L(u)$, of $M(u)$ with respect to L_{n-2}.

Returning to the $(n-1)$ bodies $M_1, \cdots, M_{n-1}$ for which we obtained (3.9), let $H^{(n)}(u)$ be the supporting function (3.1) corresponding to any bounded closed set M_n with positive volume. Then (see [2, p. 46]),

$$\text{(3.15)} \qquad V(K_1^*, \cdots, K_n^*) = n^{-1} \int_{\Omega_n} H^{(n)} D(H^{(1)}, \cdots, H^{(n-1)}) d\omega_u^n, \ |u| = 1 ,$$

where $V(K_1^*, \cdots, K_n^*)$ is the mixed volume of $K_1^*, \cdots, K_n^*$. Using (3.9), (3.15), (3.1) and the integration technique of Busemann in [1] where it is shown that

$$dV_{p_1}^n \cdots dV_{p_{n-1}}^n = (n-1)!\,T(z, p_1, \cdots, p_{n-1})dV_{p_1}^{n-1} \cdots dV_{p_{n-1}}^{n-1} d\omega_u^n ,$$

we obtain

$$(3.16) \qquad V(K_1^*, \cdots, K_n^*) = \frac{2^n}{|M_1| \cdots |M_n|} \int_{M_1} \cdots \int_{M_n} T(z, p_1, \cdots, p_n) dV_{p_1}^n \cdots dV_{p_n}^n .$$

Since both sides of (3.16) vary continuously with the M_i, we may extend this result to any n bounded and closed sets M_i with $|M_i| > 0$. Briefly, we may assume $z \in M_i$ and let $\varepsilon_j > 0$ be a sequence such that $\varepsilon_j \to 0$. A covering of open spheres of radius ε_j with centers in M_i may be reduced to a finite covering since M_i is compact. Conditions (a) and (b) are then satisfied for the closure of such a finite covering and the extention of (3.16) follows.

There is an alternate proof of (3.16) which proceeds directly from (3.1). We did not resort to this at the outset since the intervening results are of interest in themselves. Briefly, the alternate proof is as follows: We approximate the $H^{(i)}(u)$ of (3.1) by

$$E^{(i,k)}(u) = \frac{1}{|M_i|} \sum_{j=1}^{k} |u \cdot x^{(j)}| \Delta V_j^n$$

such that $E^{(i,k)} \to H^{(i)}$ as $k \to +\infty$. Now $|u \cdot x|$ is the supporting function of the segment $\bar{x}$ with end-points x and $-x$. Also, by induction, one shows that

$$V(\bar{x}^{(1)}, \cdots, \bar{x}^{(n)}) = 2^n T(z, x^{(1)}, \cdots, x^{(n)}) .$$

The function $E^{(i,k)}$ is the supporting function of the linear combination

$$E_{(i,k)} = \frac{1}{|M_i|} \sum_{j=1}^{k} \bar{x}^{(j)} \Delta V_j^n .$$

For $\lambda_j > 0$ the linear combination $E_k = \lambda_1 E_{(1,k)} + \cdots + \lambda_n E_{(n,k)}$ may also be expressed as a linear combination of the nk segments $\bar{x}^{(j_i)}$. Expressing the volume of E_k as a polynomial in the λ_i in two ways we have by comparing the coefficient of $\lambda_1 \cdots \lambda_n$

$$V(E_{(1,k)}, \cdots, E_{(n,k)}) = \frac{2^n}{|M_1| \cdots |M_n|} \sum_{j_1=1}^{k} \cdots \sum_{j_n=1}^{k} T(z, x^{(j_1)}, \cdots, x^{(j_n)}) \Delta V_{j_1}^n \cdots \Delta V_{j_n}^n$$

and (3.16) follows in the limit as $k \to +\infty$.

The formula (3.16) may be substituted into inequalities of mixed volumes to yield inequalities of the integrals. Since the number of times a component appears on each side of a mixed volume inequality is always the same, the coefficient on the righthand side of (3.16) cancels leaving, as in (2.16), inequalities among the integrals only. However, in this case when the uniqueness theorem (4.1) applies, the condition for equality may be passed through the K_i^* to the M_i.

In [1, p. 11], Busemann shows that if M is a nondegenerate convex body, then

$$(3.17) \qquad \int_M \cdots \int_M T(z, p_1, \cdots, p_n) d V_{p_1} \cdots d V_{p_n} \geqq \frac{2}{(n+1)!} \frac{\pi_{n+1}^{n-1}}{\pi_n^{n+1}} |M|^{n+1}$$

with equality only if M is an ellipsoid with center z. We define the expanded centroid body K of M to be the dilation of K^* about z by the factor $(n+1)\pi_n/2\pi_{n-1}$. By (3.12), we see that this is the factor which dilates the centroid body of an ellipsoid with center z into coincidence with the ellipsoid.

From (3.16) we obtain a reinterpretation of (3.17) by observing the identity $n!\pi_n\pi_{n-1} = 2^n\pi^{n-1}$:

(3.18) If K is the expanded centroid body of a nondegenerate convex body M, then $|K| \geqq |M|$ with equality only if M is an ellipsoid with center z.

The convexity of M is not an essential feature in (3.18) and the Steiner symmetrization used to prove (3.17) may be extended to include nonconvex sets.

Using the expanded centroid bodies K_i of M_i, we may write (3.16) as

$$(3.19) \qquad \begin{aligned} &|M_1| \cdots |M_n| \; V(K_1, \cdots, K_n) \\ &\quad = \frac{(n+1)!\pi_n^{n+1}}{2\pi_{n+1}^{n-1}} \int_{M_1} \cdots \int_{M_n} T(z, p_1, \cdots, p_n) d V_{p_1}^n \cdots d V_{p_n}^n \end{aligned}$$

and if we define K_i to be the point z if $|M_i| = 0$ then (3.19) holds for any bounded closed sets M_i.

Substituting (3.19) into (1.1) we have

(3.20) THEOREM. *If $K_i(u)$ is the expanded centroid body of $M_i(u)$ in the $(n-1)$-dimensional space $L(u)$, then*

$$\begin{aligned} &|M_1| \cdots |M_{n-1}| \\ &\quad = \frac{1}{n} \frac{\pi_n^{n-2}}{\pi_{n-1}^n} \int_{\Omega_n} |M_1(u)| \cdots |M_{n-1}(u)| \; V(K_1(u), \cdots, K_{n-1}(u)) d\omega_u^n . \end{aligned}$$

The inequality $V^{n-1}(K_1(u), \cdots, K_{n-1}(u)) \geqq |K_1(u)| \cdots |K_{n-1}(u)|$ (see [2, p. 50]) and (3.18) reproduces (1.3).

There are two special cases of (3.20) of particular geometric interest. First, set $M = M_1 = \cdots = M_{n-1}$, then

$$(3.21) \qquad |M|^{n-1} = \frac{1}{n}\frac{\pi_n^{n-2}}{\pi_{n-1}^n}\int_{\Omega_n} |M(u)|^{n-1}|K(u)|\,d\omega_u^n\,.$$

Next, for $n \geqq 3$, set $M = M_1 = \cdots = M_{n-2}$, $M_{n-1} = U_n$, then

$$(3.22) \qquad |M|^{n-2} = \frac{1}{n(n-1)}\frac{\pi_n^{n-3}}{\pi_{n-1}^{n-1}}\int_{\Omega_n} |M(u)|^{n-2}S(K(u))d\omega_u^n$$

where $S(K(u))$ is the surface area of $K(u)$ in the space $L(u)$.

4. **Uniqueness theorems.** In order for K^* to determine M, additional restrictions on M are necessary as may be seen by consideration of a set M bounded by two concentric spheres.

(4.1) THEOREM. *Suppose M_i $(i = 1, 2)$ can be represented in polar coordinates by $0 \leqq r \leqq \rho_i(u)$, $u \in \Omega_n$ where $\rho_i(u)$ is an even,* i.e., *$\rho_i(u) = \rho_i(-u)$, continuous function on Ω_n. If the centroid surface of M_1(wrt z) is identical to the centroid surface of M_2(wrt z), then M_1 and M_2 are identical.*

(4.2) THEOREM. *Suppose M_i $(i = 1, 2)$ have the same representation as in* (4.1). *If $|M_1(u)| = |M_2(u)|$ for all $u \in \Omega_n$, then M_1 and M_2 are identical.*

The latter theorem is a result, for $n = 3$, of *P.* Funk [6].

We first prove (4.1). From (3.1) and the assumption on the representation of M_i we have

$$H^{(i)}(u) = \frac{1}{(n+1)|M_i|}\int_{\Omega_n} |u\cdot\tau|\,\rho_i^{n+1}(\tau)d\omega_\tau^n,\ |\tau| = 1\,.$$

Consequently, (4.1) follows from the uniqueness of the solution of an integral equation of the first kind. Namely:

(4.3) THEOREM. *Let $h(\tau)$ be an even, continuous function on Ω_n. If for unit vectors u and τ*

$$\int_{\Omega_n} |u\cdot\tau|\,h(\tau)d\omega_\tau^n = 0$$

for all $u \in \Omega_n$, then $h(\tau)$ vanishes identically.

The result (4.3) is well known for $n = 2, 3$ and the recent extension of surface harmonics to n-dimensions, in particular the Funk-Hecke theorem, enables one to prove (4.3) for all n. There are two steps in the following proof (which applies for $n \geqq 3$). First, from the completeness [5, p. 241] it suffices to show that

$$\int_{\Omega_n} S_m(\tau)h(\tau)d\omega_\tau^n = 0$$

for all the linearly independent surface harmonics $S_m(\tau)$ of degree m and for $m = 0, 1, 2, \cdots$. Since $h(\tau)$ is an even function we need only to consider, now, even m. Next, from the Funk-Hecke theorem [5, pp. 247–248] we have

$$\int_{\Omega_n} |u \cdot \tau| S_m(u)d\omega_u^n = \lambda_m S_m(\tau)$$

where

$$\lambda_m = \frac{(4\pi)^\nu m!\Gamma(\nu)}{(m + 2\nu - 1)!}\int_{-1}^{1} |x| C_m^\nu(x)(1 - x^2)^{\nu-1/2}dx \tag{4.4}$$

and $\nu = (n - 2)/2 \geqq 1/2$. Thus, we need only to verify that $\lambda_m \neq 0$ for $m = 0, 2, 4, \cdots$. For $m = 0$, $C_0^\nu(x) = 1$ and $\lambda_0 \neq 0$. For $m > 0$,

$$C_m^\nu(x) = a_{m,\nu}(1 - x^2)^{-\nu+1/2}\frac{d^m}{dx^m}[(1 - x^2)^{m+\nu-1/2}]$$

where $a_{m,\nu} \neq 0$. See [5, p. 236] for the explicit expression of the coefficient $a_{m,\nu}$. Thus the integral in (4.4) is

$$I_{m,\nu} = 2a_{m,\nu}\int_0^1 x\frac{d^m}{dx^m}[(1 - x^2)^{m+\nu-1/2}]dx$$

and using integration by parts

$$I_{m,\nu} = 2a_{m,\nu}(-1)^{\frac{m-2}{2}}(m - 2)!\binom{m + \nu - 1/2}{\frac{m}{2} - 1} \neq 0$$

for $m = 2, 4, 6, \cdots$ which completes the proof.

The result (4.2) is clearly a consequence of the following spherical integration theorem.

(4.5) Let $f(\tau)$ be a continuous even function defined on Ω_n. If

$$\int_{\Omega_n(u)} f(\tau)d\omega_\tau^{n-1} = 0$$

for all $u \in \Omega_n$, then $f(\tau)$ vanishes identically.

A proof of (4.5) for $n = 3$ can be found in [4, pp. 136–138]. However, a proof for all $n \geqq 3$ is easily obtained from (4.1). To see this, set $g(\tau) = f(\tau) - [\min f(\tau)] + 1 > 0$. Let $\rho(\tau) = [g(\tau)]^{1/(n+1)}$ and let M be the set whose polar coordinates satisfy $0 \leqq r \leqq \rho(\tau)$, $\tau \in \Omega_n$. Using (**3.13**) for $m = 1$, the sum of the principal radii of curvature of the centroid surface of M *wrt* z may be expressed, in this case, by

$$R_1 + \cdots + R_{n-1} = \frac{2}{(n+1)\,|\,M\,|}\int_{\Omega_n(u)} g(\tau)d\omega_\tau^{n-1}$$

and, by hypothesis, this is a positive constant for $u \in \Omega_n$. However, this implies (see [4, pp. 117–118]) that the centroid surface is a sphere and by (4.1), M is a solid sphere and $g(\tau)$ is a constant which completes the proof.

REFERENCES

1. H. Busemann, *Volume in terms of concurrent cross-sections*, Pacific J. Math., **3** (1953). 1-12.
2. ———, *Convex Surfaces*, Interscience Publishers, New York, 1958.
3. H. Busemann and E. G. Straus, *Area and normality*, Pacific J. Math., **10** (1960), 35-72.
4. T. Bonnesen and W. Fenchel, *Theorie der konvexen Körper*, Chelsea, New York, 1948.
5. A. Erdélyi, et al., *Higher transcendental functions*, Bateman Manuscript Project, vol. 2, New York, 1953.
6. P. Funk, *Über Flächen mit lauter geschlossenen geodätischen Linien*, Math. Ann., **74** (1913), 278-300.

APPLIED MATHEMATICS, LOCKHEED MISSILES AND SPACE DIVISION, SUNNYVALE, CALIFORNIA

ASYMPTOTIC ESTIMATES FOR LIMIT CIRCLE PROBLEMS

C. A. SWANSON

1. Preliminaries. Characteristic value problems will be considered for the second order, ordinary, linear differential operator L defined by

$$Lx = \frac{1}{k(s)}\left\{-\frac{d}{ds}\left[p(s)\frac{dx}{ds}\right] + q(s)x\right\} \tag{1.1}$$

on the open interval $\omega_- < s < \omega_+$, where k, p, q are real-valued functions on this interval with the properties that

(i) p is differentiable;

(ii) k and q are piecewise continuous; and

(iii) k and p are positive-valued. The points ω_- and ω_+ are in general singularities of L; the possibility that they are $\pm\infty$ is not excluded. It will be convenient to use the notations

$$(x, y)_s^t = \int_s^t x(u)\bar{y}(u)k(u)du, \qquad \omega_- \leqq s < t \leqq \omega_+ , \tag{1.2}$$

$$[xy](s) = p(s)[x(s)\bar{y}'(s) - x'(s)\bar{y}(s)] . \tag{1.3}$$

Then Green's symmetric formula for L has the form

$$(Lx, y)_s^t - (x, Ly)_s^t = [xy](t) - [xy](s) . \tag{1.4}$$

The symbols $[xy](\pm)$ will be used as abbreviations for the limits of $[xy](s)$ as $s \to \omega_\pm$, and (x, y) will be used for the left member of (1.2) when s, t have been replaced by ω_-, ω_+. Let $\mathfrak{H}$, $\mathfrak{H}_{ab}$ denote the Hilbert spaces which are the Lebesgue spaces with respective inner products (x, y), $(x, y)_a^b$ and norms $||x|| = (x, x)^{1/2}$, $||x||_a^b = [(x, x)_a^b]^{1/2}$, $\omega_- \leqq a < b \leqq \omega_+$.

Let a_0 and b_0 be fixed numbers satisfying $\omega_- < a_0 < b_0 < \omega_+$ and let R_0 be the rectangle in the $a-b$-plane described by the inequalities $\omega_- < a \leqq a_0$, $b_0 \leqq b < \omega_+$. Every closed, bounded subinterval $[a, b]$ of the basic interval (ω_-, ω_+) can be associated in a one-to-one manner with a point in R_0. For every such $[a, b]$ we shall consider the regular Sturm-Liouville problem

$$Ly = \mu y, \qquad U_a y = U_b y = 0 \tag{1.5}$$

on $[a, b]$, where U_a, U_b are the linear boundary operators

$$\begin{aligned} U_a y &= \alpha_0(a)y(a) + \alpha_1(a)y'(a) \\ U_b y &= \beta_0(b)y(b) + \beta_1(b)y'(b) , \end{aligned} \tag{1.6}$$

with α_0, α_1 real-valued functions not both 0 for any value of a on $(\omega_-, a_0]$,

Received January 5, 1961.

and with β_0, β_1 real-valued and not both 0 on $[b_0, \omega_+]$. Our problem is to obtain estimates for each characteristic value $\mu = \mu_{ab}$ of (1.5) for a, b near ω_-, ω_+ under hypotheses that will ensure that the limits of μ_{ab} as $a, b \to \omega_-, \omega_+$ will exist. Also, we shall obtain estimates for the corresponding characteristic functions $y = y_{ab} = y_{ab}(s)$ on $a \leqq s \leqq b$. Results like this for differential operators having a singularity at one endpoint were obtained previously by an integral equations approach [8], [9]. The present paper contains extensions of some of these results to operators (1.1) which have singularities at both endpoints. Furthermore, the present approach to the problem will be different; the estimates will now be obtained by means of projection mappings on suitable Hilbert spaces. The method arises from an idea communicated by Professor H. F. Bohnenblust, and affords an elegant and abstract approach to the type of perturbation problem at hand [1]. Also, the present method is powerful enough to handle a variety of domain-perturbed problems that arise in the study of elliptic partial differential equations. Some of these have been considered already [10] and the author has several others in preparation.

Here the method will be illustrated in the case that both of the singularities $\omega_\pm$ of the operator (1.1) are *limit circle* singularities in the well-known classification of H. Weyl [2, p. 225]. In another paper we shall consider the limit point cases (and mixed cases) in which some additional hypotheses are needed on the growth of the coefficient functions in (1.1) as $s \to \omega_\pm$ to ensure the existence of isolated characteristic values λ of L on (ω_-, ω_+); however, very general boundary operators U_a, U_b will then permit convergence of μ_{ab} to λ. For additional details, see [8]. In the limit circle case herein under consideration, no special assumptions will be imposed on the nature of L at $\omega_\pm$, but the generality of the boundary operators must be sacrificed in order to ensure the convergence of μ_{ab}. Our purpose here is to obtain asymptotic estimates rather than asymptotic expansions for the characteristic values and functions as $a, b \to \omega_-, \omega_+$. Asymptotic formulae and expansions will be published elsewhere.

2. Basic and perturbed problems. Rather than general spectral theory, we are interested in cases that the limits of μ_{ab} as $a, b \to \omega_-, \omega_+$ exist in the elementary sense. Thus, characteristic values of suitable singular boundary value problems for L on (ω_-, ω_+) are supposed to exist. These singular problems are described differently according as the points $\omega_\pm$ are in the limit point or limit circle categories. The description is made as follows when both are limit circle singularities [2], [6]: choose a complex number l_0 with $\operatorname{Im} l_0 \neq 0$, and let L_0 be the differential operator $L - l_0$. A theorem of Weyl [6] states that there exist linearly independent solutions $\varphi_\pm \in \mathfrak{H}$ of $L_0\varphi = 0$ such that

$$(2.1) \qquad [\varphi_-\varphi_-](-) = [\varphi_+\varphi_+](+) = 0, \qquad [\varphi_+\bar{\varphi}_-](s) = 1 .$$

Let $\mathfrak{D}$ denote the domain consisting of all $x \in \mathfrak{H}$ which have the following properties:

(a) x is differentiable on (ω_-, ω_+) and x' is absolutely continuous on every closed subinterval of this interval:

(b) $Lx \in \mathfrak{H}$

(c) x satisfies the end conditions

$$[x\varphi_-](-) = [x\varphi_+](+) = 0 \ . \tag{2.2}$$

Then L on $\mathfrak{D}$ is real and essentially self-adjoint [6]. The *basic* characteristic problem

$$Lx = \lambda x, \qquad x \in \mathfrak{D} \tag{2.3}$$

is known to have a denumerable set of characteristic values λ_n and corresponding characteristic functions x_n which are orthonormal and complete in $\mathfrak{H}$ $(n = 1, 2, \cdots)$.

Two classes of perturbation problems (1.5) will be considered. The limiting behaviour of *class* 1 *boundary operators* U_a, U_b as $a, b \to \omega_-, \omega_+$ is rather arbitrary (see §5) while the limiting behaviour of *class* 2 *operators* (§§2, 3, 4) is restricted as follows:

$$\begin{aligned} U_a y &= [y\varphi_-](a)[1 + o(1)] \qquad \text{as } a \to \omega_- \\ U_b y &= [y\varphi_+](b)[1 + o(1)] \qquad \text{as } b \to \omega_+ \end{aligned} \tag{2.4}$$

for every differentiable function y. A *perturbed domain* $\mathfrak{D}_{ab}$ is defined for each $[a, b] \in R_0$ to be the set of all y in the subspace $\mathfrak{F}_{ab}$ of $\mathfrak{H}$ which satisfy the following conditions:

(a) y is differentiable and y' is absolutely continuous on $[a, b]$;

(b) $Ly \in \mathfrak{F}_{ab}$

(c) y satisfies the homogeneous boundary conditions (1.5) where the boundary operators U_a, U_b have the limiting behaviour (2.4).

The perturbed characteristic value problem that corresponds to this domain is the regular Sturm-Liouville problem

$$Ly = \mu y, \qquad y \in \mathfrak{D}_{ab} \ . \tag{2.5}$$

In addition, we define a domain $\mathfrak{D}_a$ for each a on $(\omega_-, a_0]$ to be the set of all $z \in \mathfrak{F}_{a\omega_+}$ which satisfy the following:

(a) z is differentiable and z' is absolutely continuous on every closed subinterval of $[a, \omega_+)$;

(b) $Lz \in \mathfrak{F}_{a\omega_+}$

(c) z satisfies the conditions

$$U_a z = 0, \quad [z\varphi_+](+) = 0 \ . \tag{2.6}$$

The characteristic value problem

$$(2.7) \qquad Lz = \nu z, \qquad z \in \mathfrak{D}_a$$

on the half-open interval $[a, \omega_+)$ may be regarded as intermediate between (2.3) and (2.5), and will be called a semi-perturbed problem.

In order to obtain estimates for the difference between the characteristic values and functions of (2.5) and (2.3), we shall proceed in two steps: (i) the comparison of (2.5) with (2.7), and (ii) the comparison of (2.7) with (2.3). The details of (i) and (ii) are included in §§3 and 4 respectively. Each comparison has independent interest because it is typical for a boundary variational problem when only one endpoint is varied and the unchanged endpoint is (i) an ordinary point; (ii) a singular point of the differential operator. Type (ii) variational problems arise for example in the theory of enclosed quantum mechanical systems[4], [5].

3. **Comparison of the y and z problems.** The characteristic value problems (2.5) and (2.7) will be compared, with (2.7) regarded as basic and (2.5) regarded as a perturbation on (2.7). In this case, the singular boundary condition $[z\varphi_+](+) = 0$ is replaced by the regular condition $U_b z = 0$ at the point b. We are going to estimate the variation of characteristic values and functions under this perturbation, and show that this variation has the limit 0 as $b \to \omega_+$. The ordinary endpoint a remains fixed in this section.

Let $G_{ab}(s, t)$ be the Green's function for the operator kL_0 associated with the boundary conditions (1.5), and let G_{ab} be the linear transformation on $\mathfrak{H}_{ab}$ defined by the equation

$$(3.1) \qquad G_{ab}y = \int_a^b G_{ab}(s, t)y(t)k(t)dt, \qquad y \in \mathfrak{H}_{ab} .$$

Let $\nu = \nu_a$ be a characteristic value for (2.7) and let z_a be the corresponding characteristic function. Define a function f on $[a, b]$ by[1]

$$(3.2) \qquad f = z_a - \gamma_a G_{ab} z_a \qquad \text{where } \gamma_a = \nu_a - l_0 .$$

It is easily verified because of the linearity of all the operators involved that f is a solution of the boundary value problem

$$(3.3) \qquad L_0 f = 0, \qquad U_a f = 0, \qquad U_b f = U_b z_a .$$

The solution ψ_a of $L_0 y = 0$ that is given by

$$(3.4) \qquad \psi_a(s) = \varphi_-(s) U_a \varphi_+ - \varphi_+(s) U_a \varphi_-$$

satisfies the boundary condition $U_a y = 0$. Hence the unique solution of (3.3) is

[1] The function on $[a, b]$ which coincides with z_a on this interval will also be denoted by z_a.

$$f(s) = (U_b z_a / U_b \psi_a)\psi_a(s), \qquad a \leqq s \leqq b . \tag{3.5}$$

In fact, if g is any solution of (3.3), then the function $h = g - f$ satisfies $L_0 h = 0$, $U_a h = U_b h = 0$. This implies that h is the zero function, or $g = f$.

It follows from (2.1) that $[\varphi_+\varphi_+](b) \to 0$ as $b \to \omega_+$ and $[\varphi_-\varphi_-](a) \to 0$ as $a \to \omega_-$. The identity

$$[\varphi_+\varphi_+](t) - [\varphi_+\varphi_+](s) = (l_0 - \bar{l}_0)(\|\varphi_+\|_s^t)^2$$

is a consequence of (1.4), and since $\varphi_+ \in \mathfrak{H}$, the limit $[\varphi_+\varphi_+](-)$ exists. Similarly $[\varphi_-\varphi_-](+)$ exists. From (2.1) and the identity [6]

$$|[\varphi_+\bar{\varphi}_-](a)|^2 = [\varphi_-\varphi_-](a)[\varphi_+\varphi_+](a) + |[\varphi_+\varphi_-](a)|^2$$

we deduce that $|[\varphi_+\varphi_-](a)| \to 1$ as $a \to \omega_-$. Similarly $|[\varphi_+\varphi_-](b)| \to 1$ as $b \to \omega_+$. It has then been established that

$$\begin{aligned} &U_a\varphi_- \to 0, \quad |U_a\varphi_+| \to 1 \quad \text{as } a \to \omega_- ; \\ &U_b\varphi_+ \to 0, \quad |U_b\varphi_-| \to 1 \quad \text{as } b \to \omega_+ \end{aligned} \tag{3.6}$$

where (2.4) has been used. Since $\varphi_\pm \in \mathfrak{H}$, it follows from (3.4) that $\|\psi_a\|_a^b$ is uniformly bounded for $[a, b] \in R_0$. We obtain from (3.4) that

$$U_b\psi_a = U_b\varphi_- U_a\varphi_+ - U_b\varphi_+ U_a\varphi_-$$

and hence there are numbers a_0, b_0 (we may suppose that they coincide with the original choices of a_0, b_0) such that $U_b\psi_a$ is bounded away from zero on $a \leqq a_0$, $b_0 \leqq b$. These considerations enable us to deduce from (3.2), (3.5) that there exists a constant[2] C on R_0 such that

$$\|z_a - \gamma_a G_{ab} z_a\|_a^b \leqq C|U_b z_a| \, \|z_a\|_a^b, \qquad [a, b] \in R_0 . \tag{3.7}$$

Let $\mu^i = \mu^i_{ab}$ denote the ith characteristic value of the regular problem (2.5), $\mu^1 < \mu^2 < \cdots$, and let y^i denote the corresponding characteristic function, chosen so that $\{y^i\}$ is an orthonormal basis in $\mathfrak{F}_{ab}$. The following fundamental lemma was obtained by H. F. Bohnenblust in [1] by applying the Parseval completeness relation to the set $\{y^i\}$. An outline of the proof is reproduced below.

LEMMA. *Let $P(\delta)$ be the projection mapping from the Hilbert space $\mathfrak{F}_{ab}$ onto its subspace $\mathfrak{F}_{ab}(\delta)$ $(\delta > 0)$ spanned by all characteristic functions y^i of (2.5) such that their corresponding μ^i satisfy $|\mu^i - \nu_a| \leqq \delta$. Then for any $w \in \mathfrak{F}_{ab}$,*

$$\|w - P(\delta)w\|_a^b \leqq (1 + |\gamma_a|/\delta)\, \|w - \gamma_a G_{ab} w\|_a^b .$$

[2] The letter C will be used throughout as a generic notation for the image of a constant function from R_0 into the positive numbers.

Proof. The subscripts a, b will be omitted in this proof. Let $\alpha_i = (Gw, y^i)$. It is easily verified that $(w - \gamma Gw, y^i) = (\mu^i - \nu)\alpha_i$, and hence by the Parseval identity,

$$||w - \gamma Gw||^2 = \sum_i |\mu^i - \nu|^2 |\alpha_i|^2 \geqq \delta^2 \sum_i {}^* |\alpha_i|^2 ,$$

where the * denotes summation over only those indices i such that $|\mu^i - \nu| > \delta$. Then

$$||Gw - P(\delta)Gw||^2 = \sum{}^* |\alpha_i|^2 \leqq \delta^{-2} ||w - \gamma Gw||^2 ,$$

and the conclusion of the lemma follows easily from the Minkowski inequality.

The notation $\rho_b = C|\gamma_a U_b z_a|$ will be used. It follows from (2.4) and (2.6) that $\rho_b \to 0$ as $b \to \omega_+$ for each fixed a. With the choice $\delta = 2\rho_b$, we apply the lemma to $w = z_a$ (see footnote 1) and use (3.7) to obtain

$$||z_a - P(2\rho_b)z_a||_a^b \leqq (C|U_b z_a| + \tfrac{1}{2})\, ||z_a||_a^b .$$

We may suppose that b_0 has been selected so that $C|U_b z_a| \leqq 1/4$ on $b_0 \leqq b < \omega_+$. Hence $P(2\rho_b)z_a = 0$ implies that $z_a = 0$ on $[a, b]$, and therefore $\mathfrak{F}_{ab}(2\rho_b)$ has dimension $\geqq 1$. Hence there exists at least one characteristic value $\mu = \mu_{ab}$ of (2.5) which satisfies

$$|\mu_{ab} - \nu_a| \leqq 2\rho_b, \qquad [a, b] \in R_0 . \tag{3.8}$$

To prove that there is exactly one, we conclude from the maximum-minimum principle for characteristic values [3], [7] that the absolute value of the ith characteristic value ν_a^i of (2.7) cannot decrease when a boundary condition at b is adjoined, and hence $|\nu_a^i| \leqq |\mu_{ab}^i|$ $(i = 1, 2, \cdots)$. Since the numbers ν_a^i do not accumulate and since $\rho_b \to 0$ as $b \to \omega_+$, there is a constant b_0 such that $2\rho_b$ is less than the minimum of all the differences $|\nu_a^j - \nu_a^i|$, $(i, j = 1, 2, \cdots; i \neq j)$ whenever $b \geqq b_0$. If $0 < \nu_a^1 < \nu_a^2$, it follows from (3.8) that exactly one characteristic value μ_{ab} of (2.5) lies in the interval $[\nu_a^1, \nu_a^1 + 2\rho_b]$. A similar statement applies to the case that one or both of ν_a^1, ν_a^2 are negative.

In order to prove by induction that there is exactly one μ_{ab}^i which satisfies $|\mu_{ab}^i - \nu_a^i| \leqq 2\rho_b$ $(i = 1, 2, \cdots)$, assume that this is true for each integer $i \leqq n$. In the case that $|\nu_a^{n+1}| < |\nu_a^{n+2}|$ there are at most $n + 1$ characteristic values μ_{ab}^i which satisfy $|\mu_{ab}^i| \leqq |\nu_a^{n+1}| + 2\rho_b$ since $|\mu_{ab}^i| \geqq |\nu_a^i|$ for each i. It then follows from the induction assumption that there is at most one characteristic value μ_{ab}^{n+1} satisfying $|\mu_{ab}^{n+1} - \nu_a^{n+1}| \leqq 2\rho_b$, and hence exactly one by (3.8). In the other case $\nu_a^{n+2} = -\nu_a^{n+1}$, it follows similarly that there are at most two μ_{ab}^i satisfying $|\nu_a^{n+1}| < |\mu_{ab}^i| \leqq |\nu_a^{n+1}| + 2\rho_b$, and again by (3.8) there is exactly one μ_{ab}^i near each of ν_a^{n+1}, ν_a^{n+2}.

THEOREM 1. *If the singularity* ω_+ *of* (1.1) *is the limit circle type, then for every characteristic value* ν_a *of* (2.7) *there exists a rectangle* R_0 *and a constant* C *on* R_0 *such that*[3] *a unique characteristic value* μ_{ab} *of the perturbed problem* (2.5) *lies in the interval* $|\mu_{ab} - \nu_a| \leqq C|U_b z_a|$ *whenever* $[a, b] \in R_0$.

This shows in particular that for each fixed a, there is a unique μ_{ab} of (2.5) such that $\mu_{ab} \to \nu_a$ as $b \to \omega_+$. In addition, the estimate of the theorem is valid uniformly on $\omega_- < a \leqq a_0$. One also finds for the characteristic functions y_{ab} and z_a associated with μ_{ab} and ν_a respectively that the estimate

$$||y_{ab} - z_a||_a^b \leqq C|U_b z_a|, \quad ||y_{ab}||_a^b = ||z_a||_a = 1$$

is valid on R_0.

4. Comparison of the z and x problems. The characteristic value problems (2.7) and (2.3) will now be compared, with (2.7) regarded as a perturbation of the basic problem (2.3). The perturbation arises from the singular end condition $[x\varphi_-](-) = 0$ being replaced by a homogeneous boundary condition at the point a. The novelty of this section is due to the singular nature of the unchanged endpoint ω_+.

Let λ be a characteristic value of (2.3) and let x be the corresponding normalized characteristic function. Let G_a be the linear integral operator on $\mathfrak{H}_{a\omega_+}$ whose kernel is the Green's function for kL_0 associated with the boundary conditions (2.6). This operator is defined similarly to the operator G_{ab} in (3.1) [6]. Let a function g on $[a, \omega_+)$ be defined by[4]

$$g = x - \gamma G_a x \qquad \text{where } \gamma = \lambda - l_0 . \tag{4.1}$$

The analogue of (3.5) is

$$g(s) = (U_a x / U_a \varphi_+)\, \varphi_+(s), \qquad a \leqq s < \omega_+ . \tag{4.2}$$

It follows from the postulated boundary conditions (2.2) at ω_- that $[x\varphi_-](a) \to 0$ as $a \to \omega_-$, and hence by (2.4) that $U_a x \to 0$ as $a \to \omega_-$. It was proved above (3.6) that $|U_a \varphi_+| \to 1$ as $a \to \omega_-$, and since $\varphi_+ \in \mathfrak{H}$, we obtain the inequality

$$||x - \gamma G_a x||_a \leqq C|U_a x|\, ||x||_a \tag{4.3}$$

for some constant C. The analogue of the lemma in §3 with $\mathfrak{H}_{ab}$ replaced by $\mathfrak{H}_{a\omega_+}$ leads to

$$\begin{aligned} ||x - P(\delta)x||_a &\leqq (1 + |\gamma|/\delta)\, ||x - \gamma G_a x||_a \\ &\leqq (1 + |\gamma|/\delta)\, C|U_a x|\, ||x||_a , \end{aligned}$$

[3] See footnote 2.

[4] The function on $[a, \omega_+)$ which coincides with x on this interval will also be denoted by x.

and the following theorem is obtained.

THEOREM 2. *If the singularities* $\omega_\pm$ *of* (1.1) *are both of the limit circle type, then for every characteristic value* λ *of the basic problem* (2.3) *there exist constants* a_0 *and* C *such that a unique characteristic value* ν_a *of* (2.7) *lies in the interval* $|\nu_a - \lambda| \leqq C\,|U_a x|$ *whenever* a *satisfies* $\omega_- < a \leqq a_0$. *If* x, z_a *are characteristic functions corresponding to* λ, ν_a *respectively with norms* $||x|| = ||z_a||_a = 1$, *then*

$$||z_a - x||_a \leqq C\,|U_a x|, \quad \omega_- < a \leqq a_0\,, \tag{4.4}$$

and in particular $||z_a - x||_a \to 0$ *as* $a \to \omega_-$.

We shall next prove the following consequence of (4.4):

$$U_b z_a = U_b x + (|U_a x| + |U_b x|)o(1)\,, \tag{4.5}$$

the order symbol being valid as $b \to \omega_+$ uniformly on $\omega_- < a \leqq a_0$. We use formula (1.4) to obtain

$$\begin{aligned} [z_a\varphi_+](+) - [z_a\varphi_+](b) &= (\nu_a - \bar{l}_0)(z_a, \varphi_+)_b\,, \\ [x\varphi_+](+) - [x\varphi_+](b) &= (\lambda - \bar{l}_0)(x, \varphi_+)_b\,. \end{aligned}$$

Since $[x\varphi_+](+) = [z_a\varphi_+](+) = 0$ by (2.2), (2.6), we deduce from the Schwarz inequality on $\mathfrak{F}_{b\omega_+}$ that

$$\begin{aligned} |[z_a\varphi_+](b) - [x\varphi_+](b)| &\leqq |(\nu_a - \bar{l}_0)(z_a - x, \varphi_+)_b| + |(\nu_a - \lambda)(x, \varphi_+)_b| \\ &\leqq |\nu_a - \bar{l}_0|\,||z_a - x||_b\,||\varphi_+||_b \\ &\quad + |\nu_a - \lambda|\,||x||\,||\varphi_+||_b\,. \end{aligned}$$

The desired conclusion (4.5) then follows from Theorem 2 and (2.4). The following abbreviation will be used:

$$\rho_{ab} = |U_a x| + |U_b x|\,. \tag{4.6}$$

THEOREM 3. *If both singularities* $\omega_\pm$ *are of the limit circle type, then for every characteristic value* λ *of* (2.3) *there exists a rectangle* R_0 *and a constant* C *on* R_0 *such that exactly one characteristic value* μ_{ab} *of the perturbed problem* (2.5) *lies in the interval* $|\mu_{ab} - \lambda| \leqq C\rho_{ab}$ *for every* $[a, b] \in R_0$. *For the characteristic functions* x, y_{ab} *associated with* λ, μ_{ab} *respectively, normalized by* $||x|| = ||y_{ab}||_a^b = 1$, *the estimate* $||y_{ab} - x||_a^b \leqq C\rho_{ab}$ *is valid.*

Proof. It follows from Theorems 1 and 2 that

$$\begin{aligned} |\mu_{ab} - \lambda| &\leqq |\mu_{ab} - \nu_a| + |\nu_a - \lambda| \\ &\leqq C(|U_b z_a| + |U_a x|)\,. \end{aligned}$$

The first statement of the theorem is then a consequence of (4.5) and (4.6). The proof of the second statement is similar and will be omitted.

Finally, we shall obtain uniform estimates for the difference $y_{ab}(s) - x(s)$ on $a \leqq s \leqq b$. We remark in passing that the asymptotic result $y_{ab}(s) = x(s)[1 + o(1)]$ as $a, b \to \omega_-, \omega_+$ cannot be valid for s near the boundaries a, b nor can it be valid near any zeros of $x(s)$. Uniform estimates will now be derived by the same technique that proves useful in certain domain-perturbed problems concerning elliptic partial differential equations [1], when $\varphi_\pm(s)$ are bounded on (ω_-, ω_+).

First it will be shown that $(\lambda - l_0)G_{ab}x(s)$ gives a uniform estimate for $y_{ab}(s)$ on $a \leqq s \leqq b$. Let $\psi_a(s)$ be the function (3.4) and let $\psi_b(s)$ be defined by

$$\psi_b(s) = \varphi_-(s)U_b\varphi_+ - \varphi_+(s)U_b\varphi_- .$$

Then

$$\begin{aligned} G_{ab}(s, t) &= \sigma^{-1}\psi_a(t)\,\psi_b(s) \quad \text{if } a \leqq t \leqq s \leqq b , \\ &= \sigma^{-1}\psi_a(s)\,\psi_b(t) \quad \text{if } a \leqq s \leqq t \leqq b , \end{aligned}$$

where

$$\sigma = U_a\varphi_-\, U_b\varphi_+ - U_a\varphi_+\, U_b\varphi_- .$$

Then $|\sigma| \to 1$ as $a, b \to \omega_-, \omega_+$, and the function defined by

$$(\|G_{ab}\|_a^b)^2 = \int_a^b |G_{ab}(s, t)|^2 k(t)dt$$

is a bounded function of s, a, and b. Hence the inequality

$$\begin{aligned} |y_{ab}(s) - (\lambda - l_0)G_{ab}x(s)| &= |G_{ab}[(\mu_{ab} - l_0)y_{ab}(s) - (\lambda - l_0)x(s)]| \\ &\leqq \|G_{ab}\|_a^b(|\mu_{ab} - l_0|\,\|y_{ab} - x\|_a^b + |\mu_{ab} - \lambda|\,\|x\|) , \end{aligned}$$

and Theorem 3 show that there exists a constant C such that

$$(4.6) \qquad |y_{ab}(s) - (\lambda - l_0)G_{ab}x(s)| \leqq C\rho_{ab} \quad a \leqq s \leqq b .$$

Let h be the uniquely determined solution of the boundary value problem

$$L_0h = 0, \quad U_ah = U_ax, \quad U_bh = U_bx \qquad \text{on } a \leqq s \leqq b .$$

Let the function f on $[a, b]$ be defined by

$$f(s) = (\lambda - l_0)G_{ab}x(s) - x(s) + h(s) .$$

Since f satisfies $L_0f = 0$, $U_af = U_bf = 0$, f is identically zero. The following uniform estimate is then a direct consequence of (4.6):

$$(4.7) \qquad y_{ab}(s) = x(s) - h(s) + O(\rho_{ab}), \quad a \leqq s \leqq b .$$

It can be verified without much difficulty that $h(s) = O(\rho_{ab})$ on a fixed closed subinterval I_0 of $[a, b]$, valid for $[a, b] \in R_0$. The following uniform result on I_0 is therefore a special case of (4.7):

$$y_{ab}(s) = x(s) + O(\rho_{ab}) \qquad [a, b] \in R_0 .$$

5. Class 1 boundary operators. Instead of the restrictive limiting behaviour (2.4) of the boundary operators U_a, U_b, the limiting behaviour of class 1 boundary operators is essentially arbitrary. In regard to the perturbation $a \to \omega_-$, a class 1 boundary operator U_a is defined as follows. Let φ_+ be the function defined in §2 and let x be a characteristic function of the basic problem (2.3) corresponding to the characteristic value λ. Class 1 perturbation problems are possible when the singularity ω_- is not an accumulation point of the zeros of φ_+ and

$$x(s)/\varphi_+(s) = o(1) \qquad \text{as } s \to \omega_- . \tag{5.1}$$

In this event, U_a is said to be a *class* 1 *boundary operator* on $(\omega_-, a_0]$ whenever the ratio $\varphi_+(a)U_a x/x(a)U_a\varphi_+$ is bounded on this interval. This rather mild restriction on U_a implies that

$$\varepsilon_a = |U_a x/U_a\varphi_+| = o(1) \qquad \text{as } a \to \omega_- . \tag{5.2}$$

An example is given in [8, pages 838–840] when $\omega_- = 0$ is a regular singularity of L, with $p(s) = 1$. In this event, a sufficient condition that the boundedness requirement above (5.2) be satisfied is that the limit $\sigma = \lim_{a\to 0} [a\alpha_0(a)/\alpha_1(a)]$ exists (finite or ∞) and $\sigma \neq -\rho$, where ρ denotes the smaller of two real, distinct exponents at the singularity 0.

Let g be defined by (4.1). Then (4.2) is valid but under the assumptions of this section, (4.3) is replaced by

$$||x - \gamma G_a x||_a \leqq C\varepsilon_a ||x||_a \tag{5.3}$$

where ε_a is defined by (5.2). In the notation of §§2, 3,

$$||x - P(\delta)x||_a \leqq (1 + |\gamma|/\delta)\, C\varepsilon_a ||x||_a .$$

Since $\varepsilon_a = o(1)$ as $a \to \omega_-$, Theorems 2 and 3 are valid with the replacement ε_a instead of $|U_a x|$. A similar statement is appropriate in the event that U_b is a class 1 boundary operator.

In the example of a regular singularity $\omega_- = 0$ with real exponents ρ_1, ρ_2, it turns out that $\varepsilon_a = O(a^{\rho_1 - \rho_2})$ if $\rho_1 > \rho_2$ and $\varepsilon_a = O(1/ln\, a)$ if $\rho_1 = \rho_2$ $(0 < a \leqq a_0)$.

REFERENCES

1. H. F. Bohnenblust et al, *The Laplacian operator on perturbed domains,* Unpublished manuscript.

2. E. A. Coddington and N. Levinson, *Theory of ordinary differential equations*, McGraw-Hill (1955).

3. R. Courant and D. Hilbert, *Methods of mathematical physics* I, Interscience (1953).

4. R. B. Dingle, *The solution of the Schrödinger equation for finite systems, etc.,* Proc. Camb. Philos. Soc., **49** (1953), 102-114.

5. T. E. Hull and R. S. Julius, *Enclosed quantum mechanical systems,* Canadian J. Phys., **34** (1956), 914-919. Additional references.

6. F. Rellich, *Spectral theory of a second order ordinary differential operator*, New York University (1953).

7. F. Riesz and B. Sz-Nagy. *Functional Analysis*, Blackie and Son (1956).

8. C. A. Swanson, *Differential operators with perturbed domains*, J. of Rat. Mech. Anal. **6** (1957), 823-846.

9. ———, *Asymptotic perturbation series for characteristic value problems*, Pacific J. Math., **9** (1959), 591-608.

10. ———, *Linear transformations with perturbed domains*, to appear in 1962.

THE UNIVERSITY OF BRITISH COLUMBIA

ON ESSENTIAL ABSOLUTE CONTINUITY

ROBERT J. THOMPSON

Throughout this paper $\boldsymbol{D}$ will denote a bounded domain in Euclidean n-space R^n, and T will be a bounded, continuous, single-valued transformation from $\boldsymbol{D}$ into R^n. For such transformations, concepts of essential bounded variation and essential absolute continuity have been defined and studied by Rado and Reichelderfer ([3], IV. 4). In this paper a characterization of essential absolute continuity will be given. The characterization suggests a definition of uniform essential absolute continuity and some of the consequences of this definition will be investigated.

1. For every point x in R^n a multiplicity function $K(x, T, \boldsymbol{D})$ is defined ([3], II. 3.2). T is said to be essentially of bounded variation (briefly eBV) in $\boldsymbol{D}$ provided $K(x, T, \boldsymbol{D})$ is Lebesgue summable in R^n ([3], IV. 4.1, Definition 1). Let $X_\infty = X_\infty(T, \boldsymbol{D})$ denote the set of points x in R^n for which $K(x, T, \boldsymbol{D})$ is infinite. Thus if T is eBV in $\boldsymbol{D}$, then $\mathscr{L}X_\infty = 0$ (if A is a subset of R^n, then $\mathscr{L}A$ denotes its exterior Lebesgue measure). Since $K(x, T, \boldsymbol{D})$ is a lower semicontinuous function of x ([3], II. 3.2, Remark 10), X_∞ is a Borel set and, by Theorem 1 of [3], IV. 1.1, the set $T^{-1} X_\infty$ is also a Borel set.

2. If x is a point in R^n and C is a component of $T^{-1}x$ which is closed relative to R^n, then C is termed a maximal model continuum $(x, T, \boldsymbol{D})$ ([3], II. 3.1, Definition 1). Denote by $\mathfrak{C} = \mathfrak{C}(T, \boldsymbol{D})$ the class composed of all sets C for which TC is a point in R^n and C is a maximal model continuum for $(TC, T, \boldsymbol{D})$. Let $\mathfrak{E} = \mathfrak{E}(T, \boldsymbol{D})$ be the subset of $\mathfrak{C}$ consisting of those elements C each of which is an essential maximal model continuum (briefly e.m.m.c.) for $(TC, T, \boldsymbol{D})$ ([3], II. 3.3, Definition 1); the set $E = E(T, \boldsymbol{D}) = \cup\, C, C \in \mathfrak{E}$ ([3], II. 3.6). Let $\mathfrak{E}_i = \mathfrak{E}_i(T, \boldsymbol{D})$ be the subset of $\mathfrak{E}$ consisting of those elements C each of which is an essentially isolated e.m.m.c. (briefly e.i. e.m.m.c.) for $(TC, T, \boldsymbol{D})$ ([3], II. 3.3, Definition 2); the set $E_i = E_i(T, \boldsymbol{D}) = \cup\, C, C \in \mathfrak{E}_i$ ([3], II. 3.6.). Finally, let $\mathfrak{E}_i^p = \mathfrak{E}_i^p(T, \boldsymbol{D})$ be the subset of $\mathfrak{E}_i$ consisting of those elements of $\mathfrak{E}_i$ which consist of single points; the set $E_i^p = E_i^p(T, \boldsymbol{D}) = \cup\, C, C \in \mathfrak{E}_i^p$ ([3], II. 3.6). The sets E, E_i and E_i^p are Borel sets ([3], II. 3.6, Theorem 1).

If T is eBV in $\boldsymbol{D}$, then a necessary and sufficient condition that T be essentially absolutely continuous (briefly eAC) in $\boldsymbol{D}$ ([3], IV. 4.2) is

Received December 15, 1960. The results reported here were included in a dissertation presented in partial fulfillment of the requirements for the degree Doctor of Philosophy at The Ohio State University. The author wishes to express his gratitude to Professor P. V. Reichelderfer for his generous help and advice.

that T satisfies the condition (N) on the set $E(T, \mathbf{D})$ ([3], IV. 4.2, Theorem 3) i.e., if $S \sqsubseteq E$ and $\mathscr{L}S = 0$, then $\mathscr{L}TS = 0$.

DEFINITION 1. T will be said to satisfy the (ε, δ) condition on a subset A of $\boldsymbol{D}$ if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that if $S \sqsubseteq A$ and $\mathscr{L}S < \delta$, then $\mathscr{L}TS < \varepsilon$. Clearly if T satisfies the (ε, δ) condition on each of a finite number of subsets of $\boldsymbol{D}$, then T satisfies the (ε, δ) condition on any subset of their union. Also, if A is a Borel subset of $\boldsymbol{D}$, then T satisfies the (ε, δ) condition on A if and only if for every $\varepsilon > 0$ there is a $\delta > 0$ such that if S is a Borel subset of A and $\mathscr{L}S < \delta$, then $\mathscr{L}TS < \varepsilon$.

THEOREM 1. *Suppose T is eBV in $\boldsymbol{D}$. Then a necessary and sufficient condition that T be eAC in $\boldsymbol{D}$ is that T satisfies* the (ε, δ) *condition on the set $E(T, \boldsymbol{D})$.*

Proof. Since T is assumed to be eBV in $\boldsymbol{D}$ it suffices to prove that a necessary and sufficient condition that T satisfies the condition (N) on the set E is that T satisfies the (ε, δ) condition on E. Since the proof of the sufficiency is immediate, we proceed to a proof of the necessity. If T satisfies the condition (N) on E, then, by Lemma 4 of [3], IV. 4.2, $\mathscr{L}T(E - E_i^p) = 0$ and so T clearly satisfies the (ε, δ) condition on $E - E_i^p$. Since T is eBV in $\boldsymbol{D}$, $\mathscr{L}X_\infty = 0$ and so T satisfies the (ε, δ) condition on $T^{-1}X_\infty$. Since E is a subset of the union of the sets $E - E_i^p$, $T^{-1}X_\infty$ and $E_i^p - T^{-1}X_\infty$, in view of the remarks following Definition 1 it remains only to be shown that T satisfies the (ε, δ) condition on $E_i^p - T^{-1}X_\infty$ whenever T satisfies the condition (N) on E. Assume then that T does not satisfy the (ε, δ) condition on $E_i^p - T^{-1}X_\infty$. The proof will be completed by showing that T does not satisfy the condition (N) on E. Since E_i^p and $T^{-1}X_\infty$ are Borel sets, their difference is a Borel set. Thus the assumption that T fails to satisfy the (ε, δ) condition on $E_i^p - T^{-1}X_\infty$ implies, in view of the remarks following Definition 1, that there is an $\varepsilon_0 > 0$ such that for every positive integer k there is a Borel set $S_k \sqsubseteq E_i^p - T^{-1}X_\infty$ such that $\mathscr{L}S_k < 1/2^k$ and $\mathscr{L}TS_k \geqq \varepsilon_0$. Let $S^* = \limsup S_k$ $(= \bigcap_{n=1}^{\infty} \bigcup_{k \geqq n} S_k)$. S^* is a subset of $E_i^p - T^{-1}X_\infty$ and so

$$S^* \sqsubseteq E. \tag{1}$$

For every positive integer n, $S^* \sqsubseteq \bigcup_{k \geqq n} S_k$ and so $\mathscr{L}S^* \leqq 1/2^{n-1}$. Hence

$$\mathscr{L}S^* = 0. \tag{2}$$

Let k be a positive integer and suppose $x \in TS_k$. Since $S_k \sqsubseteq E_i^p - T^{-1}X_\infty$, $K(x, T, \boldsymbol{D}) < \infty$ and there is a point u in E_i^p such that $Tu = x$, Since $K(x, T, \boldsymbol{D}) < \infty$ there are at most a finite number of e.m.m.c.s. for $(x, T, \boldsymbol{D})$ ([3], II. 3.3, Definition 1 and II. 3.4, Theorem 3). But for every point u in E_i^p such that $Tu = x$ the set consisting of the point u is an e.m.m.c. for $(x, T, \boldsymbol{D})$. Thus there are at most a finite number of points u in $E_i^p - T^{-1}X_\infty$ for which $Tu = x$. Thus it has been shown that

(3) For every integer k, if x is in TS_k then $(E_i^p - T^{-1}X_\infty) \cap T^{-1}x$ is a finite set.

Since $\bigcup S_k \subseteq E_i^p - T^{-1}X_\infty$ it is easy to show that (3) implies that $\lim\sup TS_k = T(\lim\sup S_k)$ and so

$$(4) \qquad \mathscr{L}(\limsup TS_k) = \mathscr{L}TS^*.$$

By Theorem 4 of [3], IV. 1. 1, the sets TS_k are measurable. Since T is a bounded transformation, $\mathscr{L}(\cup TS_k)$ is finite. Thus ([5], p. 17)

$$(5) \qquad \mathscr{L}(\limsup TS_k) \geqq \limsup \mathscr{L}TS_k.$$

But $\mathscr{L}TS_k \geqq \varepsilon_0 > 0$ for all k and so

$$(6) \qquad \limsup \mathscr{L}TS_k > 0$$

By (4), (5) and (6),

$$(7) \qquad \mathscr{L}TS^* > 0$$

Now (1), (2) and (7) imply that T does not satisfy condition (N) on E.

3. Definition 2. For every positive integer j let $\boldsymbol{D}_j$ be a bounded domain in R^n and let T_j be a bounded, continuous, single-valued transformation from $\boldsymbol{D}_j$ into R^n. The transformations T_j will be termed uniformly essentially absolutely continuous (briefly UEAC) provided:

(i) For each j, $T_j \, eBV$ in $\boldsymbol{D}_j$ and

(ii) Given any $\varepsilon > 0$, there is a $\delta > 0$, depending only on ε, such that for all j the following is true: if S is a subset of $E(T_j, \boldsymbol{D}_j)$ and $\mathscr{L}S < \delta$, then $\mathscr{L}T_jS < \varepsilon$.

Note that if the transformations T_j are UEAC, then, by Theorem 1, for each j, T_j is eAC in $\boldsymbol{D}_j$.

Each point u in $\boldsymbol{D}$ is contained in a unique component of $T^{-1}Tu$ denoted by C_u. A subset U of $\boldsymbol{D}$ is termed a T set if $u \in U$ implies $C_u \subseteq U$ ([4], 1).

Theorem 2. *Let $\boldsymbol{D}$ be a bounded domain in Euclidean n-space R^n and let T be a bounded, continuous, single-valued transformation from $\boldsymbol{D}$ into R^n. For every positive integer j let $\boldsymbol{D}_j$ be a bounded domain in R^n and let T_i be a bounded, continuous, single-valued transformation from $\boldsymbol{D}_j$ into R^n,* *If*

(i) *The mappings* T_j *are* UEAC

(ii) *The mappings* T_j *converge to* T *uniformly on compact subsets of* $\boldsymbol{D}$ ([3], II. 3. 2, Remark 9) *and*

(iii) A *is a* T *set contained in* $E(T, \boldsymbol{D})$ *and* $\mathscr{L}A = 0$,

then $\mathscr{L}TA = 0$.

Proof Let $\varepsilon > 0$ be given and let δ be the corresponding positive number in (ii) of Definition 2. Since A is a subset of the open set $\boldsymbol{D}$ and $\mathscr{L}A = 0$, there is an open set O, containing A and contained in $\boldsymbol{D}$, such that $\mathscr{L}O < \delta$. Let $x \in TA$. Since $A \subseteq E(T, \boldsymbol{D})$, there is a set C, e.m.m.c. for $(x, T, \boldsymbol{D})$, such that C meets A. $C \subseteq A$ since A is a T set and so $C \subseteq O$. By Definition 1 in [3], II. 3.3 there is a set D, which contains C and whose closure $\mathscr{K}D$ is contained in O, such that D is an indicator domain for $(x, T, \boldsymbol{D})$ ([3], II. 3.2). By definition $\mathscr{K}D \subseteq \boldsymbol{D}$, x is not in $T\mathscr{B}D$ (where $\mathscr{B}D$ denotes the boundary of D) and the topological index $\mu(x, T, D)$ ([3], II. 2) is not zero. Since $T\mathscr{B}D$ is compact, the ecart of x from $T\mathscr{B}D$, $e(x, T\mathscr{B}D)$, is positive ([3], I.1.4, Exercise 3). Since $\mathscr{K}D \subseteq \boldsymbol{D}$, by (ii) there is a positive integer j_x such that, for $j > j_x$, $\mathscr{K}D \subseteq \boldsymbol{D}_j$ and $\rho(T, T_j, \mathscr{K}D)$ the deviation of T_j from T on $\mathscr{K}D$ ([3], I. 1.5, Definition 5) is less than $e(x, T\mathscr{B}D)$. Clearly $\rho(T, T_j, \mathscr{B}D) \leqq \rho(T, T_j, \mathscr{K}D)$. Thus, for $j > j_x$, $\mathscr{K}D \subseteq \boldsymbol{D} \cap \boldsymbol{D}_j$ and $\rho(T, T_j, \mathscr{B}D) < e(x, T\mathscr{B}D)$. By Theorem 6 of [3], II. 2.3, $\mu(x, T_j, D)$ is defined and equals $\mu(x, T, D)$. Thus D is an indicator domain for $(x, T_j, \boldsymbol{D}_j)$ and, by Lemma 4 of [3], II. 3.3, there is a set C_j, e.m.m.c. for $(x, T_j, \boldsymbol{D}_j)$, such that $C_j \subseteq D$. Now $C_j \subseteq O \cap E(T_j, \boldsymbol{D}_j)$ and $T_jC_j = x$. Thus $x \in T_j[O \cap E(T_j, \boldsymbol{D}_j)]$ for all $j > j_x$ and hence $x \in \liminf T_j[O \cap E(T_j, \boldsymbol{D}_j)]$. Since x was any point in TA, it has been shown that $TA \subseteq \liminf T_j[O \cap E(T_j, \boldsymbol{D}_j)]$ and so

$$\mathscr{L}TA \leqq \mathscr{L}\liminf T_j[O \cap E(T_j, \boldsymbol{D}_j)]. \tag{1}$$

Since $E(T_j, \boldsymbol{D}_j)$ is a Borel set, $O \cap E(T_j, \boldsymbol{D}_j)$ is also a Borel set and so $T_j[O \cap E(T_j, \boldsymbol{D}_j)]$ is Lebesgue measurable. Thus ([5], p. 17)

$$\mathscr{L}\liminf T_j[O \cap E(T_j, \boldsymbol{D}_j)] \leqq \liminf \mathscr{L}T_j[O \cap E(T_j, \boldsymbol{D}_j)]. \tag{2}$$

Now

$$\mathscr{L}[O \cap E(T_j, \boldsymbol{D}_j)] \leqq \mathscr{L}O < \delta. \tag{3}$$

By the choice of δ, (3) implies that $\mathscr{L}T_j[O \cap E(T_j, \boldsymbol{D}_j)] < \varepsilon$ and hence

$$\liminf \mathscr{L}T_j[O \cap E(T_j, \boldsymbol{D}_j)] \leqq \varepsilon. \tag{4}$$

By (1), (2) and (4)

$$(5) \qquad \mathscr{L}TA \leqq \varepsilon.$$

Since (5) has been proved for an arbitrary $\varepsilon > 0$, it follows that $\mathscr{L}TA = 0$.

4. Theorem 2 suggests the question: under the hypotheses of Theorem 2 does T satisfy the condition (N) on $E(T, \boldsymbol{D})$? Note that T does satisfy the condition (N) on $E_i^p(T, \boldsymbol{D})$. In the remainder of the paper some results pertinent to this question will be presented.

Reichelderfer introduced the concept of the T magnification ([4], 6). It will be useful to have the definition repeated here.

Let $\mathfrak{D}^* = \mathfrak{D}^*(T, \boldsymbol{D})$ be the class composed of all domains D for each of which $\mathscr{K}D$ is contained in $\boldsymbol{D}$ and there exists an open oriented n-cube Q in R^n such that D is a component of $T^{-1}Q$. If C is a maximal model continuum for $(x, T, \boldsymbol{D})$ for some point x in R^n, for every positive number ε define

$$\bar{d}(C, \mathscr{L}T, \varepsilon) = \text{l.u.b.}\ \mathscr{L}TD/\mathscr{L}D,\ C \mathrel{\overline{\overline{\subset}}} D \in \mathfrak{D}^*,\ \delta TD \leqq \varepsilon$$

and

$$\underline{d}(C, \mathscr{L}T, \varepsilon) = \text{g.l.b.}\ \mathscr{L}TD/\mathscr{L}D,\ C \mathrel{\overline{\overline{\subset}}} D \in \mathfrak{D}^*,\ \delta TD \leqq \varepsilon$$

(If A is a subset of R^n, δA denotes the diameter of A).

$$\bar{d}(C, \mathscr{L}T) = \lim_{\varepsilon \to 0+} \bar{d}(C, \mathscr{L}T, \varepsilon)$$

and

$$\underline{d}(C, \mathscr{L}T) = \lim_{\varepsilon \to 0+} \underline{d}(C, \mathscr{L}T, \varepsilon).$$

If $\bar{d}(C, \mathscr{L}T)$ and $\underline{d}(C, \mathscr{L}T)$ are finite and equal, their common value is denoted by $M(C, T)$ and is termed the T magnification at C.

Lemma 1. *Let p be a positive number and let A be a T set with the following properties*:

(i) *If $u \in A$, then there is a set $C \in \mathfrak{C}_i(T, \boldsymbol{D})$ such that $u \in C$ and $\underline{d}(C, \mathscr{L}T) > p$.*

(ii) *If $C \in \mathfrak{C}_i(T, \boldsymbol{D})$ and $C \mathrel{\overline{\overline{\subset}}} A$, then for every domain G in R^n which contains TC and has a sufficiently small diameter it is true that $T^{-1}G$ possesses exactly one component D which meets A. Note that D must contain C and (provided only that the diameter of G is sufficiently small) be a* m.i.d. T ([4], 4 *and* 5, *Lemma* 2).

Then $\mathscr{L}A \leqq 1/p\, \mathscr{L}TA$.

Proof. Let η be any positive number. The proof will be completed by showing that $\mathscr{L}A \leqq 1/p\,\mathscr{L}TA + \eta$.

Let $x \in TA$ (the inequality is trivial if A is empty) and let $u \in A$ such the $Tu = x$. By (i) there is a set $C \in \mathfrak{C}_i(T, \boldsymbol{D})$ such that $u \in C$ and $\underline{d}(C, \mathscr{L}T) > p$. Thus there is an $\varepsilon > 0$ such that $\underline{d}(C, \mathscr{L}T, \varepsilon) > p$ and so

$$(1) \qquad \text{If } C \sqsubseteq D \in \mathfrak{D}^* \text{ and } \delta\, TD \leqq \varepsilon, \text{ then } \mathscr{L}TD/\mathscr{L}D > p$$

Since A is a T set, $C \sqsubseteq A$ and, by (ii), there exists a positive number r such that for every domain G in R^n which contains $TC(=x)$ and for which $\delta G \leqq r$ it is true that $T^{-1}G$ possesses exactly one component which meets A and, moreover, this component is a m.i.d. T containing C. For every positive integer i let Q_i be the open oriented n-cube with center at x and diameter equal to the smaller of ε, r and $1/i$. Then $T^{-1}Q_i$ possesses exactly one component D_i which meets A and D_i is a m.i.d. T containing C. By the Lemma in [4], 4, $TD_i = Q_i$ and $\mathscr{K}D_i \sqsubseteq D$. By definition, $D_i \in \mathfrak{D}^*$ and so, with the aid of (1), $\mathscr{L}D_i < 1/p\,\mathscr{L}TD_i$. Thus

(2) For every point x in TA there is associated a sequence of open oriented n-cubes Q_i with centers at x and a corresponding sequence of domains D_i such that, for all i, $\delta Q_i \leqq 1/i$, $\mathscr{L}D_i, < 1/p\,\mathscr{L}Q_i, D_i$ is a component of $T^{-1}Q_i$ and the only component of $T^{-1}Q_i$ which meets A.

Let $\mathfrak{Q}$ be the class of all n-cubes associated with points of TA in this manner. $\mathscr{L}TA$ is finite since T is bounded, and by a theorem of Rademacher ([2], p. 190) there is a $\mathfrak{Q}^*$, countable subclass of $\mathfrak{Q}$, such that

$$(3) \qquad TA \sqsubseteq \cup\, Q^*, Q^* \in \mathfrak{Q}^*$$

and

$$(4) \qquad \Sigma \mathscr{L}Q^* \leqq \mathscr{L}TA + \eta p.$$

(Rademacher's theorem is stated in terms of a covering made up of open n-spheres, but the corresponding theorem for a covering of open n-cubes is readily obtained from it). Let Q^* be an element of $\mathfrak{Q}^*$. By (2) there is a corresponding domain D^*, D^* a component $T^{-1}Q^*$ such that $\mathscr{L}D^* < 1/p\,\mathscr{L}Q^*$ and D^* is the only component of $T^{-1}Q^*$ which meets A. In this way exactly one domain D^* is associated with each $Q^* \in \mathfrak{Q}^*$. The class of domains D^* is countable and

$$(5) \qquad \Sigma \mathscr{L}D^* \leqq 1/p\, \Sigma \mathscr{L}Q^*.$$

Let $u \in A$. Then $Tu \in TA$ and by (3) there is a $Q^* \in \mathfrak{Q}^*$ such that $Tu \in Q^*$. Since the corresponding D^* is the only component of $T^{-1}Q^*$

which meets A it must contain u. Thus $A \subseteq \cup D^*$ and

$$(6) \qquad \mathscr{L}A \leqq \Sigma \mathscr{L} D^*.$$

By (4), (5) and (6), $\mathscr{L}A \leqq 1/p\, \mathscr{L}TA + \eta$. Since η is any positive number, the conclusion of the lemma is established.

LEMMA 2. *Let $\mathfrak{H}$ be a subclass of $\mathfrak{C}_i(T, \boldsymbol{D})$ such that if $C \in \mathfrak{H}$ then $\underline{d}(C, \mathscr{L}T) > 0$. Put $H = \cup C, C \in \mathfrak{H}$. If $\mathscr{L}TH = 0$, then $\mathscr{L}H = 0$.*

Proof. If H is not empty (the equality is trivial otherwise) then $\mathfrak{C}_i(T, \boldsymbol{D})$ is not empty and hence, by the Lemma in [4], 14, the set E_i can be expressed as the union of a countably infinite sequence of T sets U_k with the following property:

(1) If $C \in \mathfrak{C}_i$ and $U_k \supseteq C$, then for every domain G in R^n which contains TC and has a sufficiently small diameter it is true that $T^{-1}G$ possesses exactly one component D which meets U_k.

For every positive integer n let $\mathfrak{H}_n$ be the subclass of $\mathfrak{H}$ consisting of those elements C for which $\underline{d}(C, \mathscr{L}T) > 1/n$. Put $H_n = \cup C, C \in \mathfrak{H}_n$ and let $H_{nk} = H_n \cap U_k$. Then $H = \cup H_n$ and, for each n, $H_n = \cup H_{nk}$. The proof will be completed by showing that $\mathscr{L}H_{nk} = 0$ for arbitrary n and k. Since H_n and U_k are T sets,

(2) H_{nk} is a T set.

Clearly

(3) If $u \in H_{nk}$, then there is a set $C \in \mathfrak{C}_i$ such that $u \in C$ and $\underline{d}(C, \mathscr{L}T) > 1/n$.

By (1) and the definition of H_{nk},

(4) If $C \in \mathfrak{C}_i$ and $C \subseteq H_{nk}$, then for every domain G in R^n which contains TC and has a sufficiently small diameter it is true that $T^{-1}G$ possesses exactly one component D which meets H_{nk}.

(2), (3), (4) and Lemma 1 imply that $\mathscr{L}H_{nk} \leqq n\mathscr{L}TH_{nk}$. Since $TH_{nk} \subseteq TH$ and $\mathscr{L}TH = 0$, $\mathscr{L}TH_{nk} = 0$ and consequently $\mathscr{L}H_{nk} = 0$. Since n and k are arbitrary, it follows that $\mathscr{L}H = 0$.

5. THEOREM 3. *Let $\boldsymbol{D}$ be a bounded domain in Euclidean n-space R^n and let T be a bounded, continuous, single-valued transformation from $\boldsymbol{D}$ into R^n. For every positive integer j let $\boldsymbol{D}_j$ be a bounded domain in R^n and let T_j be a bounded, countinuous, single-valued transformation from $\boldsymbol{D}_j$ into R^n. Let $\mathfrak{B}$ be the subclass of $\mathfrak{C}_i(T, \boldsymbol{D})$*

consisting of those elements C *for each of which* C $M(C, T)$ *exists and is positive and* C *contains more than a single point. Put* $B = \cup C$, $C \in \mathfrak{B}$. *If*

(i) *The mappings* T_j *are UEAC.*

(ii) *The mappings* T_j *converge to* T *uniformly on compact subsets* $\boldsymbol{D}$ *and*

(iii) T *is eBV in* $\boldsymbol{D}$

then the following statements are equivalent:

(iv) T *satisfies the condition* (N) *on* B,

(iv)′ $\mathscr{L} TB = 0$ *and*

(iv)″ $\mathscr{L} B = 0$

and (i), (ii) *and* (iii) *together with* (iv) *or* (iv)′ *or* (iv)″ *imply that* T *is eAC in* $\boldsymbol{D}$.

Proof. First it will be shown that (i), (ii), (iii) and (iv) imply that T is *eAC* in $\boldsymbol{D}$. By the Theorem in [4], 16, there exist T sets V' and V'' contained in $\boldsymbol{D}$ such that $\mathscr{L} V' = 0$, $\mathscr{L} TV'' = 0$ and if $C \in \mathfrak{C}_i(T, \boldsymbol{D})$ and C does not meet $V' \cup V''$, then $M(C, T)$ exists and is positive. In view of (iii), in order to conclude that T is *eAC* in $\boldsymbol{D}$ it is sufficient to prove that T satisfies the condition (N) on $E = E(T, \boldsymbol{D})$. Clearly it is sufficient to show that T satisfies the condition (N) on each of the following sets whose union is E: $S_1 = E - E_i$, $S_2 = E_i^p$, $S_3 = (E_i - E_i^p) \cap V'$, $S_4 = (E_i - E_i^p) \cap V''$ and $S_5 = (E_i - E_i^p) - (V' \cup V'')$. Since T is *eBV* in $\boldsymbol{D}$, $\mathscr{L} TS_1 = 0$ (this is proved in the first step in the proof of the theorem in [4], 18) and so T satisfies the condition (N) on S_1. Any subset of S_2 is a T set contained in E and it follows by Theorem 2 that T satisfies the condition (N) on S_2. Again by Theorem 2, $\mathscr{L} TS_3 = 0$ and so T satisfies the condition (N) on S_3. $\mathscr{L} TS_4 \leqq \mathscr{L} TV'' = 0$ and so T satisfies the condition (N) on S_4. S_5 is a subset of B and so (iv) implies that T satisfies condition (N) on S_5.

If (i), (ii), (iii) and (iv) are satisfied, then it has just been shown that T satisfies the condition (N) on $E(T, \boldsymbol{D})$. Hence, by Lemma 4 of [3], IV. 4.2, $\mathscr{L} T(E - E_i^p) = 0$. Since B is a subset of $E - E_i^p$, (iv)′ must be satisfied. On the other hand, (iv)′ clearly implies (iv). Thus if (i), (ii) and (iii) are satisfied, (iv) and (iv)′ are equivalent.

By Lemma 2, $\mathscr{L} B = 0$ if $\mathscr{L} TB = 0$. On the other hand, since B is a T set contained in $E(T, \boldsymbol{D})$, (i) and (ii) imply, by Theorem 2, that $\mathscr{L} TB = 0$ if $\mathscr{L} B = 0$. Hence if (i) and (ii) are satisfied, then (iv)′ and (iv)″ are equivalent.

6. It is reasonable to inquire whether (i), (ii) and (iii) in Theorem 3 are sufficient to conclude that T is *eAC* in $\boldsymbol{D}$. After all, each of the sets C in $\mathfrak{B}$ is a non-point continuum for which the T magnification is positive and yet whose image under T is a single point in R^n. Might not (i), (ii) and (iii) imply, say, (iv)′ (or equivalently (iv) or (iv)″)? Since the class $\mathfrak{B}$ is clearly countable when T is a transformation into R^1, TB is then a countable set. Thus (iv)′ is always satisfied when T is a transformation into R^1. However, the author has constructed an example in R^2 for which (i), (ii) and (iii) are satisfied and for which the limit transformation is not eAC ([6]). In the example the limit transformation T is modeled on an example by Cesari ([1], IV. 13.1, Example A). The transformation that Cesari defined provides an example of a plane mapping that is eBV but not eAC. The example in [6] is somewhat more complicated by the need for (i) and (ii) to be satisfied.

BIBLIOGRAPHY

1. L. Cesari, *Surface Area*, Annals of Mathematics Studies, No. 35, Princeton University Press, 1956.

2. H. Rademacher, *Eineindeutige Abbildungen und Messbarkeit*, Monatshefte fur Mathematik und Physik, **27** (1916), 183-290.

3. T. Rado, and P. Reichelderfer, *Continuous Transformations in Analysis*, Die Grundlehren der Mathematischen Wissenschaften, Vol. 75, Springer-Verlag, 1955.

4. P. Reichelderfer, *A Study of the Essential Jacobian in Transformation Theory*, Rendiconti del Circolo Matematico di Palermo, Series 2, **6** (1957), 175-197.

5. S. Saks, *Theory of the Integral*, Monografie Matematyczne, Warsaw, 1937.

6. R. Thompson, *On Essential Absolute Continuity for a Transformation*, Dissertation, The Ohio State University, 1958 (L. C. Card No. Mic 58-3467).

SANDIA CORPORATION, ALBUQUERQUE, NEW MEXICO

Correction to

"ON TERMINATING PROLONGATION PROCEDURES"

BY H. H. JOHNSON

This Journal, Vol. 10 (1960), 577-583

(Received June 8, 1961)

M. Kuranishi has kindly brought to our attention an error in Theorem 1 on page 579. Condition (2) of that theorem should be corrected to read:

$$\text{``(2)}\quad dB_{\varphi;ij;k_1\cdots k_t} \equiv 0 \textit{ modulo } (\omega^i, \theta^\alpha) \textit{ for all } t.\text{''}$$

The above equation does not follow from the original hypotheses as the author indicated.

Since the interest in Theorem 1 is in its applicability as a criterion for involutiveness, it may be helpful to mention the following conditions under which (2) holds, assuming condition (1).

Condition 1. The θ^α and ω^i span $dx^1, \cdots, dx^n$.

Condition 2. $\omega^i = dy^i$, $i = 1, \cdots, p$ and $dB_{\varphi;ij} \equiv 0$ modulo (ω^i).

Under Condition 1, there are no π^λ, hence no additional variables are introduced by the prolongation process.

Under Condition 2, $B_{\varphi;ij}$ is a function of $y^1, \cdots, y^p$ alone. Consequently $dB_{\varphi;ij} = (\partial B_{\varphi;ij}/\partial y^k)\omega^k$, hence $B_{\varphi;ij;k} = (\partial B_{\varphi;ij}/\partial y^k)$ is also a function of $y^1, \cdots y^p$ alone. In the same way every $B_{\varphi;ij;k_1\cdots k_t}$ is a function of $y^1, \cdots, y^p$ alone.

Condition 1 is satisfied in Theorem 2 on page 581. Condition 2 is satisfied in the system (S') on page 220 studied in the paper, H. H. Johnson: "On the pseudo-group structure of analytic functions on an algebra," Proc. Amer. Math. Soc. 12 (1961), 218-224. Princeton University.

Pacific Journal of Mathematics
Vol. 11, No. 4 , 1961