Pacific Journal of Mathematics

IN THIS ISSUE—

Vol. 8, No. 4 **June, 1958**

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, E. G. Straus 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, 2120 Oxford Street, Berkeley 4, 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.

THE MAXIMAL IDEALS OF CERTAIN FUNCTION ALGEBRAS

RICHARD ARENS

1. Introduction. In this paper we discover the space of maximal ideals for each Banach algebra of the following concrete type. Select an open subset G of S, the compactified complex plane, and let $H(G)$ be the class of complex functions continuous on S and moreover holomorphic on G. This is a Banach algebra, and its space of maximal ideals is shown below to be precisely S, except in that case in which G is so large as to force $H(G)$ to consist only of constant functions.

Algebras of this type were introduced and studied by J. Wermer [4] and W. Rudin [3]. Wermer pointed out that $H(G)$ need not reduce to the constants even if $S - G$ is required to be (merely) an arc. He also showed distinct points of S determined distinct maximal ideals. Rudin raised the question as to the space of maximal ideals.

K. M. Hoffman, reporting (April 18, 1958, Symposium on Banach algebras and Harmonic analysis) on work by I. M. Singer and himself jointly, showed that the space of maximal ideals of $H(G)$ is S when $S - G$ has positive upper density at each of its points. On the following day, H. L. Royden's proof was presented in which the same desired conclusion was obtained if $S - G$ has dimension zero. Our technique may be regarded as a refinement of Hoffman and Singer's.

Our methods apply equally easily to more general, although perhaps less interesting, algebras. Let Z be a compact subset of S, and let G be an open subset of Z. Let $H(G/Z)$ be the functions continuous on Z and holomorphic on G. Then Z is the space of maximal ideals, unless the algebra reduces to the constants.

For some algebras in this larger class, the problem can also be solved by an appeal to Mergelyan's theorem [5], namely for those $H(G/Z)$ where $Z \neq S$ and G is the interior of Z.

2. An approximation theorem. Let Z be a Borel set in the extended complex plane. Let G be an open set included in Z. We denote by $H(G/Z)$ the class of complex-valued functions which are defined, continuous, and bounded on Z, and are holomorphic on G. $H(G/Z)$ is evidently a Banach algebra with unit, providing that for each $f \in H(G/Z)$ the norm is defined by

$$\|f\| = \sup_{\zeta \in Z} |f(\zeta)| .$$

Received May 21, 1958.

In our study of the maximal ideals of such an algebra, we have been led to an approximation Theorem 2.7 involving functions defined as follows

$$h(z) = \int \frac{f(\zeta)}{\zeta - z} \mu(d\zeta) \tag{2.1}$$

where

(2.1.1) *μ is plane Lebesgue measure restricted to a disc, and then normalized to make the measure of that disc equal to* 1.

However, it happens that functions defined as in 2.1 are useful in studying $H(G/Z)$ in still another way. They can sometimes be used to show that there are non-constant elements in $H(G/Z)$. We consider this matter again briefly in § 4. The properties of h needed for this purpose naturally suggest a condition on μ, namely that given by 2.2.3 (which involves 2.2, 2.1.2). Now it turns out that with no added effort, and very little loss of clarity, a generalization 2.6 of our real objective, 2.7, can be proved which involves only the quantity $I(\mu)$ of 2.2.3, and hence is not confined to the case 2.1.1. It is hoped that some use for the approximation Theorem 2.6 may emerge.

Let μ be a regular Borel measure in the plane, finite on bounded sets. For $r \geqq 0$ define

$$m(r) = \sup_{\zeta} \mu\{|z - \zeta| < r\} . \tag{2.1.2}$$

Let $\|\mu\| = \sup_r m(r)$; and let

$$J(\mu) \tag{2.2}$$

be the least upper bound, for $0 < t_1 < t_2 < \cdots < t_n$ of the sums

$$\frac{m(t_1)}{t_1} + \frac{m(t_2) - m(t_1)}{t_2} + \cdots + \frac{m(t_n) - m(t_{n-1})}{t_n} . \tag{2.2.1}$$

This is nothing but the Stieltjes integral

$$\int_0^\infty r^{-1} dm(r) . \tag{2.2.2}$$

The class B of measures with which we shall deal are those for which

$$I(\mu) < \infty . \tag{2.2.3}$$

There are measures in B with 0-dimensional support (see § 4). For our immediate purpose, those given by 2.1.1 are the most important. (Their support is, of course, 2-dimensional). We note the relation of $I(\mu)$ and $\|\mu\|$ in this case.

2.3 LEMMA. *Let D be a disc of positive radius δ in the plane. For* *each Borel set E let $\mu(E)$ be the plane measure of $E \cap D$. Then $\delta \cdot I(\mu) = 2 \|\mu\|$, and $\|\mu\| = \pi\delta^2$.*

Some properties of 2.1 will now be described.

2.4 LEMMA. *Let $\mu \in B$, and let f be a bounded Borel-measurable function defined on the plane. Then*

$$h(z) = \int \frac{f(\zeta)}{\zeta - z} \mu(d\zeta)$$

defines a continuous function on the plane. It is holomorphic on each open set of μ-measure 0. If μ has bounded support, h is holomorphic, and 0, at ∞; and h is bounded.

Proof. We treat first the case $f = 1$. We will show now that the set functions defined by

$$F_z(E) = \int_E \frac{\mu(d\zeta)}{\zeta - z} \qquad (E \text{ a Borel set})$$

are uniformly absolutely continuous (see [2, p. 170]). Suppose ε is a positive number. Find a number t such that, in the notation of 2.2,

$$\int_0^t \frac{1}{r} dm(r) < \frac{\varepsilon}{2} . \tag{2.4.1}$$

Let $\delta = \varepsilon t/2$. Suppose $\mu(E) < \delta$.
Then

$$|F_z(E)| \leqq \int_E \frac{\mu(d\zeta)}{|\zeta - z|} .$$

We break E into two parts: E_1, the part on which $|\zeta - z| < t$; and E_2, the part on which $|\zeta - z| \geqq t$. It is obvious that

$$\int_{E_2} \frac{\mu(d\zeta)}{|\zeta - z|} < \frac{\delta}{t} = \frac{\varepsilon}{2} .$$

By breaking E_1 into concentric annuli, and approximating the integral

$$\int_{E_1} \frac{\mu(d\zeta)}{|\zeta - z|} = I_1 \tag{2.4.2}$$

by finite sums, it can be shown that

$$I_1 \leqq \int_0^t \frac{1}{r} dm(r) . \tag{2.4.2.1}$$

This shows that $|F_z(E)| < \varepsilon$ for $\mu(E) < \delta$, independently of z, as was to be shown. We may thus apply Proposition 29.6s of [2, p. 171] to conclude that if $z_n \to z$, then

$$\int_Z \left| \frac{1}{\zeta - z_n} - \frac{1}{\zeta - z} \right| \mu(d\zeta) \to 0 . \tag{2.4.3}$$

Returning to $h(z)$ as defined by 2.3, we observe that

$$|h(z_n) - h(z)| \leqq \|f\| \int_Z \left| \frac{1}{\zeta - z_n} - \frac{1}{\zeta - z} \right| \mu(d\zeta) .$$

In view of 2.4.3, the continuity of h is apparent. The holomorphicity at points bounded away from the support of μ is a similar, actually simpler, proposition. Thus 2.4 is substantially proved.

We now wish to show that with a modified formula we can arrive at a function which is holomorphic wherever f is, *even at points of the support of* μ.

2.5 LEMMA. *Let* μ *be a measure in* B *with bounded support. Let* $f \in H(G/Z)$. *Select any complex number* λ, $|\lambda| \leqq \|f\|$, *and define* $f(z) = \lambda$ *for* $z \notin Z$. *Define*

$$h(z) = \int \frac{f(\zeta) - f(z)}{\zeta - z} \mu(d\zeta) . \tag{2.5.1}$$

Then $h \in H(G/Z)$, *and* $\|h\| \leqq 2I(\mu)\|f\|$. *Moreover,* h *is independent of* λ.

Proof. We write

$$h_1(z) = \int \frac{f(\zeta)}{\zeta - z} \mu(d\zeta) , \quad h_2(z) = f(z) \int \frac{\mu(d\zeta)}{\zeta - z} . \tag{2.5.2}$$

By 2.4, these are continuous on Z. If f is holomorphic at ∞, then h_1, h_2 are holomorphic there. Now let z_0 be a finite point of G. Indeed, assume $z_0 = 0$. Then f is holomorphic at 0. *If* $f(0) = 0$ *then* h_1, h_2 *are both differentiable there.* Indeed

$$\frac{h_1(z) - h_1(0)}{z} = \int \frac{f(\zeta)}{\zeta} \frac{1}{\zeta - z} \mu(d\zeta)$$

and

$$\frac{h_2(z) - h_2(0)}{z} = \frac{f(z)}{z} \int \frac{\mu(d\zeta)}{\zeta - z} .$$

Each of these has a limit as $z \to 0$ because $f(t)/t$ is bounded for all t, and continuously definable near 0. If $f(0) \neq 0$, we replace f by $f - f(0)$. This does not change h. Thus h is differentiable at each $z_0 \in G$.

The main result of this section is as follows:

2.6 THEOREM. *Let* z_0 *be a point of* Z, $z_0 \neq \infty$, *and let* $\mu_1, \mu_2, \cdots$ *be a sequence of measures in* B *such that*

(2.6.1) *the support of* μ_n *lies in the* δ_n*-neighborhood of* z_0 *where*

$$\delta_n \to 0 \tag{2.6.2}$$

and such that for some $M < \infty$,

(2.6.3) $$\delta_n I(\mu_n) < M \|\mu_n\| .$$

Let $f \in H(G/Z)$, and define[1]

(2.6.4) $$h_n(z) = \frac{1}{\|\mu_n\|}\int \frac{f(\zeta) - f(z)}{\zeta - z}\mu_n(d\zeta) .$$

Then $$h_n \in H(G/Z),\ \|h_n\| \leqq 2M\delta_n^{-1}\|f\|$$

and

(2.6.5) $$\|f - f(z_0) - (z - z_0)h_n\| \to 0 .$$

Proof. It will suffice to treat the case $\|\mu_n\| = 1$, and $z_0 = 0$. At first we shall deal with just one μ_n, and therefore omit the suffix. We may also confine attention to the case $f(z_0) = 0$.

We commence our calculations by observing that

$$zh(z) - f(z) = \int\Bigl(z\frac{f(\zeta) - f(z)}{\zeta - z} - f(z)\Bigr)\mu(d\zeta)$$
$$= \int\frac{zf(\zeta) - \zeta f(z)}{\zeta - z}\mu(d\zeta) .$$

Let

$$b(r) = \sup\{|f(z)|\colon |z| \leqq r\} .$$

Now suppose μ is supported by the δ-neighborhood of $z_0 = 0$. Then we may make an estimate

(2.6.6) $$|zh(z) - f(z)| \leqq (|z|b(\delta) + \delta|f(z)|)\int\frac{\mu(d\zeta)}{|\zeta - z|} .$$

For the integral in 2.6.6 there are two possible estimates:

(2.6.7) $$\int\frac{\mu(d\zeta)}{|\zeta - z|} \leqq \begin{cases} \dfrac{1}{|z| - \delta} & \textit{when } |z| > \delta \\ I(\mu) . & \textit{in general.}\end{cases}$$

The latter of these results from 2.4.2.1 for $t \to \infty$.

Now let ε be any positive number. Select a real number $k > 1$ such that

(2.6.8) $$2\|f\| < (k - 1)\varepsilon .$$

Let μ in the preceding discussion be one of the μ_n. Then 2.6.6, 2.6.7 hold with $\delta = \delta_n$. Consider first a point $z \in Z$ such that $|z| \geqq k\delta_n$. We then obtain from 2.6.6, 2.6.7, 2.6.8, that

[1] We extend the definition of f to the whole plane, if necessary, by making it have the value $f(z_0)$ everywhere outside of Z.

$$|zh_n(z) - f(z)| \leqq \frac{|z|}{|z| - \delta_n} b(\delta_n) + \frac{\delta_n}{|z| - \delta_n} ||f||$$

$$\leqq \frac{k}{k-1} b(\delta_n) + \frac{1}{k-1} ||f|| \tag{2.6.9}$$

$$< \frac{\varepsilon}{2} + \frac{k}{k-1} b(\delta_n) \,.$$

Next consider the case $|z| < k\delta_n$. Using 2.6.6, 2.6.7 we obtain

$$|zh_n(z) - f(z)| \leqq [k\delta_n b(\delta_n) + \delta_n b(k\delta_n)]\, I(\mu_n)$$
$$< M(kb(\delta_n) + b(k\delta_n)) \,.$$

The numbers $b(\delta_n)$ and $b(k\delta_n)$ tend to 0 since f is continuous at $z_0 = 0$. Therefore there is an N such that $||zh_n - f|| < \varepsilon$ for $n > N$.

This concludes our proof of 2.6.

2.7 THEOREM. *Let z_0 be a point of $Z(z_0 \neq \infty)$, and let $f \in H(G/Z)$. Then there exist functions $h_1, h_2, \ldots$ in $H(G/Z)$ such that*

$$||f - f(z_0) - (z - z_0)h_n|| \to 0 \,.$$

This emerges from the combination of 2.6 and 2.3.

3. Application to maximal ideals.

3.1 THEOREM. *Let Z be a compact set in the extended complex plane. Let G be an open subset of Z. Then the space of maximal ideals of the Banach algebra $H(G/Z)$ is naturally homeomorphic to Z, provided $H(G/Z)$ does not reduce to the constant functions.*

Proof. If the set G if void, then the proposition reduces to a case of [1, p. 54]. If G is not void, but Z is a proper subset of the extended plane, it is natural to change coordinates so that Z lies in the finite plane. However, the most interesting case, $Z =$ the extended plane, is best treated by having G be a neighborhood of ∞. For economy, if not clarity, we perform a conformal transformation, if necessary, to make G (whenever it is not empty) a neighborhood of ∞.

Let F be a multiplicative linear functional of $H(G/Z)$. If $G(f) = f(\infty)$ for all f in $H(G/Z)$, then F (or its kernel) corresponds to ∞. Having disposed of that unique multiplicative functional, let F be some other one. Then $F(f_1) = 1$ for some $f_1 \in H(G/Z)$ such that $f_1(\infty) = 0$. Then $zf_1 \in H(G/Z)$. Let $z_0 = F(zf_1)$. Then $F((z - z_0)h) = 0$ for each $h \in H(G/Z)$, such that $(z - z_0)h \in H(G/Z)$. Because

$$F((z - z_0)h) = F((z - z_0)h)F(f_1) = F(h)F((z - z_0)f_1) = 0.$$

Now suppose $f \in H(G/Z)$. By 2.7, there exist h_n, r_n in $H(G/Z)$ such that

$$f - f(z_0) - (z - z_0)h_n = r_n \tag{3.2}$$

where $\|r_n\| \to 0$. Hence $F(r_n) \to 0$. Also, $F((z - z_0)h_n) = 0$. Hence $F(f - f(z_0)) = 0$, $F(f) = f(z_0)$. Then we say that F corresponds to z_0.

We have thus shown that to every maximal ideal, or multiplicative linear functional F, there corresponds a point z_0. There might be several such points corresponding to a given F. The situation is completely illuminated by a device of Wermer's [4, p. 269] which shows that either $H(G/Z)$ consists only of the constant functions, or some triad of functions separate all pairs of points on Z. This completes the proof of 3.1.

We can now acknowledge the relation of our argument to that of Hoffman and Singer. They construct an h_n, define r_n as in 3.2, and show $\|r_n\| \to 0$; and so forth. Their choice of $h_n(z)$,

$$f(z)\frac{1}{m_n}\int_{E_n}\frac{d\zeta\, d\eta}{z - \zeta} \qquad (\zeta = \xi + i\eta) \tag{3.3}$$

where E_n is the intersection of a δ_n-disc about z_0, with $S - G$, and m_n is the measure of m_n, is effective only when $m_n > 0$ as $\delta_n \to 0$. Hence they assume that $S - G$ has positive upper density at z_0.

4. Remarks on the dimension of H(G/Z). We return to the question, when does $H(G/Z)$ contain non-constant functions? A sufficient condition is that $S - G$ carry a measure of type B, for then 2.4 provides such functions. The formula 2.4.3 is used in [3, 4] for this very purpose, but the measures there employed are absolutely continuous. For this reason it is desirable to point out that there are measures in B that have zero-dimensional support of plane measure zero. An example can be obtained from a well-known function, which increases only at points of the Cantor set. Calling this function f, as in [2, p. 49], we form the measure on the line; and then we form the product measure μ of this measure with itself. It is not hard to see that $I(\mu) \leqq 16\sqrt{2}(1 - \lambda)^{-1}$ where $\lambda = \log 4/\log 3$.

Questions analogous to the above are discussed in [6, 7].

REFERENCES

1. L. H. Loomis, *Abstract hnrmonic analysis*, D. Van Nostrand, New York, 1953.
2. E. J. McShane, *Integration*, Princeton University Press, Princeton, 1947.
3. Walter Rudin, *Algebras of analytic functions*, Summary of Lectures and Seminars, *Summer Institute on Set-Theoretic Topology*, (Amer. Math. Soc.) University of Wisconsin, Mathematics Department, Madison **6** (1955), 137–139.
4. John Wermer, *Polynomial approximation on an arc in C^3*, Ann. of Math. Stat., **62** (1955), 269–270.

5. S. N. Mergelyan, *On the representation of functions by series of polynomials on closed sets*, Amer. Math. Soc. Translation 85, Providence (R.I.), translated from Doklady Akademii Nauk SSSR (N.S.) **78** (1951), 405-408.
6. L. Ahlfors and A. Beurling, *Function-theoretic null sets and conformal invariants*, Acta Math., **83**, (1950), 101–129.
7. Otto Frostman, *Potentiel d'équilibre et capacité* . . . (thesis) Lund (1935).

UNIVERSITY OF CALIFORNIA, LOS ANGELES

AN OPERATOR IDENTITY

GLEN BAXTER

1. Introduction. Recently, some combinatorial results by Andersen [1, 2], Spitzer [5], and others have been applied quite successfully to problems in probability theory. Many of these applications have given rise to results which are entirely analytical in nature. For example, Spitzer used a combinatorial theorem to find the distribution function for the maximum of the partial sums $S_1, S_2, \cdots, S_n$ for a sequence $\{X_k\}$ of independent, identically distributed random variables. His final result is a functional identity,

$$\sum_{n=0}^{\infty} \varphi_n(t) s^n = \exp\left\{\sum_{k=1}^{\infty} \frac{s^k}{k} \psi_k(t)\right\}, \tag{1.1}$$

where $\varphi_n(t)$ is the characteristic function of $\max(0, S_1, \cdots, S_n)$ and where $\psi_k(t)$ is the characteristic function of $\max(0, S_k)$. One of our purposes in this paper is to generalize (1.1) to an identity involving operators. Our proofs involve more or less analytical methods and thus show that the combinatorial methods hitherto employed can be avoided. We also obtain certain results concerning $\max(X_0, X_1, \cdots, X_n)$ when $\{X_k, k \geq 0\}$ forms a stationary Markov process.

To illustrate the results we consider a simple example. Let N be an $n \times n$ matrix and let N^+ be the matrix formed from N by replacing with zeros all elements of N which are either on or below the diagonal. Let $N^- = N - N^+$, and suppose that N^+ and N^- commute. Now consider the matrix equation

$$PQ = e^N = I + N + N^2/2! + \cdots \tag{1.2}$$

where P-I (I is the identity matrix) has non-zero terms only above the diagonal and where $Q - I$ has non-zero terms only on or below the diagonal. The properties of N^+ and N^- imply that

$$\begin{aligned} P &= e^{N^+} = I + N^+ + (N^+)^2/2! + \cdots, \\ Q &= e^{N^-} = I + N^- + (N^-)^2/2! + \cdots \end{aligned} \tag{1.3}$$

satisfy (1.2) and have the proper form for P and Q. In particular, $\exp(N^+)$ has the proper form for P by virtue of the fact that the product of two matrices with non-zero elements only above the diagonal is a

This paper was originally accepted by the Trans. Amer. Math. Soc. Presented to the American Mathematical Society August 28, 1957; received by the editors of the Trans. Amer. Math. Soc., July 29, 1957. This work was supported by the United States Air Force under Contract No. AF18(603)-30 monitored by the Office of Scientific Research.

matrix of the same type. A similar statement holds for $\exp(N^-)$. It is not hard to see that P and Q are uniquely determined by (1.2). Thus (1.3) is the unique solution of (1.2).

Suppose further that in some neighborhood of $s = 0$, $N = N_1 s + N_2 s^2 + \cdots$, where convergence of the infinite series of $(n \times n)$ matrices is equivalent to convergence of the series of ijth elements for all fixed i and j. Relations (1.3) may be rewritten as power series in s

$$P = \sum_{n=0}^{\infty} P_n s^n \,, \qquad Q = \sum_{n=0}^{\infty} Q_n s^n \tag{1.4}$$

which converge in some neighborhood of $s = 0$. It follows from the form of P and Q that $P_1, P_2, \cdots$ have non-zero elements only above the diagonal while $Q_1, Q_2, \cdots$ have non-zero elements only on or below the diagonal. Certain problems will lead directly to an equation of the form (1.2) where P and Q have the form (1.4). For example, in one case we will have

$$PQ = (I - sM)^{-1} = \exp\left\{\sum_{k=1}^{\infty} \frac{M^k}{k} s^k \right\} . \tag{1.5}$$

Under the appropriate commutativity conditions it will follow that

$$P = \exp\left\{\sum_{k=1}^{\infty} \frac{(M^k)^+}{k} s^k\right\} , \qquad Q = \exp\left\{\sum_{k=1}^{\infty} \frac{(M^k)^-}{k} s^k\right\} . \tag{1.6}$$

We see later that (1.6) is the operator analogue of Spitzer's identity (1.1) whenever the operator M has a special form.

Equation (1.5) is of particular importance in finding the distribution of $\max(X_0, X_1, \cdots, X_n)$ when $\{X_k, k \geq 0\}$ is a Markov process with a stationary transition probability matrix M. In this case the matrix M in (1.5) is identified (see § 4) with the stationary transition probability matrix M. Unfortunately, in the general Markov chain, the commutativity conditions which give (1.6) as the solution of (1.5) are not satisfied. Some information can be obtained directly from (1.5).

In the next section we give general definitions and a few preliminary results. The main theorems are proved in § 3 and illustrated in § 5. A probabilistic interpretation of the theorems is contained in § 4.

2. Definitions and preliminaries. Let L_0 be the space of bounded Baire functions (real-valued and Borel measurable) $f(x)$ on the infinite interval $-\infty < x < \infty$. We will deal with bounded linear operators M defined over L_0 which have the form

$$Mf \equiv \int_{-\infty}^{\infty} f(y) m(x\,; dy) \tag{2.1}$$

where $m(x\,;A)$ is a function of a real number x and a linear Borel measurable set A such that

(2.2) (i) for each fixed set A, $m(x\,;A)$ is a Baire function of x,
(ii) for each fixed x, $m(x\,;A)$ is a signed measure in A on the linear Borel sets.

The norm of the operator M is defined in the usual way in terms of the norm $\|f\| = \max|f(x)|$ in the Banach space L_0. Let $\mu(x\,;A)$ and $v(x\,;A)$ be, respectivery, the upper variation and the lower variation of the signed measure $m(x\,;A)$ (see [4, page 122]) The boundedness of M in (2.1) implies that

$$\int_{-\infty}^{\infty}[\mu(x\,;dy)+\nu(x\,;dy)] \leq \max_{-\infty<x<\infty}\int_{-\infty}^{\infty}[\mu(x\,;dy)+\nu(x\,;dy)] = \|M\| < \infty\,. \tag{2.3}$$

We call $m(x\,;A)$ the kernel of the operator M. The notation which will be used for integration with respect to a given measure is indicated in (2.1). From now on when we call M a bounded linear operater of the form (2.1), we imply that (2.2) is also satisfied. As a matter of fact, with proper understanding of the notation, (2.2) follows directly from (2.1). If M_1 and M_2 are bounded linear operators of the form (2.1) with kernels $m_1(x\,;A)$ and $m_2(x\,;A)$, respectively, then M_1M_2 is also of the form (2.1) with kernel

$$m(x\,;A) = \int_{-\infty}^{\infty} m_2(y\,;A)m_1(x\,;dy)\,. \tag{2.4}$$

We now let $[x]$ be the greatest integer less than or equal to x.

DEFINITION 2.1. Set $B_n(x) = \{y : y > [2^n x + 1]/2^n\}$. For any bounded linear operator M of the form (2.1) with kernel $m(x\,;A)$, define

$$m^+(x\,;A) \equiv \lim_{n\to\infty} m(x\,;B_n(x)A)\,, \tag{2.5}$$

and let M^+ be the operator of form (2.1) with kernel $m^+(x\,;A)$. Finally, set, $M^- = M - M^+$.

Almost directly from the definition of M^+ follow certain useful facts which we list below. The bounded, linear operators M, M_1, M_2, etc. are all of the form (2.1); I denotes the identity operator, which is also of the form (2.1); and s, α, and β, are real numbers:

(i) $I^- = I$,
(ii) $(M^+)^+ = M^+$, (iii) $(M^-)^- = M^-$,
(iv) $(M_1^+M_2^+)^+ = M_1^+M_2^+$, (v) $(M_1^-M_2^-)^- = M_1^-M_2^-$,

(vi) $\|M^+\| \leq \|M\|$, (vii) $\|M^-\| \leq \|M\|$,

(viii) $(\alpha M_1 + \beta M_2)^+ = \alpha M_1^+ + \beta M_2^+$,

(2.6) (ix) if $M_0 + M_2 + \cdots$ is a strongly convergent series of bounded, linear operators of the form (2.1), i.e. if $\|M_n + \cdots + M_m\| \to 0$ as $n, m \to \infty$, then $T = M_0 + M_1 + M_2 + \cdots$ is of the form (2.1), and $M_0^+ + M_1^+ + M_2^+ + \cdots$ and $M_0^- + M_1^- + M_2^- + \cdots$ are both convergent in the strong sense. Moreover, $T^+ = M_0^+ + M_1^+ + M_2^+ + \cdots$ and $T^- = M_0^- + M_1^- + M_2^- + \cdots$.

We prove only (ix) of (2.6). Let $T_n = M_0 + \cdots + M_n$, let $t_n(x; A)$ be the kernel of T_n, and let χ_A be the characteristic function of a measurable set A. If $T = \lim T_n$, we note that $\|T\|$ is finite. Now

$$|t_n(x;\ A) - t_m(x;\ A)| = |(T_n - T_m)\chi_A| \leq \|T_n - T_m\|,$$

so that $\lim t_n(x;\ A) = t(x; A)$ exists uniformly in A. If $A = \Sigma A_k$ where the A_k are disjoint, then by the Moore double-limit theorem

$$\sum_{k=1}^{\infty} t(x; A_k) = \lim_{N\to\infty} \lim_{n\to\infty} \sum_{k=1}^{N} t_n(x; A_k) = \lim_{n\to\infty} t_n(x; A) = t(x; A). \tag{2.7}$$

This shows that $t(x; A)$ is a signed measure. Since $T\chi_A = t(x; A)$, a simple argument shows that $t(x; A)$ is the kernel of T. Finally, since $\|T^+ - T_n^+\| \leq \|T - T_n\|$, it follows that $T^+ = \lim T_n^+$. In terms of M_n this means $T^+ = M_0^+ + M_1^+ + M_2^+ + \cdots$. A similar argument gives $T^- = M_0^- + M_1^- + M_2^- + \cdots$.

It is interesting to note that the proofs of the main theorems will depend only on the facts listed in (2.6). Before proceeding to the next section we mention two special subclasses of operators which have the form (2.1).

Case 1. Let $M = (m_{ij})$ be a matrix for which uniformly in i

$$\sum_{(j)} |m_{ij}| < C \tag{2.8}$$

for some constant C. For any Borel measurable set A and any real number x define

$$m(x; A) = \begin{cases} \sum_{j\in A} m_{ij} & x = i \text{ (an integer)} \\ 0 & x \neq [x]. \end{cases} \tag{2.9}$$

Condition (2.8) insures the existence of a bounded linear operator of form (2.1) with the kernel $m(x; A)$ of (2.9). Certainly the operator given by (2.1) in this case and the original matrix M can be identified. In fact, L_0 could be replaced here by the class of bounded, doubly infinite sequences $\{a_k\}$, that is $a_k = f(k)\ (-\infty < k < \infty)$ where $f(x) \in L_0$. It will be convenient whenever possible to think of the matrix M rather than the operator M. Note that the matrix M^+ is formed from the matrix M by replacing with zeros all elements of M either on or below the diagonal. Moreover, the matrix M^+ satisfies (2.6).

Case 2. Let $m(x, y)$ be Borel measurable and integrable over the plane and such that for some constant C

$$\int_{-\infty}^{\infty} | m(x, y) | \, dy < C \tag{2.10}$$

uniformly in x. For any Borel measurable set A and any real number x, define

$$m(x\,;\,A) \equiv \int_A m(x, y)dy\ . \tag{2.11}$$

Then, (2.1) gives a bounded, linear operator M which has the form

$$M\,\cdot\, = \int_{-\infty}^{\infty} \cdot\; m(x, y)dy\ , \tag{2.12}$$

and M^+ becomes simply

$$M^+\,\cdot\, = \int_{x}^{\infty} \cdot\; m(x, y)dy \tag{2.13}$$

with a similar formula for M^-.

3. The theorems. When we say a sequence of operators $\{M_n\}$ converges to an operator M, we mean it converges in the strong sense, that is $\| M_n - M \| \to 0$ as n becomes infinite.

LEMMA 3.1. *Let* $\{K_k\}$, $\{P_k\}$, *and* $\{Q_k\}$, $k = 1, 2, 3, \cdots$, *be sequences of bounded, linear operators of the form* (2.1) *for which* $P_k^+ = P_k$ *and* $Q_k^- = Q_k$. *For any* $| s | < s_0$, *let*

$$\begin{aligned} P &= I + P_1 s + P_2 s^2 + \cdots\ , \\ Q &= I + Q_1 s + Q_2 s^2 + \cdots\ , \\ K &= I + K_1 s + K_2 s^2 + \cdots \end{aligned} \tag{3.1}$$

converge. If $PQ = K$ *for all* $| s | < s_0$, *then* $\{P_k\}$ *and* $\{Q_k\}$ *are uniquely determined by* $\{K_k\}$.

Proof. Equating coefficients of like powers of s on the two sides of the equation $PQ = K$ we obtain

$$\sum_{k=0}^{n} P_k Q_{n-k} = K_n\ . \tag{3.2}$$

If $P_1, P_2, \cdots, P_{n-1}$ and $Q_1, Q_2, \cdots, Q_{n-1}$ have been uniquely determined by $K_1, K_2, \cdots, K_{n-1}$, then we may write (3.2) as

$$P_n + Q_n = J_n \tag{3.3}$$

where J_n is determined uniquely by $K_1, K_2, \cdots, K_n$. Since $P_n^- = Q_n^+ = 0$, we have $P_n = J_n^+$ and $Q_n = J_n^-$ and the proof follows by induction.

The next theorems give results in the direction of solving equations which involve the operation " + ". Later we give a probabilistic interpretation of these equations. As we will see, in certain cases the equations may be solved completely in terms of the known operator M.

THEOREM 3.1. *Let M be a bounded, linear operator of the form* (2.1). *Define the sequences $\{P_k\}$, $\{Q_k\}$, $\{R_k\}$, and $\{T_k\}$ by*

$$\begin{aligned} P_0 &= Q_0 = I\ , & R_0 &= T_0 = 0\ , \\ P_{n+1} &= (MP_n)^+\ , & Q_{n+1} &= (Q_nM)^-\ , \\ T_{n+1} &= (MP_n)^-\ , & R_{n+1} &= (Q_nM)^+\ , \end{aligned} \tag{3.4}$$

and let the generating functions of these sequences be

$$\begin{aligned} P &= \sum_{n=0}^{\infty} P_n s^n\ , & Q &= \sum_{n=0}^{\infty} Q_n s^n\ , \\ R &= \sum_{n=0}^{\infty} R_n s^n\ , & T &= \sum_{n=0}^{\infty} T_n s^n\ , \end{aligned} \tag{3.5}$$

Then, the series' in (3.5) *all converge for $|s| < 1/\|M\|$, and, moreover, they are the unique bounded, linear operators of the form* (2.1) *which satisfy.*

$$\begin{aligned} P &= I + s(MP)^+\ , & T &= s(MP)^-\ , \\ Q &= I + s(QM)^-\ , & R &= s(QM)^+\ . \end{aligned} \tag{3.6}$$

Proof. Let P be a bounded, linear operator of the form (2.1) which satisfies the first equation of (3.6). By iteration we may write $P = I + P_1 s + P_2 s^2 + \cdots + P_n s^n + L_n$, where $L_0 = s(MP)^+$ and $L_n = s(ML_{n-1})^+$ and where $P_1, P_2, \cdots, P_n$ are determined in (3.4). Property (vi) of (2.6) implies that $\|L_n\| \leq |s|^n \|M\|^n \|P\|$ which approaches zero as n becomes infinite for all $|s| < 1/\|M\|$. Thus, the solution (if it exists) of the first equation of (3.6) is unique. Let $\{P_k\}$ satisfy the conditions of (3.4). By property (vi) of (2.6), it follows that $\|P_n\| \leq \|M\|^n$. For $|s| < 1/\|M\|$, the power series in (3.5) for P converges and by property (ix) of (2.6)

$$P - I = \sum_{n=0}^{\infty} P_{n+1}s^{n+1} = \sum_{n=0}^{\infty} (MP_n)^+ s^{n+1} = \left(\sum_{n=0}^{\infty} MP_n s^{n+1}\right)^+ = s(MP)^+\ . \tag{3.7}$$

The proofs of the other parts of the theorem follow similarly.

THEOREM 3.2. *Let* $|s| < 1/\|M\|$ *and let* P *and* Q *be the bounded, linear operators of the form* (2.1) *which satisfy the equations of* (3.6). *Then,*

$$(3.8)\qquad \begin{aligned} &PQ = (I - sM)^{-1} \\ &sP' = P(QP - I)^{+}\,, \qquad sQ' = (QP - I)^{-}Q\,, \end{aligned}$$

where $'$ *indicates derivative with respect to* s.

Proof. From (3.6) we find that $\|Q\| \leq 1/(1 - |s|\,\|M\|)$ and

$$\|R\| \leq |s|\,\|M\|\,\|Q\| \leq |s|\,\|M\|/(1 - |s|\,\|M\|)\,.$$

Thus, for $|s| < (1 - |s|\,\|M\|)/\|M\|$, the operator $(I - R)^{-1}$ is a bounded linear operator of the form (2.1) and has a convergent power series expansion in s. But $Q = I - R + sQM$, or equivalently, $(I - R)^{-1}Q = (I - sM)^{-1}$. Similarly we show that $(I - T)^{-1}$ is a bounded linear operator of the form (2.1) which has a convergent power series expansion in s for $|s| < (1 - |s|\,\|M\|)/\|M\|$, and that $P(I - T)^{-1} = (I - sM)^{-1}$. Applying Lemma 3.1 in the common interval of convergence of $P, Q, (I - T)^{-1}$ and $(I - R)^{-1}$, we deduce that

$$(3.9)\qquad P = (I - R)^{-1}\,, \qquad Q = (I - T)^{-1}\,.$$

and hence that $PQ = (I - sM)^{-1}$. Since P, Q, and $(I - sM)^{-1}$ all converge for $|s| < 1/\|M\|$, we have finally $PQ = (I - sM)^{-1}$ for all $|s| < 1/\|M\|$. To show the second half of (3.8), we consider $(PQ)' = P'Q + PQ' = (I - sM)^{-2}M$. It follows that

$$(3.10)\qquad (PQ)^2 - s(PQ)' = (I - sM)^{-2}(I - sM) = PQ\,.$$

Multiplying on the left of (3.10) by P^{-1} and on the right by Q^{-1} (take $|s| < (1 - |s|\,\|M\|)/\|M\|$) we obtain

$$(3.11)\qquad QP - s(P^{-1}P' + Q'Q^{-1}) = I\,.$$

By properties (iv), (v), and (ix) of (2.6), it is not hard to see that $(P^{-1}P')^{+} = P^{-1}P'$ and $(Q'Q^{-1})^{-} = Q'Q^{-1}$. From (3.11) we find $sP' = P(QP - I)^{+}$ and $sQ' = (QP - I)^{-}Q$. These latter equations can certainly be extended to hold for all $|s| < 1/\|M\|$, and the theorem is proved.

THEOREM 3.3. *Let* $\{a_k\}$ *be a sequence of real numbers such that* $a_1s + a_2s^2 + a_3s^3 + \cdots$ *has a positive radius of convergence. Let* M *be a bounded, linear operator of the form* (2.1) *such that* $(M^k)^{+}M = M(M^k)^{+}$ *for all* $k = 1, 2, 3, \cdots$. *Then for* $|s|$ *such that*

$$(3.12)\qquad \sum_{k=1}^{\infty} |a_k|\,\|M\|^k\,|s|^k < 1\,,$$

there is a unique pair of bounded linear operators P and Q of the form (2.1) *which satisfy*

$$P = I + \left[\sum_{k=1}^{\infty} (a_k M^k s^k) P\right]^+ , \qquad Q = I + \left[Q \sum_{k=1}^{\infty} (a_k M^k s^k)\right]^- . \tag{3.13}$$

Moreover, the solution of (3.13) *is*

$$P = \exp\left\{\left[- \log (I - \sum_{k=1}^{\infty} a_k M^k s^k)\right]^+\right\} , \qquad Q = \exp\left\{\left[- \log (I - \sum_{k=1}^{\infty} a_k M^k s^k)\right]^-\right\} . \tag{3.14}$$

Before proving Theorem 3.3 we mention a result of particular interest which occurs when both Theorems 3.1 and 3.3 apply, i.e. when $a_1 = 1$ and $a_2 = a_3 = \cdots = 0$.

COROLLARY 3.1. *Let M be a bounded linear operator of the form* (2.1) *such that $(M^k)^+M = M(M^k)^+$ for all $k = 1, 2, 3, \cdots$, and let the sequences $\{P_k\}$ and $\{Q_k\}$ be defined as in* (3.4). *Then, for all $|s| < 1/\|M\|$, the P and Q of* (3.5) *have the form*

$$P = \exp\left\{\sum_{k=1}^{\infty} \frac{(M^k)^+}{k} s^k\right\} , \qquad Q = \exp\left\{\sum_{k=1}^{\infty} \frac{(M^k)^-}{k} s^k\right\} . \tag{3.15}$$

Proof of Theorem 3.3. Let $|s|$ satisfy the condition of (3.12), and let

$$L = \sum_{k=1}^{\infty} a_k M^k s^k , \qquad N = \log(I - \sum_{k=1}^{\infty} a_k M^k s^k) = \sum_{k=1}^{\infty} L^k/k . \tag{3.16}$$

Both L and N are bounded linear operators of the form (2.1). The commutativity of $(M^k)^+$ and M together with property (ix) of (2.6) implies that $L^+L = LL^+$. Again by property (ix) of (2.6) and the second relation of (3.16), we deduce that $N^+N = NN^+$. In terms of N the first equation in (3.13) may be written in the form

$$P = I + [(I - e^N)P]^+. \tag{3.17}$$

Using that $(\exp(- N^+))^+ = \exp(- N^+) - I$ and that $(\exp(N^-))^+ = 0$, it is easy to show by substitution that $P = \exp(- N^+)$ is a solution of (3.17). To show that this solution is unique we apply Theorem 3.1, where the operator "M" of Theorem 3.1 is now

$$\sum_{k=1}^{\infty} a_k M^k s^k \tag{3.18}$$

and the number "s" of Theorem 3.1 is now 1. In a similar manner we can show that the Q of (3.14) is the unique solution of the second equation in (3.13). This finishes the proof.

Before proceeding into the next section, we point out some implications of the theorems above. In Theorem 3.3, the operators P, Q, M, M^+, and M^- all commute. Thus, the order of the factors Q and M^k or of P and M^k in (3.13) is unimportant. In the s interval determined by (3.12), there is a power series expansion in s for the solutions of (3.13). The coefficients in this power series satisfy

$$\begin{aligned} &P_0 = Q_0 = I\,, \\ &P_{n+1} = (a_1 M P_n + a_2 M^2 P_{n-1} + \cdots + a_{n+1} M^{n+1})^+\,, \\ &Q_{n+1} = (a_1 Q_n M + a_2 Q_{n-1} M^2 + \cdots + a_{n+1} M^{n+1})^-\,. \end{aligned} \tag{3.19}$$

If the M in Theorem 3.1 is a matrix of finite order, the P and Q of (3.5) can be conveniently evaluated in terms of subdeterminants of the matrix $I - sM$ (See example 3, § 5).

4. Probabilistic interpretation. In this section we give a probabilistic interpretation of the sequences $\{P_k\}$, $\{Q_k\}$, $\{R_k\}$, and $\{T_k\}$ of Theorem 3.1. Let $m(x\,; A)$ be a function of a real number x and a linear Borel measurable set A such that

(4.1)
(i) for each fixed set A, $m(x\,; A)$ is a Baire function of x,
(ii) for each fixed x, $m(x\,; A)$ is a probability measure in A on the linear Borel measurable sets.

Let $\{X_k, k \geq 0\}$ be a stationary Markov process for which $m(x\,; A) = P\{X_{k+1} \in A \mid X_k = x\}$ is defined and satisfies the conditions of (4.1) (see [3, pp. 18, 26–27]). We deal here only with processes of this type. By (2.1) and (2.3) each Markov process under consideration has associated with it a bounded linear operator M, with $\| M \| = 1$. We call this the transition probability operator of the process.

Two subcases of special interest may be mentioned. The first one is that of a discrete Markov *chain* (countable state space). In this case the transition probabilities form a matrix $M = (m_{ij})$. The connection between the matrix M and the function $m(x\,; A)$ has already been discussed in § 2, case 1. The second type process of interest is the one for which the joint distributions have densities. In this latter case, there exists a transition probability density function $m(x, y)$, and the connection with $m(x\,; A)$ is given in § 2, case 2.

For convenience in stating the next theorem we introduce a random variable L_n.

(4.2) L_n: the index $k(=0, 1, 2, \cdots)$ for which $\max(X_0, X_1, \cdots, X_n) = X_k$ and $\max(X_0, X_1, \cdots, X_{k-1}) < X_k$.

Note in particular the meaning of the statements $L_n = n$ and $L_n = 0$. In Theorem 4.1 and thereafter we will have occasion to refer to the kernel associated with a given operator of the form (2.1). If the operator is denoted by some capital letter, the kernel will be denoted by the corresponding small letter.

THEOREM 4.1. *Let $\{X_k, k \geq 0\}$ be a stationary Markov process with transition probability operator M, and let $\{P_k\}$, $\{Q_k\}$, $\{R_k\}$ and $\{T_k\}$ be defined as in* (3.4). *Then, if the right hand members of* (4.3) *are defined and satisfy* (2.2), *we have*

$$
\begin{aligned}
p_n(x\,; A) &= P\{L_n = n, X_n \in A \mid X_0 = x\}\,, \\
q_n(x\,; A) &= P\{L_n = 0, X_n \in A \mid X_0 = x\}\,, \\
r_n(x\,; A) &= P\{L_n = n, L_{n-1} = 0, X_n \in A \mid X_0 = x\}\,, \\
t_n(x\,; A) &= P\{L_n = 0, \max(X_1, \cdots, X_{n-1}) < X_n, X_n \in A \mid X_0 = x\}\,.
\end{aligned} \tag{4.3}
$$

Proof. We prove only the first one of the relations in (4.3). Our proof is by induction. Since $P_0 = I$, it follows that

$$
p_0(x\,; A) = P\{X_0 \in A \mid X_0 = x\} = \begin{cases} 1 & x \in A \\ 0 & x \notin A\,. \end{cases} \tag{4.4}
$$

Now assume the first relation of (4.3) is true for the case n and set $B_N(x) = \{y: y > [2^N x + 1]/2^N\}$ for $N = 1, 2, 3, \cdots$. Then,

$$
\begin{aligned}
&P\{L_{n+1} = n+1, X_{n+1} \in B_N(x)A \mid X_0 = x\} \\
&\quad = \int_{-\infty}^{\infty} P\{\max(X_1, \cdots, X_n) < X_{n+1}, X_{n+1} \in B_N(x)A \mid X_1 = z\} \\
&\qquad\qquad \cdot P\{X_1 \in dz \mid X_0 = x\} \\
&\quad = \int_{-\infty}^{\infty} \{L_n = n, X_n \in B_N(x)A \mid X_0 = z\} P\{X_1 \in dz \mid X_0 = x\} \\
&\quad = \int_{-\infty}^{\infty} p_n(z\,; B_N(x)A) m(x\,; dz)\,.
\end{aligned} \tag{4.5}
$$

From (2.4) we see that the last term of (4.5) is the kernel of MP_n evaluated at x and $B_N(x)A$. Set $A_x = A \cap (x, \infty)$, and note that for any $n > 0$,

$$
P\{L_n = n, X \in A \mid X_0 = x\} = P\{L_n = n, X_n \in A_x \mid X_0 = x\}\,.
$$

Thus, by Definition 2.1 and (4.5)

$$
\begin{aligned}
p_{n+1}(x\,;A) &= \lim_{N\to\infty}\int_{-\infty}^{\infty} p_n(z\,; B_N(x)A,\, m(x\,; dz)) \\
&= \lim_{N\to\infty} P\{L_{n+1} = n+1,\, X_{n+1} \in B_N(x)A \mid X_0 = x\} \\
&= P\{L_{n+1} = n+1,\, X_{n+1} \in A_x \mid X_0 = x\} \\
&= P\{L_{n+1} = n+1,\, X_{n+1} \in A \mid X_0 = x\}\ ,
\end{aligned} \tag{4.6}
$$

and the proof follows by induction.

Combining the first and second of the relations in (4.3) we get certain additional information about $\max(X_0, \cdots, X_n)$. In fact, we can evaluate the generating function

$$\sum_{n=0}^{\infty} P\{\max(X_0, \cdots, X_n) \in A \mid X_0 = x\} s^n \tag{4.6}$$

in terms of the kernels of P and Q. Let $S = (-\infty, \infty)$. Then, by Theorem 4.1

$$
\begin{aligned}
&P\{\max(X_0, \cdots, X_n) \in A \mid X_0 = x\} \\
&\quad = \sum_{k=0}^{n} P\{L_n = k,\, \max(X_0, \cdots, X_n) \in A \mid X_0 = x\} \\
&\quad = \sum_{k=0}^{n} \int_A P\{L_{n-k} = 0 \mid X_0 = y\} P\{L_k = k,\, X_k \in dy \mid X_0 = x\} \\
&\quad = \sum_{k=0}^{n} \int_A q_{n-k}(y\,; S) p_k(x\,; dy)\ .
\end{aligned} \tag{4.7}
$$

Multiplying through (4.7) by s^n and summing over $s = 0, 1, 2, \cdots$ we obtain

$$\int_A q(y\,; S) p(x\,; dy) = \sum_{n=0}^{\infty} P\{\max(X_0, \cdots, X_n) \in A \mid X_0 = x\} s^n\ . \tag{4.8}$$

Relation (4.8) takes on a particularly simple form if $q(y\,; S)$ is independent of y (See example 2, § 5). In fact, in this special case we have the following Corollary to Theorem 4.1:

COROLLARY 4.1. *Let $\{X_k, k \geq 0\}$ be a stationary Markov process with transition probability operator M and let P and Q be defined as in Theorem* 3.1. *Furthermore, let $q(x\,; A)$ be the kernel of Q, and let Φ be the bounded, linear operator of the form* (2.1) *determined by*

$$\varphi(x\,; A) = \sum_{n=0}^{\infty} P\{\max(X_0, \cdots, X_n) \in A \mid X_0 = x\} s^n\ . \tag{4.9}$$

Then, if $q(x\,; S) = q$ is independent of x,

$$\Phi = qP\ . \tag{4.10}$$

Relation (4.10) is an operator analogue of Spitzer's identity (1.1).

5. Examples. We now give applications of the theorems to some particular examples.

EXAMPLE 1. Let the operator of form (2.1) be (See case 1, § 2)

$$M = \begin{bmatrix} a & 0 & b \\ (a-c)d/b & c & d \\ 0 & 0 & a \end{bmatrix}, \tag{5.1}$$

so that for $k = 1, 2, 3, \cdots$

$$M^k = \begin{bmatrix} a & 0 & ka^{k-1}b \\ (a^k - c^k)d/b & c^k & ka^{k-1}d \\ 0 & 0 & a^k \end{bmatrix}. \tag{5.2}$$

It is not hard to see that $(M^k)^+M = M(M^k)^+$ in this case so Corollary 3.1 applies here. The solution of $P = I + s(MP)^+$ for $|s| < 1/\|M\| < 1/|a|$ is

$$\begin{aligned} P &= \exp\left\{\sum_{k=1}^{\infty} \frac{s^k}{k}(M^k)^+\right\} = \exp\left\{\sum_{k=1}^{\infty} s^k \begin{bmatrix} 0 & 0 & a^{k-1}b \\ 0 & 0 & a^{k-1}d \\ 0 & 0 & 0 \end{bmatrix}\right\} \\ &= \exp\left\{\begin{bmatrix} 0 & 0 & bs/(1-as) \\ 0 & 0 & ds/(1-as) \\ 0 & 0 & 0 \end{bmatrix}\right\} = \begin{bmatrix} 1 & 0 & bs/(1-as) \\ 0 & 1 & ds/(1-as) \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned} \tag{5.3}$$

In a similar manner it follows that the solution of $Q = I + s(QM)^-$ for $|s| < 1/\|M\| < 1/\min(|a|, |c|)$ is

$$Q = \begin{bmatrix} 1/(1-as) & 0 & 0 \\ (a-c)ds/b(1-as)(1-cs) & 1/(1-cs) & 0 \\ 0 & 0 & 1/(1-as) \end{bmatrix}. \tag{5.4}$$

These solutions are easily checked by substitution.

EXAMPLE 2. Let $\{X_k\}(k = 1, 2, 3, \cdots)$ be a sequence of independent, identically distributed random variables with a common density function $f(x)$, and let $S_n = X_1 + \cdots + X_n$. If T_0 is any random variable independent of $\{X_k\}$, and if we set $T_n = S_n + T_0$ $(n = 1, 2, 3, \cdots)$, then $\{T_n, n \geq 0\}$ is a stationary Markov process with transition probability

$$m(x\,; A) \equiv P\{T_{k+1} \in A \mid T_k = x\} = \int_A f(y-x)dy\,. \tag{5.5}$$

The conditions (4.1) are satisfied by $m(x\,; A)$ (as well as by the right hand members of (4.3)) in this case so we so may talk about the transition probability operator M associated with $\{T_n, n \geq 0\}$. This operator has the form

$$M \cdot = \int_{-\infty}^{\infty} \cdot f(y-x)dy . \tag{5.6}$$

Using (2.4) and (5.6) it is not hard to deduce that M^k also has a kernel with a density. In fact,

$$M^k \cdot = \int_{-\infty}^{\infty} \cdot f_k(y-x)dy , \tag{5.7}$$

where $f_k(x)$ is the k-fold convolution of $f(x)$ with itself.

By (5.6), (5.7), and (2.4) we see that the kernel of $(M^k)^+M$ has a density of the form

$$\int_{-\infty}^{\infty} f(y-w)f_k^+(w-x)dw = \int_x^{\infty} f(y-w)f_k(w-x)dw . \tag{5.8}$$

We now make the change of variable $z = y + x - w$ in the second integral of (5.8) to get

$$\int_y^{\infty} f_k(y-z)f(z-x)dz = \int_{-\infty}^{\infty} f_k^+(y-z)f(z-x)dz . \tag{5.9}$$

The second term of (5.9) is the density of the kernel of $M(M^k)^+$. Thus, $(M^k)^+M = M(M^k)^+$ in this case and Corollary 3.1 applies. If P and Q are as defined in Theorem 3.1, then for $|s| < 1$ (that is $\|M\| = 1$)

$$P = \exp\left\{\sum_{k=1}^{\infty} \frac{(M^k)^+ s^k}{k}\right\} , \qquad Q = \exp\left\{\sum_{k=1}^{\infty} \frac{(M^k)^- s^k}{k}\right\} . \tag{5.10}$$

Since $(M^k)^-$ has a kernel with a density of the form $f_k(y-x)$, we deduce that Q must have a kernel with a density of the form $q(y-x)$. This means

$$q(x\,; S) = \int_{-\infty}^{\infty} q(y-x)dy = \exp\left[\sum_{k=1}^{\infty} \frac{P\{S_k \leq 0\}}{k} s_k\right] \tag{5.11}$$

is independent of x and Corollary 4.1 applies. Spitzer's identity (1.1) is found in this case from (4.10) by operating with each side on the function $g(y) = \exp(ity)$. In fact, in the notation of (1.1)

$$\begin{aligned} \Phi g &= e^{itx} \sum_{n=0}^{\infty} \int_{-\infty}^{\infty} e^{it(y-x)} P\{\max(T_0, \cdots, T_n) \in dy \mid T_0 = x\} s^n \\ &= e^{itx} \sum_{n=0}^{\infty} \int_{-\infty}^{\infty} e^{ity} P\{\max(0, S_1, \cdots, S_n) \in dy\} s^n \\ &= e^{itx} \sum_{n=0}^{\infty} \phi_n(t) s^n . \end{aligned} \tag{5.12}$$

Now in the special case of the exponential function $g(y) = e^{ity}$,

$$(M^k)^+(M^n)^+ge^{-itx} = [(M^k)^+ge^{-itx}][(M^n)^+ge^{-itx}] . \tag{5.13}$$

From (6.10), we find[1]

$$\begin{aligned} Pg &= e^{itx}\exp\left\{\sum_{k=1}^{\infty}\frac{(M^k)^+ge^{-itx}}{k}s^n\right\} \\ &= e^{itx}\exp\left[\sum_{k=1}^{\infty}\frac{s^k}{k}\int_0^{\infty}e^{ity}P\{S_k \in dy\}\right] . \end{aligned} \tag{5.14}$$

Putting (5.11), (5.12), and (5.14) into (4.10), it follows that

$$\sum_{n=1}^{\infty}\varphi_n(t)s^n = \exp\left[\sum_{k=1}^{\infty}\frac{s^k}{k}\int_0^{\infty}e^{ity}P\{\max(0, S_k) \in dy\}\right] . \tag{1.1}$$

In passing we note that the existence of a density is convenient but not necessary for the derivation of (1.1) from (4.10). In general, we can replace (5.5) by

$$m(x\,;A) = P\{(X_1 + x) \in A\} , \tag{5.15}$$

which is Borel measurable in x for each fixed set A. The conditions (4.1) are satisfied and the derivation continues in the obvious manner.

EXAMPLE 3. Let M be a matrix of finite order. We denote by D_k the subdeterminant formed from the determinant of $I - sM$ by crossing out all but the first k rows and columns. Moreover, $D_k(i\,;j)$ $(1 \leq i, j \leq k)$ will denote the cofactor of the ijth element in D_k. Finally, for any matrix N, let $N(k)$ denote the matrix formed from N by crossing out all but the first k rows and columns.

Let $\{P_n\}$, $\{Q_n\}$, $P = (p_{ij})$, and $Q = (q_{ij})$ denote the matrices defined by (3.4) and (3.5) when Theorem 3.1 is applied to M. We may also apply Theorem 3.1 to $M(k)$. It is not hard to show by induction that $\{P_n(k)\}$, $\{Q_n(k)\}$, $P(k)$, and $Q(k)$ are the matrices defined by (3.4) and (3.5) when Theorem 3.1 is applied to $M(k)$. Thus, by (3.8)

$$P(k)Q(k) = [I(k) - sM(k)]^{-1} . \tag{5.16}$$

Equating elements of the last row (the kth row) in the matrix product of (5.16), we find

$$q_{kj} = D_k(j\,;k)/D_k , \qquad j = 1, 2, \cdots, k . \tag{5.17}$$

Using (5.17) and the elements of the last column of the product in (5.16), it follows that

$$p_{ik} = D_k(k\,;i)/D_{k-1} , \qquad i = 1, 2, \cdots, k . \tag{5.18}$$

[1] The referee points out that (5.14) holds if and only if g is the exponential function.

Let M be the transition probability matrix of a stationary Markov chain $\{X_k, k \geq 0\}$ with states $\alpha_1 < \alpha_2 < \cdots < \alpha_N$. From (4.3), we find

$$P\{L_n = n, X_n = \alpha_j \mid X_0 = \alpha_i\} = D_j(j\,;\,i)/D_{j-1}\,, \qquad (i \leq j)\,, \tag{5.19}$$
$$P\{L_n = 0, X_n = \alpha_j \mid X_0 = \alpha_i\} = D_i(j\,;\,i)/D_i\,, \qquad (i \leq j)\,.$$

References

1. E. Sparre Andersen, *On the fluctuations of sums of random variables*, Math. Scand. **1** (1953), 263-285.

2. ———, *On the fluctuations of sums of random variables* II, Math. Scand. **2** (1954), 195-223.

3. J. L. Doob, *Stochastic process*, New York 1953.

4. P. R. Halmos, *Measure theory*, New York 1958.

5. Frank Spitzer, *A combinatorial lemma and its application to probability theory*, Trans. Amer. Math. Soc. **82** (1956), 323-339.

University of Minnesota

AUTOMORPHIC GROUP REPRESENTATIONS

ROBERT J. BLATTNER

1. Introduction In this paper we investigate certain representations of groups as * -automorphisms of rings of operators. More particularly, we are interested in finding conditions on the group, representation, and ring which guarantee the production of outer automorphisms of the ring. The exhibition of outer automorphisms has been considered before, notably by Singer in a paper [9] which intensively analyses the automorphism group of one of the finite factors constructed by von Neumann. Although we also shall be concerned with finite rings, our results do not overlap Singer's.

Segal, in [8], introduced the notion of skew distribution over a real Hilbert space $\mathfrak{H}$. He singled out one in particular, the Clifford distribution over $\mathfrak{H}$, which admits a representation Γ of the orthogonal group $\mathcal{O}(\mathfrak{H})$ of $\mathfrak{H}$ into the automorphism group of the ring $\mathfrak{A}$ associated to the distribution. Section 2 of the present paper gives (mostly without proofs) a variant of Segal's construction of $\mathfrak{A}$ and Γ, which is more suitable for our calculations. Section 3 states and begins the proof of Theorem 1, which completely classifies (*vis à vis* innerness) the automorphisms of $\mathfrak{A}$ arising from Γ. The proof is completed in § 4 and 5.

In § 6, we introduce the notion of a continuous automorphic group representation and show that any locally compact group satisfying the second axiom of countability may be represented as outer automorphisms of the Clifford distribution ring. Finally, Theorem 2 shows that any continuous automorphic representation of an open simple Lie group on a finite ring is essentially outer.

We shall make free use of the standard theory of operators and rings of operators as found in [6] and [3]. For the theory of measurable operators and gage spaces see [7].

The author would like to thank I. E. Segal for bringing the problems solved in this paper to his attention.

2. Preliminaries. Let $\mathfrak{H}$ be a real Hilbert space, $\mathfrak{T}$ the tensor algebra over $\mathfrak{H}$, $\mathfrak{J}$ the ideal generated by elements of the form $x \otimes x - (x, x)1$. Set $\mathfrak{C} = \mathfrak{T}/\mathfrak{J}$, the Clifford algebra over $\mathfrak{H}$ with respect to the quadratic form $x \to (x, x)$, and $* =$ the main anti-automorphism of $\mathfrak{C} =$ the anti-automorphism of $\mathfrak{C}$ arising from the anti-automorphism of $\mathfrak{T}$ which sends

$$x_1 \otimes \cdots \otimes x_n \to x_n \otimes \cdots \otimes x_1 .$$

Received April 10, 1958.

As usual, we will consider $\mathfrak{H}$ as embedded in $\mathfrak{C}$.

Any central linear functional θ on $\mathfrak{C}$ has the property that $\theta(x_1 \cdots x_n) = 0$ whenever $\{x_1, \cdots, x_n\}$ are an orthogonal set in $\mathfrak{H}$. Since those elements, together with 1, span $\mathfrak{C}$, there is (up to a multiplicative factor) at most one such θ on $\mathfrak{C}$. Let us produce one. Following Chevalley [2], we let $\mathfrak{E}$ = the exterior algebra over $\mathfrak{H}$, multiplication indicated by $\wedge$, and $\mathscr{M}$ = the algebra of endomorphisms of $\mathfrak{E}$. $\mathfrak{H}$ is considered embedded in $\mathfrak{E}$. For each $x \in \mathfrak{H}$, let δ_x be the unique antiderivation of $\mathfrak{E}$ such that $\delta_x y = (x, y)1$ for all $y \in \mathfrak{H}$, and let l_x be the operator of left multiplication by x in $\mathfrak{E}$. The mapping $x \to l_x + \delta_x$ of $\mathfrak{H}$ into $\mathscr{M}$ extends to a homomorphism Ψ of $\mathfrak{C}$ into $\mathscr{M}$. We let $\tau(u) = \Psi(u)1$ for $u \in \mathfrak{C}$. It is easy to show that τ is a one-to-one linear map of $\mathfrak{C}$ onto $\mathfrak{E}$.

Now the inner product $(\cdot\,,\cdot)$ on $\mathfrak{H}$ extends to a real Hilbert inner product, also called $(\cdot,\cdot)$, on $\mathfrak{E}$. We set $\theta(u) = (\tau(u), 1)$. It is clear that θ is linear and that $\theta(1) = 1$. We shall show that $\theta(v^* u) = (\tau(u), \tau(v))$. This will establish the centrality of θ and will also show that $(u, v) = \theta(v^* u)$ is a Hilbert inner product on $\mathfrak{C}$ making τ into an isometry. It suffices to prove the above when $u = x_1 \cdots x_n y_1 \cdots y_r$ and $v = x_1 \cdots x_n z_1 \cdots z_s$ where $\{x_1, \cdots, x_n, y_1, \cdots, y_r, z_1, \cdots, z_s\}$ (n, r, s possibly 0) form an orthonormal set in $\mathfrak{H}$, since the $u \otimes v$ for all such pairs u, v span $\mathfrak{C} \otimes \mathfrak{C}$. But

$$\begin{aligned}\theta(v^* u) &= \theta(z_s \cdots z_1 x_n \cdots x_1 x_1 \cdots x_n y_1 \cdots y_r) = \theta(z_s \cdots z_1 y_1 \cdots y_r) \\ &= z_s \wedge \cdots \wedge z_1 \wedge y_1 \wedge \cdots \wedge y_r, 1) = 1 \text{ or } 0\end{aligned}$$

according as $r = s = 0$ or not. Thus

$$\begin{aligned}\theta(v^* v) &= (x_1 \wedge \cdots \wedge x_n \wedge y_1 \wedge \cdots \wedge y_r, x_1 \wedge \cdots \wedge x_n \wedge z_1 \wedge \cdots \wedge z_s) \\ &= (\tau(u), \tau(v))\,,\end{aligned}$$

as desired.

Let $\mathfrak{D}$ be the complexification of $\mathfrak{C}$ and extend the inner product on $\mathfrak{C}$ in the usual way to a (complex) Hilbert inner product on $\mathfrak{D}$. Let *$\mathfrak{K}$ be the completion of $\mathfrak{D}$. * may be extended by conjugate linearity* and closure to be a conjugation on $\mathfrak{K}$. We note that if $\{e_i\}$ is an orthonormal basis for $\mathfrak{H}$, then $\{e_{i_1} e_{i_2} \cdots e_{i_r}\}(i_1 < i_2 < \cdots < i_r\,;\ r = 0, 1, \cdots)$ is an orthonormal base for $\mathfrak{K}$, where the indices i have been linearly ordered in some fashion. We shall adopt the notation e_A, A a finite set of indices, to mean $e_{i_1} e_{i_2} \cdots e_{i_r}$ where $i_1 < \cdots < i_r$ and $A = \{i_1, \cdots, i_r\}$. Conventionally $e_\phi = 1$.

For any element $u \in \mathfrak{D}$, let L''_u be the operator with domain $\mathfrak{D}$ defined by $L''_u a = ua$, $a \in \mathfrak{D}$. It is easily seen that L''_x, x a unit vector of $\mathfrak{H}$, is an isometry of $\mathfrak{D}$ onto $\mathfrak{D}$. Since $\mathfrak{D}$ is spanned algebraically by products of the form $x_1 \cdots x_n$, x_i unit vectors in $\mathfrak{H}$, we conclude that L''_u, $u \in \mathfrak{D}$, is a continuous operator on the normed linear space $\mathfrak{D}$. Thus $(\mathfrak{K}, \mathfrak{D}, {}^*)$ forms a Hilbert algebra, the left ring $\mathfrak{A}$ of which is a factor of type II, when $\mathfrak{H}$ is infinite dimensional, which is the only case we shall consider [8]. Let $\mathfrak{B}$ be the algebra of all bounded elements of $(\mathfrak{K}, \mathfrak{D}, {}^*)$. L_a and R_a, $a \in \mathfrak{B}$, will denote the closure of the left and right multiplication operators respectively by a on the domain $\mathfrak{B}$. The maps $a \to L_a$ and $a \to R_a$ are an isomorphism and anti-isomorphism of $\mathfrak{B}$ onto $\mathfrak{A}$ and $\mathfrak{A}'$ respectively. For $x \in \mathfrak{K}$, we define L'_x as the operator with domain $\mathfrak{B}$ such that $L'_x a = R_a x$, $a \in \mathfrak{B}$. Then L_x is defined to be $(L'_{x^*})^*$, the notation agreeing with the above when $x \in \mathfrak{B}$. L_x is always measurable with respect to $\mathfrak{A}$ [7].

3. The representation Γ. Any orthogonal transformation U on $\mathfrak{H}$ extends canonically to an automorphism of $\mathfrak{T}$ which leaves $\mathfrak{J}$ invariant and thus induces an automorphism of $\mathfrak{C}$, which we denote $\Gamma(U)$. $\Gamma(U)$ is defined by

$$\Gamma(U)(x_1 \cdots x_n) = (Ux_1 \cdots Ux_n) \text{ for } x_1, \cdots, x_n \in \mathfrak{H} .$$

Clearly $\Gamma(U)$ commutes with *. The functional $\theta \circ \Gamma(U)$ is again a central linear functional on $\mathfrak{C}$ and $\theta(\Gamma(U)1) = \theta(1) = 1$. Hence $\theta \circ \Gamma(U) = \theta$ so that $\Gamma(U)$ is an isometry on $\mathfrak{C}$. The automorphism $\Gamma(U)$ then extends to an automorphism of $(\mathfrak{K}, \mathfrak{D}, {}^*)$ which leaves $\mathfrak{D}$ invariant. Γ is clearly a faithful representation of the orthogonal group of $\mathfrak{H}$ into the automorphism group of $(\mathfrak{K}, \mathfrak{D}, {}^*)$. Ω will denote the automorphism $\Gamma(-I)$. $\Omega^2 = I$, so that $\mathfrak{K}$ is the direct sum of two subspaces $\mathfrak{K}^+$ and $\mathfrak{K}^-$ defined by $\Omega x = x$ or $-x$ according as $x \in \mathfrak{K}^+$ or $\mathfrak{K}^-$.

$\Gamma(U)$ is an *inner* automorphism if there exists a unitary element $u \in \mathfrak{B}$ such that $\Gamma(U) = L_u R_{u^*}$. Since $\mathfrak{A}$ is a factor, u is determined up to a multiplicative constant of modulus 1 by $\Gamma(U)$. Now

$$\begin{aligned}\Gamma(U) &= \Gamma((-I)U(-I)) = \Omega L_u R_{u^*} \Omega = (\Omega L_u \Omega)(\Omega R_{u^*} \Omega) \\ &= L_{\Omega u} R_{\Omega u^*} = L_{\Omega u} R_{(\Omega u)^*} .\end{aligned}$$

Hence $\Omega u = \lambda u$, λ a constant. Clearly $\lambda = \pm 1$ so that either $u \in \mathfrak{K}^+$ or $u \in \mathfrak{K}^-$. In the former case $\Gamma(U)$ is called *even*, in the latter, *odd*. Those inner automorphisms $\Gamma(U)$ which are even form a subgroup of the group of all inner automorphisms of the type $\Gamma(U)$.

In order to classify the automorphisms $\Gamma(U)$ according to the above categories, we introduce the following notation: Let $\mathscr{G}^+$ be the set of orthogonal transformations U on $\mathfrak{H}$ such that $I - U$ is of Hilbert-Schmidt class and whose eigenspace belonging to -1 has even dimension; let $\mathscr{G}^-$ contain all those U such that $I + U$ is Hilbert-Schmidt and whose eigenspace belonging to $+1$ has odd dimension. Set $\mathscr{G}_0 = \mathscr{G}^+ \cup \mathscr{G}^-$.

THEOREM 1. *$\Gamma(U)$ is inner if and only if $U \in \mathscr{G}_0$. If $\Gamma(U)$ is inner, it is even if and only if $U \in \mathscr{G}^+$.*

Let $\{e_i\}_{i \in J}$ be an orthonormal basis of $\mathfrak{H}$, where J is a totally ordered index set, and let U be a fixed orthogonal transformation. Set $f_i = Ue_i$ and $V_i = L_{f_i}R_{e_i}$.

LEMMA 1. *The subspace $\mathfrak{L}$ of vectors left invariant by all the V_i has dimension $= 0$ or 1. If* $\dim \mathfrak{L} = 0$, *$\Gamma(U)$ is outer; if* $\dim \mathfrak{L} = 1$, *$\mathfrak{L}$ contains a unitary element $u \in \mathfrak{B}$ such that $\Gamma(U) = L_u R_{u^*}$.*

Proof. Suppose $\Gamma(U) = L_u R_{u^*}$, u a unitary in $\mathfrak{B}$. Then $\Gamma(U)e_i = L_u R_{u^*} e_i$ or $f_i = Ue_i = ue_iu^*$, all i. Thus $f_iue_i = u$ or $V_iu = u$, all i. Therefore $\dim \mathfrak{L} \geqq 1$ if $\Gamma(U)$ is inner.

Next suppose $\dim \mathfrak{L} \geqq 1$ and let $0 \neq x \in \mathfrak{L}$. Then $R_{f_i}x^* = L_{e_i}x^*$ so that $L'_{L_{e_i}x^*} = L'_{R_{f_i}x^*}$. For any element $a \in \mathfrak{B}$, $L'_{L_{e_i}x^*}a = R_a L_{e_i}x^* = L_{e_i}R_a x^* = L_{e_i}L'_{x^*}a$ so that $L'_{L_{e_i}x^*} = L_{e_i}L'_{x^*}$. Similarly $L'_{R_{f_i}x^*} = L'_{x^*}L'_{f_i}$. Taking adjoints, we have

$$(L'_{L_{e_i}x^*})^* \supseteqq (L'_{x^*})^* L_{e_i}^{\;*} = L_x L_{e_i} \text{ and } (L'_{R_{f_i}x^*})^* \supseteqq (L'_{f_i})^*(L'_{x^*})^* = L_{f_i}L_x .$$

Therefore, using $\cdot$ to indicate strong product [7], $L_x \cdot L_{e_i} = L_{f_i} \cdot L_x$. Again taking adjoints, $L_{e_i} \cdot L_x^* = L_x^* \cdot L_{f_i}$. This implies $L_{f_i} \cdot L_x \cdot L_x^* = L_x \cdot L_x^* \cdot L_{f_i}$; that is, the positive measurable operator $L_xL_x^*$ commutes with each L_{f_i}. Thus every spectral projection of $L_xL_x^*$ commutes with each L_{f_i}. But the $\{L_{f_i}\}$ are a self-adjoint set of generators for $\mathfrak{A}$ and $\mathfrak{A}$ is a factor. Therefore each spectral projection of $L_xL_x^*$ is either 0 or I, whence $L_xL_x^* = \lambda I$, λ a positive constant. A similar argument shows that $L_x^*L_x = \lambda I$. Thus we have shown that L_x is bounded and that $\lambda^{-1/2}L_x$ is unitary so that $\lambda^{-1/2}x$ is a unitary element in $\mathfrak{B} \cap \mathfrak{L}$.

Let $u = \lambda^{-1/2}x$. Then $f_iue_i = u$, and hence $ue_iu^* = f_i$. Therefore the automorphisms $\Gamma(U)$ and $L_uR_{u^*}$ agree on the $\{e_i\}$ which are a set of generators for $(\mathfrak{K}, \mathfrak{D}, {}^*)$; that is, $\Gamma(U) = L_uR_{u^*}$. Suppose now that $0 \neq y \in \mathfrak{L}$ also. Then $y \in \mathfrak{B}$, $yy^* = y^*y = \mu 1$, μ a positive constant. $v = \mu^{-1/2}y$ is also a unitary $\in \mathfrak{B} \cap \mathfrak{L}$ and $\Gamma(U) = L_vR_{v^*}$. But this implies that $v = \zeta u$, ζ a constant; that is, $y = \mu^{1/2}\zeta u$. Hence $\dim \mathfrak{L} = 1$.

DEFINITION 1. For any orthogonal transformation U on $\mathfrak{H}$, the subspace $\mathfrak{L}$ of $\mathfrak{K}$ is called its *characteristic subspace*.

It is clear from Lemma 1 that the characteristic subspace depends only on U and not on the choice of a basis $\{e_i\}$ for $\mathfrak{H}$.

4. The determinant condition. In this section we will show that if $\Gamma(U)$ is an even inner automorphism, then $I - U$ is Hilbert-Schmidt.

This will be achieved in a series of lemmas. We adhere to the notation of § 2.

LEMMA 2. *Let U be an orthogonal transformation on $\mathfrak{H}$ and let P be the projection on its characteristic subspace. Then*

$$\lim_{A} \det_{k,l\in A} (2^{-1}(I+U)e_k, e_l) = (P1, 1) .$$

(Here "lim" means limit according to the set of finite subsets A of J, directed upward by inclusion. The determinant is expanded with respect to the total order on the elements of A.)

Proof. We first introduce some notation: f_A will denote $\Gamma(U)e_A$; $P_i=$ the projection on the invariant subspace of V_i (see Lemma 1); if $A=\{i_1, \cdots, i_r\}(i_1<\cdots<i_r)$ then $P_A=P_{i_1}\cdots P_{i_r}$.

Since the V_i mutually commute, the P_A are projections which mutally commute. Clearly $P=$ strong-$\lim_A P_A$. In addition, $V_i^2=I$, all i. This imples that $P_i=2^{-1}(I+V_i)$; that is, $P_i a=2^{-1}(a+f_i a e_i)$ for $a\in\mathfrak{B}$. Iterating this, we calculate that

$$P_A 1 = 2^{-r}\sum_{B\subseteq A} f_B e_B^*$$

where r is the cardinality of A. Hence $(P_A 1, 1)=2^{-r}\sum_{B\subseteq A}(e_B, f_B)$.

Fix $B\subseteqq A$ and suppose $B=\{j_1,\cdots,j_s\}(j_1<\cdots<j_s)$. Then

$$\begin{aligned}(e_B, f_B) &= (\tau(e_B), \tau(f_B)) \text{ (see § 2)} = (e_{j_1}\wedge\cdots\wedge e_{j_s}, f_{i_1}\wedge\cdots f_{j_s}) \\ &= \det_{k,l\in B}(e_k, f_l) = \det_{k,l\in B}(e_k, Ue_l) = \det_{k,l\in B}(Ue_k, e_l) .\end{aligned}$$

Hence

$$(P_A 1, 1) = 2^{-r}\sum_{B\subseteq A}\det_{k,l\in B}(Ue_k, e_l) ,$$

which we recognize to be

$$\det_{k,l\in A}(2^{-1}(I+U)e_k, e_l) .$$

Passing to the limit on A, we have the lemma.

Note that the lemma shows that $\lim_A \det_{k,l\in A}(2^{-1}(I+U)e_k, e_l)$ depends only on U and not on the particular choice of basis. Hence we may write $\det(2^{-1}(I+U))$ without fear of confusion. This motivates the following.

DEFINITION 2. An operator T on $\mathfrak{H}$ will be said to *have a determinant* if, for every choice of an orthonormal basis $\{e_i\}_{i\in J}$ (J totally ordered), $\lim_A \det_{k,l\in A}(Te_k, e_l)$ exists and is independent of the choice of basis. We write $\det(T)$ for the common limit. (Cf. the treatment in [5].)

To make use of the conclusion of Lemma 2, we must prove a short preliminary result. For any operator T, $\sigma(T)$ will denote the spectrum of T.

LEMMA 3. *Let S be a self-adjoint operator on $\mathfrak{H}$ and let $\{S_\alpha\}$ be a net of self-adjoint operators and $\{Q_\alpha\}$ a net of projections (same index set) such that:*

(1) $S = \text{strong-lim}_\alpha S_\alpha$;
(2) $S_\alpha = Q_\alpha S_\alpha Q_\alpha$ and $I = \text{strong-lim}_\alpha Q_\alpha$.
Then $\sigma(S) \subseteqq$ topological $\liminf_\alpha \sigma(S_\alpha | Q_\alpha \mathfrak{H})$.

Proof. Let $\lambda \in \sigma(S)$ and $\varepsilon > 0$ be given. Since S is self-adjoint, we can find a unit vector $x \in \mathfrak{H}$ such that $\|(S - \lambda I)x\| < \varepsilon/4$. We may then find an index α_0 such that $\alpha > \alpha_0$ implies $\|(S_\alpha - S)x\| < \varepsilon/4$ and $\|Q_\alpha x\| \geqq 1/2$. Then $\|(S_\alpha - \lambda I)x\| < \varepsilon/2$, whence $\|(S_\alpha - \lambda Q_\alpha)(Q_\alpha x)\| < \varepsilon/2$. Set $y_\alpha = Q_\alpha x/\|Q_\alpha x\|$. Then y_α is a unit vector in $Q_\alpha \mathfrak{H}$. We have shown that $\|(S_\alpha - \lambda Q_\alpha)y_\alpha\| < \varepsilon$. This implies that $\sigma(S_\alpha | Q_\alpha \mathfrak{H})$ contains a point within ε of λ, $\alpha > \alpha_0$.

We shall apply this to the situation in the following lemma.

LEMMA 4. *Let the operator T on $\mathfrak{H}$ have a determinant* $\det(T) = c \neq 0$ *and suppose $\|T\| \leqq 1$. Then $I - T^*T$ is of trace class.*

Proof. Chose a basis $\{e_i\}_{i \in J}$, J totally ordered. We shall take as our index set the set of all finite subsets A of J. $Q_A =$ the projection on the subspace of $\mathfrak{H}$ spanned by the e_i, $i \in A$. Clearly $Q_A \to I$ strongly. Set $S = T^*T$, $S_A = Q_A T^* Q_A T Q_A$, and $T_A = Q_A T Q_A / Q_A \mathfrak{H}$. Then $S_A \to S$ strongly and $S_A / Q_A \mathfrak{H} = T_A^* T_A$. Lemma 3 asserts that $\sigma(T^*T) \subseteqq$ topological $\liminf_A \sigma(T_A^* T_A)$.

Now $\det(T_A^* T_A) = (\det T_A)^2$ so that the hypotheses of the lemma assert that $\lim_A \det(T_A^* T_A) = c^2 \neq 0$ and $\|T^*T\| \leqq 1$, implying each $\|T_A^* T_A\| \leqq 1$. Clearly $\sigma(T^*T)$ and $\sigma(T_A^* T_A) \subseteqq [0, 1]$. Given $\varepsilon < 1$, let $N(A, \varepsilon) =$ cardinality of $\sigma(T_A^* T_A) \cap [0, \varepsilon)$ and $N(\varepsilon) =$ cardinality of $\sigma(T^*T) \cap [0, \varepsilon)$. Choose a set A_0 such that $\det(T_A^* T_A) \geqq 2^{-1}c^2$ for $A \supseteqq A_0$. Since $\det(T_A^* T_A)$ is the product of the eigenvalues of $T_A^* T_A$, we must have that $N(A, \varepsilon) \leqq \log 2^{-1}c^2/\log \varepsilon$ for $A \supseteqq A_0$. Therefore $N(\varepsilon) \leqq \log 2^{-1}c^2/\log \varepsilon$. This shows that $\sigma(T^*T)$ is pure point spectrum except possibly for the value 1 and that $\sigma(T^*T)$ has only 1 as a cluster point.

$\mathfrak{H}$ is the direct sum of the eigenspaces of T^*T. Choose a new basis, again called $\{e_i\}_{i \in J}$, adapted to this direct decomposition of $\mathfrak{H}$. We use the notation of the previous paragraphs (with respect to the new basis). Each e_i is an eigenvector belonging to an eigenvalue λ_i of T^*T. For every finite subset A of J we have

$$\prod_{i \in A} \lambda_i = \det(T^*T \,|\, Q_A \mathfrak{H}) \ .$$

But

$$T^*T \,|\, Q_A\mathfrak{H} - T_A^* T_A = Q_A T^*(I - Q_A) T Q_A \,|\, Q_A \mathfrak{H} \ ,$$

a positive operator on $Q_A\mathfrak{H}$. Now the determinant of the sum of two positive operators on a finite dimensional Hilbert space is greater than the determinant of either operator. Hence

$$\prod_{i \in A} \lambda_i \geqq \det(T_A^* T_A) \ .$$

As $A \uparrow$, $\prod_{i \in A} \lambda_i \downarrow$ since $0 \leqq \lambda_i \leqq 1$ for all $i \in J$. It follows that $\prod_i \lambda_i$ exists. Moreover, $\prod_i \lambda_i \geqq c^2 > 0$ since $\lim_A \det(T_A^* T_A) = c^2$. Therefore $\prod_i \lambda_i$ converges absolutely, so that $\sum_i (1 - \lambda_i) < \infty$; that is, $I - T^*T$ is of trace class.

LEMMA 5. *Let $\Gamma(U)\, L_u R_{u^*}$, u a unitary operator of $\mathfrak{B} \cap \mathfrak{K}^+$. Then $I - U$ is Hilbert-Schmidt.*

Proof. Fix a basis $\{e_i\}_{i \in J}$, J totally ordered. Each e_i is a self-adjoint unitary element of $\mathfrak{B}$. Hence for any finite subset A of J, e_A (notation as in § 2) is a unitary in $\mathfrak{B}$. For each e_i we define U_{e_i} to be that orthogonal transformation on $\mathfrak{H}$ which leaves e_i invariant and multiplies the other elements of the basis by -1. It is easy to see that $\Gamma(U_{e_i}) = L_{e_i} R_{e_i}$. In general, we define U_{e_A} to be $U_{e_{i_1}} \cdots U_{e_{i_r}}$ where $A = \{i_1, \cdots, i_r\} (i_1 < \cdots < i_r)$. Then $\Gamma(U_{e_A}) = L_{e_A} R_{e_A^*}$; and $U_{e_A} \in \mathscr{G}^+$ or $\mathscr{G}^-$ and $\Gamma(U_{e_A})$ is even or odd according as A has even or odd cardinality. It is clear that all the U_{e_A} are self-adjoint.

Let $\Gamma(U) = L_u R_{u^*}$, u a unitary element of $\mathfrak{B} \cap \mathfrak{K}^+$. Then $u = \sum_A \lambda_A e_A$, the summation being extended over the A of even cardinality. Pick a $\lambda_B \neq 0$. Then $u = e_B(\sum_A \lambda'_A e_A)$; λ'_A 's different constants, Setting $v = \sum_A \lambda'_A e_A$, v is a unitary in $\mathfrak{B} \cap \mathfrak{K}^+$ such that $(v, 1) = \lambda'_\phi = \lambda_B \neq 0$. Set $V = U_{e_B} U$. Then

$$\Gamma(V) = \Gamma(U_{e_B})\Gamma(U) = (L_{e_B^*} R_{e_B})(L_u R_{u^*}) = L_{e_B^* u} R_{(e_B^* u)^*} = L_v R_{v^*} \ .$$

Since $U_{e_B} \in \mathscr{G}^+$, $I - U$ will be Hilbert-Schmidt if and only if $I - V$ is. Thus we may assume without loss of generality that $(u, 1) \neq 0$.

If P is the projection on the characteristic subspace of U, our assumption implies that $(P1, 1) \neq 0$. Setting $T = 2^{-1}(I + U)$, we conclude from Lemmas 2 and 4 that $I - T^*T$ is of trace class. This says that

$$I - \frac{1}{4}\Big(I + U^*\Big)\Big(I + U\Big) = \frac{1}{4}\Big(2I - U - U^*\Big)$$

is of trace class. Hence $2I - U - U^* = (I - U)^*(I - U)$ is of trace class so that $I - U$ is Hilbert-Schmidt.

5. Completion of proof.

LEMMA 6. *If* $U \in \mathscr{G}^+$ *[respectively* $\mathscr{G}^-$*], then* $\Gamma(U)$ *is an even [odd] inner automorphism.*

Proof. We use the notation of Lemma 5. Let $U \in \mathscr{G}^+$ [respectively $\mathscr{G}^-$]. Then the eigenspace $\mathfrak{M}$ of U belonging to the eigenvalue -1 [$+1$] is of even [odd] dimension. Let $\{e_i\}_{i \in A}$ be a basis for $\mathfrak{M}$. Then $U_{e_A} \in \mathscr{G}^+$ [$\mathscr{G}^-$]. Set $V = U_{e_A}U$. It is easily seen that $V \in \mathscr{G}^+$ and that the eigenspace belonging to -1 has dimension 0. If the lemma can be proved for V, it will follow for U since $U = U_{e_A}V$ will then be the product of an even [odd] and an even inner automorphism. Thus we may assume without loss of generality that $U \in \mathscr{G}^+$ and that U has no eigenvectors belonging to -1.

$\mathfrak{H}$ is the direct sum of the eigenspaces of $(I - U)^*(I - U)$. These spaces are all finite dimensional except possibly that belonging to 0. These subspaces all reduce U and on the 0-eigenspace $U = I$. Using the classical reduction of an orthogonal transformation on a Euclidean space and remembering that U has no eigenvectors belonging to -1, we see that $\mathfrak{H}$ is the direct sum of a countable number of 2-dimensional subspaces $\mathfrak{H}_n (n = 1, 2, \cdots)$ and a subspace $\mathfrak{H}_0$, each of which reduces U, and such that U is irreducible on every $\mathfrak{H}_n$ and $U = I$ on $\mathfrak{H}_0$. Let $\{e_i\}_{i \in J}$ be a basis for $\mathfrak{H}$ adapted to this direct decomposition. With respect to this basis

$$U|\mathfrak{H}_n = \begin{pmatrix} \cos\theta_n & -\sin\theta_n \\ \sin\theta_n & \cos\theta_n \end{pmatrix}, \qquad 0 < \theta_n < \pi$$

(where the basis elements of $\mathfrak{H}$ have been suitably ordered). We readily calculate $\det(2^{-1}(I + U))$ to be

$$\prod_n \det\left(\frac{1}{2}\Big(I + U\Big)\Big|\mathfrak{H}_n\right) = \prod_n \left(\frac{1 + \cos\theta_n}{2}\right).$$

For any operator T, we denote its Hilbert-Schmidt norm by $\|T\|_2$. Then

$$\|I - U\|_2^2 = \sum_n \|(I - U)|\mathfrak{H}_n\|_2^2 = 4\sum_n (1 - \cos\theta_n) = 8\sum_n \left(1 - \frac{1 + \cos\theta_n}{2}\right).$$

Hence

$$\prod_n \left(\frac{1 + \cos\theta_n}{2}\right)$$

converges absolutely. By Lemma 2, this says that $(P1, 1) \neq 0$, where P is the projection on the characteristic subspace of U. Thus $P \neq 0$ and $\Gamma(U)$ is inner by Lemma 1. Finally, $\Gamma(U)$ must be even. In fact, suppose $\Gamma(U) = L_u R_{u^*}$, u unitary in $\mathfrak{B} \cap \mathfrak{K}^-$. Since P is the projection on the subspace generated by u and since $1 \in \mathfrak{K}^+$, we conclude $P1 = 0$, a contradiction.

LEMMA 7. *If $\Gamma(U)$ is inner, then $U \in \mathscr{G}_0$.*

Proof. Suppose $\Gamma(U)$ is odd. Let e_1 be a unit vector in $\mathfrak{H}$. Then $\Gamma(U_{e_1})$ (notation as in Lemma 5) is odd. It follows that $\Gamma(U_{e_1}U)$ is even, so that $I - U_{e_1}U$ is Hilbert-Schmidt. Since $I + U_{e_1}$ is Hilbert-Schmidt, so is $I + U = (I + U_{e_1}) - U_{e_1}(I - U_{e_1}U)$.

Let now $\Gamma(U)$ be even [odd]. We know that $I - U$ [respectively $I + U$] is Hilbert-Schmidt. Suppose that the eigenspace of U belonging to -1 [$+1$] is of odd [even] dimension. Then $-U \in \mathscr{G}^-$[respectively $\mathscr{G}^+$] so that $\Gamma(-U)$is also inner by Lemma 6. Therefore $\Omega = \lambda\Gamma(U^*(-U))$ is inner. We shall be through if we can show that Ω is outer.

Suppose then that Ω is inner. Since $I - (-I)$ is not Hilbert-Schmidt, Lemma 5 implies that Ω must be odd. Let $\Omega = L_u R_{u^*}$, u unitary in $\mathfrak{B} \cap \mathfrak{K}^-$. Picking a basis $\{e_i\}_{i \in J}$, J totally ordered, for $\mathfrak{H}$, $u = \sum \lambda_A e_A$, the summation being extended over all finite $A \subseteqq J$ of odd cardinality. We let $\mathfrak{L}$ be the characteristic subspace of $-I$ and adopt the notation of Lemmas 1 and 2. We have $P_i u = u$ for each $i \in J$. Now

$$P_i u = \sum_A \lambda_A \cdot \frac{1}{2}\left(e_A - e_i e_A e_i\right).$$

Since A has odd cardinality, $e_i e_A e_i = e_A$ or $-e_A$ according as $i \in A$ or $i \notin A$. Hence $\lambda_A = 0$ unless $i \notin A$ for all i; that is, all $\lambda_A = 0$, which is ridiculous. This concludes the proof of Lemma 7 and, with it, Theorem 1.

6. Automorphic representations of topological groups. The mapping Γ is a representation of the orthogonal group $\mathscr{O}(\mathfrak{H})$ of $\mathfrak{H}$ as automorphisms of the Hilbert algebra ($\mathfrak{H}$, D, $*$). Every automorphism of the Hilbert algebra gives rise to an automorphism of its left ring $\mathfrak{A}$ via the isomorphism $b \leftrightarrow L_b$ of the bounded algebra $\mathfrak{B}$ and $\mathfrak{A}$. Conversely, every $*$ -automorphism of $\mathfrak{A}$ gives rise to one of the Hilbert algebra (by the uniqueness of the normalized central trace on $\mathfrak{B}$) and the correspondance is univalent. Henceforth we identify these two types of automorphisms.

Let $\mathscr{A}$ be the $*$ -automorphism group of $\mathfrak{A}$. If $\alpha \in \mathscr{A}$, $T \in \mathfrak{A}$, T^α denotes the image of T by α. We have then

$$L_b^{\Gamma(u)} = L_{\Gamma(u)b} = \Gamma(U) L_b \Gamma(U)^*$$

for $U \in \mathscr{O}(\mathfrak{H})$ and $b \in \mathfrak{B}$. Now it is easy to check that the maps $U \to \Gamma(U)b$, $b \in \mathfrak{D}$, are continuous in the norm of $\mathfrak{R}$ when $\mathscr{O}(\mathfrak{H})$ is given the strong operator topology (as henceforth it shall be). Hence Γ is continuous from $\mathscr{O}(\mathfrak{H})$ to the unitaries of $\mathfrak{R}$ in the strong topology. It follows that, for each $T \in \mathfrak{A}$, the map

$$U \to T^{\Gamma(u)} = \Gamma(U)T\Gamma(U)^*$$

is continuous from $\mathscr{O}(\mathfrak{H})$ to $\mathfrak{A}$ in the strong, and hence the weak topology. Γ is thus a continuous automorphic representation, in the following sense.

DEFINITION 3. Let $\mathfrak{A}$ be a ring of operators, $\mathscr{A}$ its * -automorphism group, G a topological group. A representation ρ of the abstract group G into $\mathscr{A}$ is called a *continuous automorphic representation on* $\mathfrak{A}$ if, for every $T \in \mathfrak{A}$, the map $g \to T^{\rho(g)}$ of G into $\mathfrak{A}$ (in the weak topology) is continuous.

This continuity restriction is the weakest that can reasonably be imposed on ρ and is independent of the particular spacial representation of $\mathfrak{A}$. We note that if ρ is continuous in the above sense, then $g \to T^{\rho(g)}$ is strongly continuous. In fact, let $g_\alpha \to g$. Then

$$\begin{aligned}&(T^{\rho(g_\alpha)} - T^{\rho(g)})^* \ (T^{\rho(g_\alpha)} - T^{\rho(g)})\\ &= (T^*T)^{\rho(g_\alpha)} - (T^*)^{\rho(g_\alpha)}T^{\rho(g)} - (T^*)^{\rho(g_\alpha)}T^{\rho(g)} + (T^*T)^{\rho(g)} \to 0 \text{ weakly.}\end{aligned}$$

Let now G be a topological group, τ a continuous representation of G into $\mathscr{O}(\mathfrak{H})$, $\mathfrak{A}$ the left ring of our Hilbert algebra, and $\mathscr{A}$ its automorphism group. Any $\alpha \in \mathscr{A}$ which leaves $\mathfrak{H}$ invariant will be called *special*. $\rho = \Gamma \circ \tau$ is then a continuous representation of G as special automorphisms. Conversely, let ρ be a special continuous automorphic representation on $\mathfrak{A}$. Then $\rho = \Gamma \circ \tau$, where τ is a representation of the abstract group G into $\mathscr{O}(\mathfrak{H})$ (merely restrict ρ to $\mathfrak{H}$). Let $a, b \in \mathfrak{H}$. Then $g \to (\tau(g)a, b) = (L_a^{\rho(g)}1, b)$ is continuous so that τ is continuous.

With this in mind, we see that Theorem 1 has the following easy consequence.

COROLLARY. *Let G be a locally compact group satisfying the second axiom of countability. Then G has a special continuous automorphic representation ρ such that if $\rho(g)$ is inner, then $g = e$.*

Proof. Let τ_1 be a faithful continuous representation of G into some orthogonal group $\mathscr{O}(\mathfrak{H}_1)$, $\mathfrak{H}_1$ a real Hilbert space of countable dimension; e.g., the left regular representation of G (real functions). Let $\aleph$ be the cardinality of a basis for $\mathfrak{H}$. Then $\mathfrak{H}$ is the $\aleph$-fold direct copy of $\mathfrak{H}_1$ so that the direct sum of $\aleph$ copies of τ_1, call it τ, can be taken as a representatation of G into $\mathscr{O}(\mathfrak{H})$. Since $\aleph$ is infinite, it is clear that $\tau(g)\varepsilon\mathscr{E}_0$ if and only if $\tau(g) = I$; that is, $g = e$. Set $\rho = \Gamma \circ \tau$.

It can also be deduced from Theorem 1 that for any special continuous automorphic representation ρ of an open simple Lie group on $\mathfrak{A}$, ρ^{-1} (inner automorphisms) is central. But this will follow from the more general statement in Theorem 2.

THEOREM 2. *Let G be an open simple Lie group, $\mathfrak{A}$ a finite ring of operators, and ρ a non-trivial continuous automorphic representation of G on $\mathfrak{A}$. Then $\rho(g)$ is inner only if g is central.*

The essential step is Lemma 8 below. It, together with an extension of a result of Kadison and Singer [4] to continuous projective unitary representation (for definitions, see below), will imply our theorem. Let $\mathfrak{A}$ be any ring of operators, $\mathscr{A}$ its automorphism group, $\mathscr{A}_0$ the subgroup of inner automorphisms, and $\mathscr{U}$ the group of unitary operators in $\mathfrak{A}$. The map $\pi: \mathscr{U} \to \mathscr{A}_0$ defined by $T^{\pi(U)} = UTU^*$, $U \in \mathscr{U}$, $T \in \mathfrak{A}$, is a homomorphism onto $\mathscr{A}_0$ whose kernel is the center $\mathscr{C}$ of $\mathscr{U}$. If $\mathscr{U}$ is given the strong operator topology, then π is a continuous automorphic representation. We set $\mathscr{P} = \mathscr{U}/\mathscr{C}$ with the topology induced from $\mathscr{U}$ and, with some abuse of language, call $\mathscr{P}$ the *projective unitary group of* $\mathfrak{A}$; any continuous group representation into $\mathscr{P}$ will be called a *continuous projective unitary representation on* $\mathfrak{A}$. π clearly induces a continuous isomorphism $\bar{\pi}$ of $\mathscr{P}$ onto $\mathscr{A}_0$.

Let τ be a continuous projective unitary representation of the topological group G on $\mathfrak{A}$. Then $\rho = \bar{\pi} \circ \tau$ is an inner continuous automorphic representation of G on $\mathfrak{A}$. Lemma 8 gives a partial converse.

LEMMA 8. *Let G be a locally compact group satisfying the second axiom of countability and let $\mathfrak{A}$ be a ring of operators on a separable Hilbert space $\mathfrak{H}$. Let ρ be an inner continuous automorphic representation of G on $\mathfrak{A}$. Then ρ is induced by a continuous projective unitary repretation on $\mathfrak{A}$.*

Proof. Let $\mathfrak{A}_1$ be the unit sphere of $\mathfrak{A}$ in the weak topology. $\mathfrak{A}_1$ is a compact separable metric space (since $\mathfrak{H}$ is separable). G is a complete separable metric space and bears its left Haar measure μ. Let $\{T_i\}$ be a dense sequence in $\mathfrak{A}_1$, and let $\mathfrak{S}$ be the set of all $(g, U) \in G \times \mathfrak{A}_1$ such that U is unitary and $UT_i = T_i^{\rho(g)}U$ for all i. The argument of Lemma 2(b), section 6, § 2, Chapter II of [3] shows that $\mathfrak{S}$ is a Borel, and hence an analytic, subset of $G \times \mathfrak{A}_1$. Let ζ be the canonical projection of $G \times \mathfrak{A}_1$ on G. Since ρ is inner, $\zeta(\mathfrak{S}) = G$. By Appendix V of [3], there exists a measurable map (in the sense of Bourbaki [1]) $g \to U_g$ of G into $\mathfrak{A}_1$ such that $(g, U_g) \in \mathfrak{S}$, $g \in G$.

Thus we have a measurable map $g \to U_g$ of G into $\mathscr{U}$ in the weak, and hence strong, topology such that $T_i^{\rho(g)} = U_g T_i U_g^*$, all i. Taking weak limits of $\{T_i\}$, we have $T^{\rho(g)} = U_g T U_g^*$ for all $T \in \mathfrak{A}_1$, and hence in $\mathfrak{A}$. Let ψ be the canonical projection of $\mathscr{U}$ on $\mathscr{P}$ and set $\tau(g) = \psi(U_g)$. We know that $g \to \rho(g) = \pi(U_g) = \bar{\pi}(\tau(g))$ is a representation. Since $\bar{\pi}$ is univalent, τ is a representation. Since ψ is continuous, τ is measurable from G into $\mathscr{P}$. Hence there exists a compact $K \subseteq G$ such that $\mu(K) > 0$ and $\tau \mid K$ is continuous. Using the representation property of τ we see easily that $\tau \mid KK^{-1}$ is continuous. But KK^{-1} is a neighborhood of e in G [10]. Therefore τ is continuous, as desired.

Proof of Theorem 2. For any ring of operators, the group of inner automorphisms is invariant in the full group of automorphisms. Let G be an open simple Lie group and C its center. Since every proper abstract invariant subgroup of G is contained in C, the theorem asserts the non-existence of non-trivial inner continuous automorphic representations of G on finite rings. Suppose, on the contrary, that ρ is such a representation on the finite ring $\mathfrak{A}$ on the Hilbert space $\mathfrak{H}$. We shall first show that $\mathfrak{H}$ may be assumed separable.

For eace $g \in G$ choose a unitary $U_g \in \mathfrak{A}$ such that $U_g T U_g^* = T^{\rho(g)}$, $T \in \mathfrak{A}$. Let $\{g_i\}$ be a dense sequence in G and let $\mathfrak{B}_0$ be the $*$ -algebra (with unit) generated by the U_{g_i} over the complex rationals. $\mathfrak{B}_0$ is countable. Let $\mathfrak{B}$ be the weak closure of $\mathfrak{B}_0$, a finite ring of operators. Clearly each $\rho(g_i)$ leaves $\mathfrak{B}_0$, and hence $\mathfrak{B}$, setwise invariant. The continuity of ρ then implies that $\rho(g)$ leaves $\mathfrak{B}$ setwise invariant, $g \in G$. Set $\sigma(g) = \rho(g) \mid \mathfrak{B}$. Then σ is a continuous automorphic representation on $\mathfrak{B}$ and each $\sigma(g_i)$ is inner by construction. The argument of the last paragraph shows that $\sigma(g)$ is inner, $g \in G$.

We next assert that $U_{g_j}^{\sigma(g_i)} U_{g_j}{}^* \neq I$ for some i, j. Otherwise $U_{g_i} U_{g_j} = U_{g_j} U_{g_i}$, whence $\rho(g_i g_j) = \rho(g_j g_i)$, all i, j. Taking limits, we see that ρ maps the commutator subgroup of G ($= G$ itself) into the identity, a contradiction. Let then $x \in \mathfrak{H}$ be a vector not invariant under all the $U_{g_j}^{\sigma(g_i)} U_{g_j}{}^*$. Set $\mathfrak{K}$ equal to the closure of $\mathfrak{B}x$ and P equal to the smallest central projection of $\mathfrak{B}$ such that $Px = x$. $\mathfrak{K}$ reduces every $T \in \mathfrak{B}$ and the homomorphism $T \to T \mid K$ of $\mathfrak{B}$ onto $\mathfrak{B} \mid \mathfrak{K}$ is faithful and onto on the direct summand $\mathfrak{B}P$ of $\mathfrak{B}$. Since each $\sigma(g)$ is inner, each leaves the center of $\mathfrak{B}$ elementwise invariant. Therefore σ induces by restriction an automorphic representation $\bar{\sigma}$ of G on $\mathfrak{B}P$ and hence, via the above isomorphism, on $\mathfrak{B} \mid \mathfrak{K}$. Clearly $\bar{\sigma}$ is continuous and inner. It is also non-trivial, since $x \in \mathfrak{K}$. Lastly, $\mathfrak{B}_0 x$ is a countable dence subset of $\mathfrak{K}$.

We thus see, returning to the first paragraph of the proof, that we can assume $\mathfrak{H}$ separable. Lemma 8 then implies the existence of a non-trivial continuous representation of G into the projective unitary group $\mathscr{P} = \mathscr{U} / \mathscr{C}$ of $\mathfrak{A}$. We follow now the methods of [4]. Since

$\mathfrak{H}$ is separable, $\mathfrak{A}$ is countably decomposable and hence carries a faithful normal positive (finite) central trace ω. The space $\mathfrak{A}$ is then a pre-Hilbert space in the norm $\| T \|_2^2 = \omega(T^*T)$ and the unitaries of $\mathfrak{A}$ form a topological group $\mathscr{U}_2$ in the metric $d(U, V) = \| U - V\|_2$. The identity map of $\mathscr{U}$ onto $\mathscr{U}_2$ is continuous, hence so is the identity map of $\mathscr{P}$ onto $\mathscr{P}_2 = \mathscr{U}_2 / \mathscr{C}$. Therefore G has a non-trivial continuous representation into $\mathscr{P}_2$. The metric g on $\mathscr{U}_2$ is both left and right invariant. Hence $\mathscr{P}_2$ has a metric similarly invariant. The Lemma in [4] shows that G has arbitrarily small invariant neighborhoods of the identity, an impossibility. This contradiction proves Theorem 2.

REFERENCES

1. N. Bourbaki, *Intégration*, Paris, 1952.

2. C. Chevalley, *The algebraic theory of spinors*, New York, 1954.

3. J. Dixmier, *Les algèbres d'opérateurs dans l'espace hilbertien*, Paris, 1957.

4. R. V. Kadison and I. M. Singer, *Some remarks on representations of connected groups*, Proc. Nat. Acad. Sci. U. S. A. , **38** (1952), 419–423.

5. F. Riesz, *Équations linéaraires à une infinité d'inconnues*, Paris, 1913.

6. F. Riesz and B. Sz-Nagy, *Functional analysis*, London, 1956.

7. I. E. Segal, *A non-commutative extension of abstract integration*, Ann. of Math., **57** (1953), 401–475.

8. I. E. Segal, *Tensor algebras over Hilbert spaces* II, Ann. of Math., **63** (1956), 160–175.

9. I. M. Singer, *Automorphisms of finite factors*, Amer. J. Math., **77** (1955), 117–133.

10. A. Weil, *L'intégration dans les groups topologiques et ses applications*, Paris, 1940.

THE UNIVERSITY OF CALIFORNIA AT LOS ANGELES

ISOMORPHISM ORDER FOR ABELIAN GROUPS

STEVEN BRYANT

In the theory of isometric embedding in metric spaces the following theorem is proved: Let M be a metric space every $n+3$ points of which can be mapped isometrically into Euclidean n-space, then there exists an isometry from all of M into Euclidean n-space. Because of this theorem Euclidean n-space is said to have *congruence order* $n+3$. [1].

L. M. Blumenthal has raised the question as to whether a notion analogous to that of congruence order could be developed for algebraic systems. In this paper a definition of *isomorphism order* is introduced for groups and a complete description of all Abelian groups having *finite* or *hyperfinite isomorphism order* is obtained.

First a well known definition to avoid any possible misunderstanding of the use of the concept of *rank*.

DEFINITION. A group G is said to have *rank* n if every finitely generated subgroup can be generated by n or fewer elements and n is the smallest natural number with this property.

For convenience we introduce the following definition.

DEFINITION. If k elements $g_1, g_2, \cdots, g_k$ of a group G generate a subgroup of G which is isomorphic to a subgroup of a group H, we will say that $g_1, g_2, \cdots, g_k$ are *embeddable* in H and that the subgroup generated by the g's is *embeddable* in H.

Now we are ready for the definition of isomorphism order.

DEFINITION. A group G is said to have *isomorphism order* k if and only if any group H is embeddable in G whenever every k of its elements are embeddable in G.

In the above definition k may be any cardinal number, however, in this paper k will always stand for a natural number.

If A and B are two cardinal numbers such that A is less than or equal to B then it is easy to see that if a group G has isomorphism order A then G has isomorphism order B.

Every group has some isomorphism order, since if G is a group of cardinality M then G has isomorphism order N where N is any cardinal

Presented to the Society November 17, 1956; received by the editors May 14, 1958.

number which is larger than M. Since the cardinals can be well ordered every group has a smallest isomorphism order. However, in what is to follow, if we say G has isomorphism order k we will not mean that k is the smallest isomorphism order of G unless we explicitly say so.

The following lemmas lead to a theorem describing all Abelian groups having finite isomorphism order.

LEMMA 1. *Let k be a natural number and p a fixed prime. Let G be a direct sum of k groups each of which is a cyclic group of order a power of p or a group isomorphic to $Z\ (P\,\infty)$. Then G has isomorphism order $k+1$.*

Proof. Let H be a group every $k+1$ elements of which are embeddable in G. H is primary and has rank k. From this the conclusion easily follows. (Exercise 49, [2])

LEMMA 2. *An Abelian torsion group G has isomorphism order k if and only if G is a direct sum of fewer than k subgroups of the rationals mod one.*

Proof. Let G be an Abelian torsion group having isomorphism order k. Write G as a direct sum of primary groups that is $G = \sum G_p$, where p ranges over the primes and G_p consists of all elements whose order is a power of p. Now G_p does not contain the integers mod p taken k times for, if it did, arbitrarily large groups constructed by taking direct sums of the integers mod p would (by hypothesis) be embeddable in G. From this it follows that G_p has rank less than k. Hence (exercise 49, [2]) G_p is a direct sum of fewer than k subgroups of $Z(P\,\infty)$, and therefore G is a direct sum of fewer than k subgroups of the rationals mod one by rearrangement of summands.

Conversely, let, G be a direct sum of fewer than k subgroups of the rationals mod one. Let H be a group every k elements of which are embeddable in G, so that H is torsion. Write $H = \sum H_p$ and consider H_p. Every k elements of H_p are embeddable in G_p, but by Lemma 1, G_p has isomorphism order k, hence H_p is embeddable in G_p and so H is embeddable in G.

LEMMA 3. *A torsion free Abelian group has isomorphism order k if and only if it is a vector space over the rationals of dimension less than k.*

Proof. Let G be a torsion free Abelian group having isomorphism order k. Now G does not contain the direct sum of the integers taken k times, for, if it did, the group consisting of the direct sum of the integers taken a greater number of times than the cardinality of G would have every k elements embeddable in G and hence by hypothesis would be embeddable in G, a contradiction.

Let m be the maximal number of elements of G which are independent over the integers. By what was just said m must be less than k. Any m dimensional vector space over the rationals is embeddable in G, by hypothesis. So G contains a vector space over the rationals of dimension m, call this space V. The space V is a divisible subgroup of G and hence is a direct summand so $G = A + V$. Let a be a nonzero element of A. Since m is the maximal number of independent elements of G, na is in V for some nonzero integer n, but since na is in A it is zero and therefore a is zero and so $G = V$.

Conversely, if G is a vector space over the rationals of dimension less than k and H is a group every k elements of which are embeddable in G then H is embeddable in G. To see this, observe that H can be embedded in a vector space over the rationals consisting of all couples of the form $(n, h,)$ when n is a nonzero integer and equivalence is defined in the natural way, and the dimension of this space is less than k for if not, there exist k elements of H not embeddable in G, which completes the proof.

THEOREM 1. *An Abelian group G has isomorphism order k if and only if G is the direct sum of two groups, one torsion, the other torsion free. The torsion free group is a vector space over the rationals of dimension less than k, while the torsion group can be written as a direct sum of fewer than k subgroups of the rationals mod one.*

Proof. Let G be an Abelian group having isomorphism order k. The theorem follows from the lemmas if G is torsion or torsion free. Now G contains a vector space V over the rationals of dimension n less than k where n is the maximal number of elements of G which are independent over the integers. This holds by an application of the argument of Lemma 3. Regard V as a group, then V is a direct summand of G since V is divisible. So $G = A + V$ and A is torsion, for if x is in A then mx is in V for some nonzero integer m, hence $mx = 0$. Now apply Lemma 2 to A and obtain the necessity of the theorem.

To prove the sufficiency, let G be an Abelian group such that $G = T + V$ where $T = A_1 + A_2 + \cdots + A_s$ and each A_i is a subgroup of the rationals mod one and $s < k$, and V is a vector space over the rationals of dimension less than k.

We must show that if H is an Abelian group, every k (or fewer) elements of which are embeddable in G, then H is embeddable in G.

H does not contain k elements which are independent over the integers. Hence H contains at least one subgroup H_0 such that $h \in H$ implies $rh \in H_0$ for some natural number r and such that H_0 is embeddable in G.

Let T^* be the direct sum of the rationals mod one taken s times. Let $G^* = T^* + V$. We will show that if ϕ is an isomorphism from H_0 into G^* then if $H_0 \neq H$, ϕ can be proporly extended. Then the embeddability of H in G^* can be obtained by a transfinite argument. Finally, we will see that H is embeddable in G.

So let H_0 be a subgroup of H such that $h \in H$ implies $rh \in H_0$ for some integer r and let F be an isomorphism from H_0 into G^*. If $H_0 = H$ we are done, if not, let $h \notin H_0$, and m the smallest natural number such that $mh \in H_0$.

Case 1. $m = p$, p a prime. Let $M = [z \mid pz = F(ph), z \in G^*]$. For convenience, we will refer to M as the set of all the "pth roots" of $F(ph)$, and note that M is finite, and that the number of elements in M is exactly the number of "pth roots" of 0 in G^*. Now, not every element of M is in $F(H_0)$, for if so, a glance at the inverse images will show that the inverse image of every element of M is a "pth root" of ph. But $F(ph)$ has at least as many "pth roots" in G^* as ph has in H. Hence h itself is in H_0 a contradiction.

We conclude that some element of M, call it z, is not in $F(H_0)$. Furthermore, if $0 < n < p$, then $nz \notin F(H_0)$ and hence F can be extended in the natural way.

Case 2. m not a prime, then $m = qt$ where q is a prime. Apply the argument of Case 1 to the set of all qth roots of $F(mh)$.

This shows that H is embeddable in G^*. But by Lemma 2, if T' is the torsion subgroup of H, T' is embeddable in T. Hence it is easily seen that H is actually embeddable in G, which completes the proof.

In the above theorem, nothing has been said about smallest isomorphism order. However, it is easy to see that, if G has smallest isomorphism order k then either the torsion free summand of G has rank k-1 or the torsion summand cannot be written as a direct sum of fewer than k-1 subgroups of the rationals mod 1.

The next step up in the hierarchy of isomorphism order is given by the following definition.

DEFINITION. A group G is said to have *hyperfinite isomorphism order* if, whenever every finitely generated subgroup of a group H is embeddable in G, then H is embeddable in G.

The proof of the next theorem is similar to that of Theorem 1, and rests on the fact that a torsion group has hyperfinite isomorphism order if and only if the rank of each primary subgroup is finite, while a torsion free group has hyperfinite isomorphism order if it is a finite dimensional vector space over the rationals.

THEOREM 2. *An Abelian group G has hyperfinite isomorphism order if and only if it is the direct sum of two groups, one torsion, the other torsion free. The torsion free group is a finite dimensional vector space over the rationals while the torsion summand has no primary subgroup of infinite rank.*

REMARK. If the smallest isomorphism order G has is hyperfinite, then there is no upper bound on the ranks of the primary subgroups of G.

This concludes the analysis of Abelian groups having finite or hyperfinite isomorphism order.[2] In a subsequent paper, we hope to give some results concerning Abelian groups having transfinite isomorphism order.

Also, this notion can be carried over to other systems, such as rings, a direction in which some preliminary results have been obtained.

REFERENCES

1. L. M. Blumenthal, *Theory and applications of distance geometry*, Oxford at the Clarendon Press, 1953.

2. I. Kaplansky, *Infinite Abelian groups*, University of Michigan Press, Ann Arbor, 1954.

FRESNO STATE COLLEGE
UNIVERSITY OF MISSOURI

[2] The author wishes to acknowledge his indebtedness to Joseph Zemmer for help in the initial stages of this work, and to Paul Cohen for part of the proof of Theorem 1.

MODULES WHOSE ANNIHILATORS ARE DIRECT SUMMANDS

CHARLES W. CURTIS

Introduction. Let B be a ring with an identity element, and let M be a right B-module. The set of all elements b in B such that $Mb = (0)$ is called the annihilator of M, and will be denoted by $(0 : M)$. It is a natural question to ask under what circumstances the ideal $(0 : M)$ is a direct summand of B. If B is a semi-simple ring with minimum condition, for example, then every ideal is a direct summand, and there is no problem. We shall be concerned with a ring B, not assumed to be semi-simple, which is a crossed product $\Delta(G, H, \rho)$ of a finite group G and a division ring Δ, with factor set ρ. In particular, B may be the group algebra of a finite group with coefficients in a field. The purpose of this note is to obtain necessary and sufficient conditions on the structure of the module M in order that its annihilator $(0 : M)$ be a direct summand of B.

Our interest in the problem stems chiefly from the fact that the the modules whose annihilators are direct summands turn out to be precisely the modules for which the pairing defined in § 2 of [1] is regular in the sense of [1, p. 281]. The main results of [1], given in § 5 and § 6, are based upon the assumption that the pairing is regular, and establish a connection between the structure of the module M relative to the set of B-endomorphisms of M and the structure of a certain ideal in B, called the nucleus of M, which is the uniquely determined complementary ideal to $(0 : M)$ when $(0 : M)$ is a direct summand.

2. Familiarity with crossed products and their connection with projective representations of finite groups is assumed (see [1, § 2]). In this section we recall some of the properties of a crossed product, and introduce, in a more general, and at the same time, much simpler fashion, the pairing defined in a special case by formula (7) of [1]. Let $G = \{1, s, t, \cdots\}$ be a finite group, Δ a division ring and $B = \Delta(G, H, \rho)$ a crossed product of G and Δ with correspondence $s \rightarrow \bar{s} = s^{H}$ from G to the group of automorphisms of Δ, and factor set $\{\rho_{s,t}\}$. There exist elements $\{b_1, b_s, \cdots\}$ in B in one-to-one correspondence with the elements of G, such that every element of B can be expressed uniquely in the form $\sum b_s\xi_s$, with coefficients ξ_s in Δ. The multiplication in B is determined by the equations

$$b_s b_t = b_{st}\rho_{s,t}\ ;\ \xi b_s = b_s \xi^{\bar{s}}, \qquad \xi \in \Delta . \tag{1}$$

This paper was originally accepted by the Trans. Amer. Math. Soc. Presented to the Society April 20, 1957; received by the editors of the Trans. Amer. Math. Soc., August 12, 1957.

The fact that B is an associative ring implies that the factor set $\{\rho_{s,t}\}$ satisfies the equations

$$(2)\qquad \rho_{s,tu}\rho_{t,u} = \rho_{st,u}\rho_{s,t}^{\bar{u}} ,$$

for all s, t, u in G. We shall assume that the factor set ρ is normalized so that $\rho_{1,t} = \rho_{t,1} = 1$ for all t in G; then b_1 is the identity element in B.

The additive group of B is a right vector space over Δ which we shall denote by $B^{(r)}$, if we define scalar multiplication by $\xi \in \Delta$ by means of the right multiplication $\xi_r : x \rightarrow x\xi$. Similarly the additive group of B can be regarded as a left vector space $B^{(l)}$ over Δ. The elements $b_1, b_s, \cdots$ form bases for both of these spaces. Because both spaces are finite dimensional, B satisfies both chain conditions for left and right ideals.

The mapping $\lambda : \sum b_s\xi_s \rightarrow \xi_1$ is a linear function on both vector spaces $B^{(r)}$ and $B^{(l)}$ whose kernel contains no left or right ideal different from zero. Therefore the mapping $\Lambda : \Lambda(a, b) = \lambda(ab)$ is a non-degenerate bilinear form on $B^{(l)} \times B^{(r)} \rightarrow \Delta$. Using the bilinear form Λ it is easy to verify (cf. [1, p. 279]) that B is a quasi-Frobenius ring, that is, B satisfies the minimum condition, and every right ideal in B is the right annihilator of its left annihilator, and similarly for left ideals.

A right B-module[1] M is a fortiori a right vector space over Δ since $\Delta \subset B$. For each s in G, the mapping $T_s : x \rightarrow xb_s$ is a semi-linear transformation belonging to the automorphism $\bar{s}$ in this vector space. The correspondence $s \rightarrow T_s$ defines a projective representation of G. Each transformation T_s has an inverse T_s^{-1} which is a semilinear transformation with automorphism $\bar{s}^{-1}$. Let M' be any left vector space over Δ which is paired with M to Δ by a non-degenerate bilinear form f. Let us assume also that the semi-linear transformations T_s all possess transposes T_s^* with respect to the form f, such that

$$(3)\qquad f(\psi, xT_s) = f(T_s^*\psi, x)^{\bar{s}} ,$$

for all $x \in M$, $\psi \in M'$. If we define $(\sum b_s\xi_s)\psi = \sum T_s^*(\xi_s\psi)$, then M' becomes a left B-module (see [1, p. 274]). When these conditions are satisfied, we shall call the system (M', M, f) a *pair of dual B-modules.*

LEMMA 1. *Let (M', M, f) be a pair of dual B-modules. Then the function*

$$(4)\qquad \tau_f(\psi, x) = \sum_{s \in G} f(\psi, xT_s)b_s^{-1}$$

is a non-degenerate B-bilinear function on $M' \times M \rightarrow B$ (cf. [1, Proposition 1]).

[1] We shall assume that the identity element of B acts as the identity operator on all modules we shall consider.

Proof. For any $u \in G$ we have

$$\begin{aligned} b_u^{-1}\tau_f(T_u^*\psi, x) &= \sum_{s\in G} b_u^{-1}f(T_u^*\psi, xT_s)b_s^{-1} \\ &= \sum_{s\in G} f(\psi, xT_sT_u)b_u^{-1}b_s^{-1} = \tau_f(\psi, x) \end{aligned}$$

by (1) and (3). Similarly, for all u,

$$\tau_f(\psi, xT_u)b_u^{-1} = \tau_f(\psi, x) .$$

Since the function τ_f is obviously bilinear as far as Δ is concerned, these calculations establish that for all $b \in B$,

$$b\tau_f(\psi, x) = \tau_f(b\psi, x) \text{ and } \tau_f(\psi, xb) = \tau_f(\psi, x)b .$$

The non-degeneracy of τ_f follows at once from the non-degeneracy of f.

To each right B-module M corresponds a two-sided ideal B_M in B, defined as follows. Find a left B-module M' which is paired with M to Δ by a non-degenerate bilinear form f such that (M', M, f) is a pair of dual B-modules (for example, the space M' of all linear functions on M can be used). Then by Lemma 1, the set B_M consisting of all finite sums $\sum \tau_f(\psi_i, x_i)$, $\psi_i \in M'$, $x_i \in M$, is a two-sided ideal in B. We shall call B_M the *nucleus* of M. We leave it to to the reader to verify that, as our notation indicates, B_M is independent of the choice of M' and f.

We now define a right B-module M to be a *regular module* if B_M contains an element ε such that $\varepsilon b = b\varepsilon = b$ for all $b \in B_M$. We remark that the statement that M is a regular module is equivalent to the statement, in the terminology of [1], that (M', M, τ_f) is a regular pairing (see [1, p. 281]).

3. This section contains some lemmas on regular modules. We remark first that if M_1 and M_2 are isomorphic B-modules, then $B_{M_1} = B_{M_2}$, and hence regularity is preserved under isomorphism.

LEMMA 2. *The nucleus B_M and the annihilator $(0 : M)$ of a regular module M are two-sided ideals in B generated by central idempotents, and $B = (0 : M) \oplus B_M$.*[2]

Proof. Let (M', M, f) be a pair of dual B-modules, where M is the given regular module. By Theorem 1, p. 282, of [1], we have $B_r = (B_M)_r$, and consequently $B = B_M + (0 : M)$. Let $\varepsilon = \sum \tau_f(\psi_i, x_i)$ be the identity element in B_M. Then $a \in B_M \cap (0 : M)$ implies $a = \varepsilon a = \sum \tau_f(\psi_i\, x_i a) = 0$, and the sum is direct. We have $\varepsilon' = 1 - \varepsilon \in (0 : M)$, and because B_M and $(0 : M)$ are ideals whose intersection is zero, ε and ε' are orthogonal central idempotents which generate B_M and $(0 : M)$

[2] We take this opportunity to correct an error in [1]. The assertion made in example (c) in § 11, p. 291, of [1] that $0 \neq (0 : M) \subset B_M$ for a certain regular module M is false and the assertion (c) should be deleted from [1].

respectively.

LEMMA 3. *Let M be a right B-module such that $M = M_1 \oplus M_2$, where M_1 and M_2 are submodules. Let M' be the space of all linear functions on M, paired with M to Δ by the function f defined by $f(\psi, x) = \psi(x)$, $\psi \in M', x \in M$. Then (M', M, f) is a pair of dual B-modules, Let $M_1^\perp$ and $M_2^\perp$ be the subspaces of M' which annihilate M_1 and M_2 respectively. Then $M' = M_1^\perp \oplus M_2^\perp$; the restrictions f_1 and f_2 of f to $M_2^\perp \times M_1$ and $M_1^\perp \times M_2$, respectively, are non-degenerate; and $(M_2^\perp, M_1, f_1)$, and $(M_1^\perp, M_2, f_2)$ are pairs of dual B-modules.*

Proof. The semi-linear transformations T_s all possess transposes T_s^* relative to the form f, such that formula (3) holds, and consequently (M', M, f) is a pair of dual B-modules. The sets $M_1^\perp$ and $M_2^\perp$ are subspaces of M' such that $M_1^\perp \cap M_2^\perp = (0)$. If $\psi \in M'$, then $\psi | M_1 = \psi_1$ is a linear function on M_1, which can be extended to a linear function ψ_1 in M' by setting $\psi_1 | M_2 = 0$. Similarly we define ψ_2. Then $\psi = \psi_1 + \psi_2$, and we have proved that $M' = M_2^\perp \oplus M_1^\perp$. The restrictions f_1 and f_2 defined in the statement of the lemma are clearly non-degenerate. Finally, since M_1 and M_2 are B-submodules, it follows from (3) that $T_s^*(M_i^\perp) \subseteq M_i^\perp$, $i = 1, 2$, and hence $T_s | M_i$ has the transpose $T_s^* | M_{2-i}^\perp$, $i = 1, 2$, and the proof is complete.

LEMMA 4. *Let $M = M_1 \oplus M_2$, where M_1 and M_2 are B-submodules of M. Then $B_M = B_{M_1} + B_{M_2}$.*

Proof. Let M' be the space of all linear functions on M, and define f, f_1, f_2 as in Lemma 3. Let τ_f, τ_{f_1}, τ_{f_2} be the corresponding functions defined by (4). For $x \in M_1$, $\psi \in M_2^\perp$, we have $\tau_{f_1}(\psi, x) = \tau_f(\psi, x)$ and $B_{M_1} \subseteq B_M$. Similarly $B_{M_2} \subseteq B_M$. Now let $x \in M$, and write $x = x_1 + x_2$, $x_i \in M_i$; and let $\psi \in M'$, $\psi = \psi_1 + \psi_2$, $\psi_1 \in M_2^\perp$, $\psi_2 \in M_1^\perp$. Then since M_1 and M_2 are submodules we have

$$\begin{aligned} \tau_f(\psi, x) &= \sum f(\psi_1 + \psi_2, (x_1 + x_2)b_s)b_s^{-1} \\ &= \sum f_1(\psi_1, x_1 b_s)b_s^{-1} + \sum f_2(\psi_2, x_2 b_s)b_s^{-1} \\ &= \tau_{f_1}(\psi_1, x_1) + \tau_{f_2}(\psi_2, x_2) , \end{aligned}$$

and the lemma is proved.

LEMMA 5. *Let $M = M_1 \oplus M_2$ where M_1 and M_2 are regular B-modules. Then M is a regular B-module.*

Proof. By Lemma 4, $B_M = B_{M_1} + B_{M_2}$. By Lemma 2, we have $B_{Mi} = \varepsilon_i B$ where ε_i is a central idempotent, $i = 1, 2$. Then $\varepsilon = \varepsilon_1 + \varepsilon_2 - \varepsilon_1\varepsilon_2 \in B_M$, and $b\varepsilon = \varepsilon b = b$ for all $b \in B_M$, proving our assertion.

LEMMA 6. *Let e be an idempotent in B. Then (Be, eB, Δ) is a pair of dual B-modules.*

Proof. We recall from § 2 that Δ is a non-degenerate bilinear form on $B^{(l)} \times B^{(r)} \to \Delta$. The restriction of Λ to $Be \times eB$ is also non-degenerate (see [1], p. 279). It remains to verify that for all c, d in B,

$$\Lambda(c, db_s) = \Lambda(b_s c, d)^{\bar{s}} . \tag{5}$$

For this it is sufficient to prove that if $a = \sum \xi_u b_u = \sum b_u \xi_u^{\bar{u}}$, then $\lambda(ab_s) = \lambda(b_s a)^{\bar{s}}$ for all $s \in G$. We have $\lambda(ab_s) = \xi_{s^{-1}} \rho_{s^{-1},s}$, while

$$\lambda(b_s a)^{\bar{s}} = \rho_{s,s^{-1}}^{\bar{s}} \xi_{s^{-1}}^{\overline{s^{-1}}\bar{s}} = \rho_{s,s^{-1}}^{\bar{s}} \rho_{s^{-1},s}^{-1} \xi_{s^{-1}} \rho_{s^{-1},s}$$

by formula (2) of [1], and by (2) above we have

$$\rho_{1,s} \rho_{s,s^{-1}}^{\bar{s}} = \rho_{s,1} \rho_{s^{-1},s} ,$$

and the formula (5) is proved.

4. Now we shall formulate and prove our main result. Because B satisfies the minimum condition, $B = B_1 \oplus \cdots \oplus B_r$, where the B_i are uniquely determined indecomposable two-sided ideals, called the block ideals[3] of B. If we write $1 = \varepsilon_1 + \cdots + \varepsilon_r$, $\varepsilon_i \in B_i$, then the ε_i are mutually orthogonal idempotents belonging to the center of B, and ε_i is the identity element in the block ideal to which it belongs. For any right B-module M, $M\varepsilon_i$ is a submodule of M, and M is the direct sum of the modules $M\varepsilon_i$. These submodules are called the *block components* of M; the block component $M\varepsilon_i$ can also be described as the set of elements of M which are left fixed by ε_i. The block components of $(B, +)$, where $(B, +)$ is viewed as a right B-module in the obvious way, are the block ideals $B\varepsilon_i$. Each block component $B\varepsilon_i$ of B can be expressed as a direct sum of the indecomposable right ideals $e_k B$, $e_k^2 = e_k$, which belong to the block. It is known that two indecomposable right ideals eB and $e'B$ belonging to distinct blocks have no isomorphic composition factors. The direct sum of a full set of non-isomorphic indecomposable right ideals $e_k B$ belonging to the ith block component $B\varepsilon_i$ of B, or any right B-module isomorphic to this module, is called a *reduced block component* of B.

Our theorem is stated as follows.

THEOREM. *Let M be a right B-module with annihilator* $(0 : M)$. *The following statements are equivalent.*

(A) $(0 : M)$ *is a direct summand of B*;

(B) *every non-zero block component $M\varepsilon_i$ of M contains the* ith *reduced block component of B as a direct summand*;

(C) *M is a regular module.*

Proof. The implication (C) $\to$ (A) is the content of Lemma 2. We prove next that (A) $\to$ (B). Let B' be a two sided ideal in B such that

[3] For the concepts of block ideals and block components see [3], and the references given there.

$B = B' \oplus (0 : M)$. By the uniqueness of the decomposition of B into block ideals, B' is a direct sum of certain of the block ideals $B\varepsilon_i$. Let $M\varepsilon_i$ be a non-zero block component of M; then $B\varepsilon_i \subseteqq B'$, and $M\varepsilon_i$ is a faithful $B\varepsilon_i$ module. Let eB be an indecomposable right ideal belonging to the ith block. By Proposition 4 of [1], eB contains a unique minimal right ideal $N \neq (0)$. There exists an element $x \in M$ such that $xN \neq (0)$. It follows that $u \rightarrow xu$ is a B-isomorphism of eB onto the submodule $P = xeB$ of $M\varepsilon_i$. We shall prove that there exists a submodule Q of $M\varepsilon_i$ such that $M\varepsilon_i = Q \oplus P$. Let M' be the set of all linear functions on $M\varepsilon_i$, paired with $M\varepsilon_i$ to Δ by the non-degenerate bilinear form f, so that $(M', M\varepsilon_i, f)$ is a pair of dual B-modules. Let $P^\perp$ be the submodule of M' consisting of all elements $\psi \in M'$ such that $f(\psi, P) = (0)$. Then $(M'/P^\perp, P, \bar{f})$ is a pair of dual B-modules, where $\bar{f}$ is the induced mapping on $M'/P^\perp \times P$. On the other hand, by Lemma 6, (Be, eB, Δ) is a pair of dual B-modules. Using the fact that eB is a finite dimensional space, it is easily verified that Be and $M'/P^\perp$ are isomorphic left B-modules. By Theorem 1 of [2], Be is an (M_0)-module, and consequently there exists a B-submodule Q' of M' such that $M' = P^\perp \oplus Q'$. Let $Q = \{x \mid x \in M\varepsilon_i,\ f(Q', x) = (0)\}$. Then Q is a submodule such that $P \cap Q = (0)$. Moreover

$$M = (P^\perp \cap Q')^\perp = P^{\perp\perp} + (Q')^\perp = P + Q \, ,$$

since P is finite dimensional and $Q = (Q')^\perp$.

The proof that $M\varepsilon_i$ contains the reduced block component of $B\varepsilon_i$ as a direct summand is now proved by induction. Let $M\varepsilon_i = R \oplus S$, where R is isomorphic to a direct sum of a finite number of non-isomorphic indecomposable right ideals belonging to the ith block, and let eB be an indecomposable right ideal in $B\varepsilon_i$ not isomorphic to any of the direct summands of R. Let N be the unique minimal subideal of eB. If $RN \neq (0)$, then by the previous argument R contains a direct summand isomorphic to eB, which contradicts the Krull-Schmidt theorem. Thus $RN = (0)$, and $SN \neq (0)$, so that S contains a direct summand isomorphic to eB. This completes the proof of the induction step, and the implication (A) $\rightarrow$ (B) is established.

Finally we prove that (B) $\rightarrow$ (C). By Lemma 5, it is sufficient to prove that each block component $M\varepsilon_i$ of M is a regular module, and for this it is sufficient to show that $\varepsilon_i \in B_{Mi}$ whenever $M\varepsilon_i \neq (0)$. Let us consider a non-zero component $M\varepsilon_i$. Let $e_1B, \cdots, e_sB$ be a full set of non-isomorphic indecomposable right ideals belonging to the i^{th} block. For each j, $1 \leqq j \leqq s$, there exists a B-direct summand P_j of $M\varepsilon_i$ such that $P_j \cong e_jB$. By Lemma 4, $B_{P_j} = B_{e_jB} \subseteqq B_{M\varepsilon i}$. We prove that $e_j \in B_{e_jB}$. By Lemma 6, (Be_j, e_jB, Δ) is a pair of dual B-modules. We assert that

$$e_j = \tau_\Delta(e_j, e_j) \, . \qquad (6)$$

In fact, $\tau_\Lambda(e_j, e_j) = \sum \Lambda(e_j, e_jb_s)b_s^{-1}$, and if $e_j = \sum \xi_u b_u$, then

$$\Lambda(e_j, e_jb_s) = \lambda(e_jb_s) = \xi_{s^{-1}}\rho_{s^{-1},s}$$

while from $b_{s^{-1}}b_s = b_1\rho_{s^{-1},s}$ we have $b_s^{-1} = \rho_{s^{-1},s}^{-1}b_{s^{-1}}$. From these remarks (6) follows.

We have shown that $e_j \in B_{M\varepsilon i}$. Since ε_i is a sum of idempotents e such that eB is isomorphic to one of the ideals e_jB, $1 \leqq j \leqq s$, we have $\varepsilon_i \in B_{M\varepsilon i}$, and $M\varepsilon_i$ is a regular module. This completes the proof of the theorem.

References

1. C. W. Curtis, *Commuting rings of endomorphisms*, Can. J. Math., **8** (1956), 271-292.
2. H. Nagao, and T. Nakayama, *On the structure of* (M_0) *and* (M_u) *modules*, Math. Zeit., **59** (1953), 164-170.
3. C. Nesbitt, and R. Thrall, *Some ring theorems with applications to modular representations*, Ann. of Math,. **47** (1946), 551-567.

UNIVERSITY OF WISCONSIN

ON THE RADICAL OF A GROUP ALGEBRA

W. E. DESKINS

A basic result in the study of group algebras and characters states that the group algebra $\mathfrak{A}(\mathscr{G})$ of a finite group $\mathscr{G}$ over the field $\mathfrak{F}$ of characteristic $p \neq 0$ has a nonzero radical $\mathfrak{R}$ if and only if p is a divisor of $o(\mathscr{G})$, the order of $\mathscr{G}$. This suggests that $\mathfrak{R}$ is related in some manner to the Sylow p-groups of $\mathscr{G}$ and that it may be possible to define $\mathfrak{R}$ in terms of these subgroups. In [6] Jennings showed that if $o(\mathscr{G}) = p^a$, then $\mathfrak{R}$ is of dimension $p^a - 1$ and has as a basis the set of elements $P_i - 1$. As a generalization of this define $\mathfrak{R}'$ to be the intersection of all the left ideals of $\mathfrak{A}(\mathscr{G})$ generated by the radicals of the group algebras of the Sylow p-groups of $\mathscr{G}$. Then $\mathfrak{R}'$ is a nilpotent ideal of $\mathfrak{A}(\mathscr{G})$ (cf. [2]), and Lombardo-Radici has shown [8] that $\mathfrak{R}' = \mathfrak{R}$ provided $\mathscr{G}$ has a unique Sylow p-group or $o(\mathscr{G}) = pq$ where q is also a prime. Also, in [9] he demonstrated that if $\mathscr{G}$ is the simple group of order 60 and if $p = 2$ or 3 then $\mathfrak{R}'$ is a proper subideal of $\mathfrak{R}$. In this paper it will be shown that $\mathfrak{R}' = \mathfrak{R}$ if one of the following conditions is satisfied:

(A) $\mathscr{G}$ is homomorphic with a Sylow p-group of $\mathscr{G}$.

(B) $\mathscr{G}$ is a super-solvable group.

(C) $\mathscr{G}$ is a solvable group with $(o(\mathscr{G}), p^2) = p$.

In the last section of the paper an application to a related problem is made. If $\mathscr{G}$ contains an invariant p-group then $\mathfrak{A}(\mathscr{G})$ is bound to its radical $\mathfrak{R}$ (i.e., if a in $\mathfrak{A}(\mathscr{G})$ is an element such that $a\mathfrak{R} = \mathfrak{R}a = 0$, then a is in $\mathfrak{R}$). This raises the question: If $\mathfrak{A}(\mathscr{G})$ is bound to its radical $\mathfrak{R}$, does $\mathscr{G}$ contain an invariant p-group? This is equivalent to the question: Does $\mathscr{G}$ contain an invariant p-group if $\mathscr{G}$ possesses no irreducible representation of highest kind? (An irreducible representation of highest kind is one whose dimension is divisible by the highest power of p which divides $o(\mathscr{G})$.) It is shown that if $\mathscr{G}$ is a group such that $\mathfrak{R}' = R$ and if the Sylow p-groups of $\mathscr{G}$ are cyclic, then the above question is answered affirmatively. Also an example is given where the answer is negative.

1. Type A. Let $\mathscr{G}$ be a group of order of order $g = hp^a$, $(h, p) = 1$, with a normal subgroup $\mathscr{H}$ of order h. And let $\mathfrak{F}$ be an algebraically closed field of characteristic p. (The requirement that $\mathfrak{F}$ be algebraically closed is only a convenience since the dimension of $\mathfrak{R}'$ is

Received by the editors June 19, 1957 and, in revised form, April 9, 1958. Presented to the American Mathematical Society, April 21, 1957.

unaffected by any extension of the ground field.)

THEOREM 1. *The radical* $\mathfrak{R}$ *of the group algebra* $\mathfrak{A}(\mathscr{G})$ *of the group* $\mathscr{G}$ *over the field* $\mathfrak{F}$ *equals* $\mathfrak{R}'$, *the intersection of all the left ideals of* $\mathfrak{A}(\mathscr{G})$ *generated by the radicals of the group algebras of the Sylow p-groups of* $\mathscr{G}$.

Let $\mathscr{P}$ be a Sylow p-group of $\mathscr{G}$: then $\mathscr{G}/\mathscr{H}$ is isomorphic with $\mathscr{P}$ and $\mathscr{G}$ is an extension of $\mathscr{H}$ by $\mathscr{P}$. Now $\mathfrak{A}(\mathscr{P})$, the group algebra of $\mathscr{P}$ over $\mathfrak{F}$, has the radical $\mathfrak{N}$ which is of dimension $p^a - 1$ over $\mathfrak{F}$ and has as a basis the differences $P_i - 1$, all $P_i \in P$. Form $\mathfrak{M}$, the left ideal of $\mathfrak{A}(\mathscr{G})$ generated by $\mathfrak{N}$. The ideal $\mathfrak{M}$ is of dimension $h(p^a - 1)$ over $\mathfrak{F}$, and we propose to show that $\mathfrak{R}$, the radical of $\mathfrak{A}(\mathscr{G})$, is contained in $\mathfrak{M}$.

Now $\mathfrak{A}(\mathscr{H})$, the group algebra of $\mathscr{H}$ over $\mathfrak{F}$, is expressible as $\mathfrak{B}_1 \oplus \cdots \oplus \mathfrak{B}_n$ where $\mathfrak{B}_i$ is a simple ideal of $\mathfrak{A}(\mathscr{H})$. Let $\mathfrak{B}$ be one of these, and let $\mathscr{P}'$ be the subgroup of $\mathscr{P}$ consisting of elements P_i such that $P_i\mathfrak{B}P_i^{-1} = \mathfrak{B}$, with $o(\mathscr{P}') = r = p^c$, $o \leqq c \leqq a$. The elements H of $\mathscr{H}$ are represented by $\bar{H}$ in $\mathfrak{B}$ and the $\bar{H}$ form a group $\bar{H}$ homomorphic with $\mathscr{H}$. Furthermore the elements of $\mathfrak{B}$ can be expressed linearly in terms of the elements of $\bar{\mathscr{H}}$.

If $P \in \mathscr{P}'$, then P corresponds to an automorphism of $\mathfrak{B}$ since $P\mathfrak{B}P^{-1} = \mathfrak{B}$, and since $\mathfrak{B}$ is central simple this automorphism is an inner automorphism of $\mathfrak{B}$. Thus P corresponds to a sum of elements of $\bar{\mathscr{H}}$ and so leaves the conjugate classes of $\bar{\mathscr{H}}$ invariant since these classes commute with the individual elements of $\bar{\mathscr{H}}$. Basically, therefore, we are dealing with an extension $\bar{\mathscr{G}}$ of $\bar{\mathscr{H}}$ by a p-group $\mathscr{P}'$ in which each element of $\mathscr{P}'$ induces an automorphism A of $\bar{\mathscr{H}}$ which leaves the conjugate classes invariant. Since the order of $\bar{\mathscr{H}}$ is prime to p it is well-known [**11**, p. 123] that A is an inner automorphism of $\bar{\mathscr{H}}$. Now a result due to M. Hall [**4**, Theorem 6.1] implies that $\bar{\mathscr{G}}$ is a direct product of $\mathscr{P}'$ and $\bar{\mathscr{H}}$, and this leads to the conclusion that the elements of $\mathscr{P}'$ commute elementwise with $\mathfrak{B}$. If $\mathfrak{Q} = \sum_{P_i \in \mathscr{P}'} P_i\mathfrak{B}$, then the radical $\mathfrak{Q}'$ of $\mathfrak{Q}$ equals $\mathfrak{B}$ times the radical of $\mathfrak{A}(\mathscr{P}_{,})$, and therefore $\mathfrak{Q}'$ is contained in $\mathfrak{M}$.

If $t = p^{a-c}$ is the index of $\mathscr{P}'$ in $\mathscr{P}$, then there are t distinct ideals $\mathfrak{B}_i$ in the decomposition of $\mathfrak{A}(\mathscr{H})$ which form a set of transitivity $\mathbf{T}$ for $\mathscr{P}$, with $\mathfrak{B}_1 = \mathfrak{B}$. That is, $P_i\mathfrak{B}_jP_i^{-1} \in \mathbf{T}$ if $\mathfrak{B}_j \in \mathbf{T}$ and $P_i \in \mathscr{P}$, and furthermore, if $\mathfrak{B}_i, \mathfrak{B}_j \in \mathbf{T}$, then there is a $P_k \in \mathscr{P}$ such that $\mathfrak{B}_i = P_k\mathfrak{B}_jP_k^{-1}$. Then the algebra $\mathfrak{T} = \sum P_i\mathfrak{B}_j$, all $P_i \in \mathscr{P}$ and $\mathfrak{B}_j \in \mathbf{T}$, is an ideal of $\mathfrak{A}(\mathscr{G})$, and we assert that its radical is contained in $\mathfrak{M}$. To see this consider the coset expansion of $\mathscr{P}$ relative to $\mathscr{P}'$, $\mathscr{P} = \sum S_i \mathscr{P}' = \sum \mathscr{P}' S_i$. Then clearly the algebra $\mathfrak{T}' = \sum_{i,j}^t S_i \mathfrak{Q}' S_j$ is a nilpotent ideal of $\mathfrak{T}$, while the transitivity of $\mathbf{T}$ implies that $\mathfrak{T} - \mathfrak{T}'$ is a simple algebra. Thus $\mathfrak{T}'$ is the radical of $\mathfrak{T}$ and obviously is contained in $\mathfrak{M}$.

As the choice of $\mathfrak{B}$ was arbitrary in the decomposition of $\mathfrak{A}(\mathscr{H})$, clearly the process above leads to the conclusion that $\mathfrak{R}$ is contained in $\mathfrak{M}$. Since the choice of $\mathscr{P}$ was arbitrary this enables us to conclude that $\mathfrak{R}' \supseteqq \mathfrak{R}$. However $\mathfrak{R}'$ is known to be nilpotent (cf [2]), hence $\mathfrak{R}' = \mathfrak{R}$.

2. Type B. A group $\mathscr{G}$ is defined to be *super-solvable* if it possesses a sequence of subgroups $\mathscr{G}_0 = \mathscr{G} \supset \mathscr{G}_1 \supset \cdots \supset \mathscr{G}_s = 1$ such that $\mathscr{G}_i$ is normal in $\mathscr{G}$ and $\mathscr{G}_i / \mathscr{G}_{i+1}$ is cyclic. If in addition each $\mathscr{G}_i / \mathscr{G}_{i+1}$ is contained in the center of $\mathscr{G} / \mathscr{G}_{i+1}$, then $\mathscr{G}$ is called a *nilpotent* group. A basic result concerning nilpotent groups states that a nilpotent group is a direct product of its Sylow groups. And a principal theorem on super-solvable groups states that a super-solvable group is an extension of a nilpotent group by a nilpotent group. (For these results see Kurosch [7, pp. 216 and 228])

THEOREM 2. *The radical $\mathfrak{R}$ of the group algebra $\mathfrak{A}(\mathscr{G})$ of a super-solvable group $\mathscr{G}$ over the field $\mathfrak{F}$ equals $\mathfrak{R}'$.*

By the theorems quoted above $\mathscr{G}$ contains a normal nilpotent subgroup $\mathscr{G}_1$ such that $\mathscr{G} / \mathscr{G}_1$ is nilpotent while $\mathscr{G}_1$ has a normal Sylow p-group $\mathscr{P}_1$. Evidently $\mathscr{P}_1$ is normal in $\mathscr{G}$ since $\mathscr{G}_1$ is a direct product of its Sylow groups. Then the radical of $\mathfrak{A}(\mathscr{P}_1)$ generates a nilpotent ideal $\mathfrak{R}_1$ of $\mathfrak{A}(\mathscr{G})$ and $\mathfrak{A}(\mathscr{G}) - \mathfrak{R}_1$ is isomorphic with the group algebra $\mathfrak{A}(\mathscr{G} / \mathscr{P}_1)$ of $\mathscr{G} / \mathscr{P}_1$. Now the group $\mathscr{G} / \mathscr{P}_1$ is a group of Type A which was discussed in the preceding section. So if $\mathfrak{J}$ is a left ideal of $\mathfrak{A}(\mathscr{G})$ generated by the radical of the group algebra of $\mathscr{P}$, a Sylow p-group of $\mathscr{G}$, then $\mathfrak{A}(\mathscr{G}) - \mathfrak{J}$ is a completely reducible left $\mathfrak{A}(\mathscr{G})$-module since $\mathscr{P} / \mathscr{P}_1$ is a Sylow p-group of $\mathscr{G} / \mathscr{P}_1$. Hence $\mathfrak{R} = \mathfrak{R}'$.

3. Type C. Let $\mathscr{G}$ be a solvable group whose order is divisible by p to the first power only. Then $\mathscr{G}$ possesses a sequence of subgroups $\mathscr{G}_0 = \mathscr{G} \supset \mathscr{G}_1 \supset \cdots \supset \mathscr{G}_n = 1$ such that $\mathscr{G}_{i+1}$ is normal in $\mathscr{G}_i$ and $\mathscr{G}_i / \mathscr{G}_{i+1}$ is a group of order q where q is a prime.

THEOREM 3. *The radical $\mathfrak{R}$ of the group algebra $\mathfrak{A}(\mathscr{G})$ of the group $\mathscr{G}$ over the field $\mathfrak{F}$ equals $\mathfrak{R}'$.*

The proof will be by induction on n, the length of the series defined above. If $n = 1$ the theorem is trivally true; so assume the result to be true for groups of length less than n. Now consider $\mathscr{G}_1$, which is of length $n-1$. If $\mathscr{G}/\mathscr{G}_1$ is of order p, then the order of $\mathscr{G}_1$ is prime to p and the result follows by Theorem 1. So we shall restrict our attention to the case where $\mathscr{G}/\mathscr{G}_1$ is of order q, $(p, q) = 1$.

Now by a theorem due to P. Hall [5] $\mathscr{G}$ contains a group $\mathscr{H}$ of order t, where $pt = g$, the order of $\mathscr{G}$. If $\mathscr{P}$ is a Sylow p-group $\mathscr{G}$ of form $\mathfrak{J}$, the left ideal of $\mathfrak{A}(\mathscr{G})$ generated by the radical of $\mathfrak{A}(\mathscr{P})$. Then $\mathfrak{A}(\mathscr{G}) - \mathfrak{J} = \mathfrak{Q}$ is a left $\mathscr{G}$-module representable by $\mathfrak{A}(\mathscr{H})$ and is a completely reducible $\mathfrak{A}(\mathscr{G}_1)$-module. For $\mathfrak{R}_1$, the radical of $\mathfrak{A}(\mathscr{G}_1)$, is such that $\mathfrak{R}_1\mathfrak{A}(\mathscr{G})$ is contained in $\mathfrak{J}$ and so $\mathfrak{R}_1\mathfrak{Q} = 0$. So let $\mathfrak{Q}_1$ be an irreducible left $\mathscr{G}$-submodule of $\mathfrak{Q}$. Then $\mathfrak{Q}$ may be written $\mathfrak{Q} = \mathfrak{Q}_1 + \mathfrak{Q}_2$ where $\mathfrak{Q}_2$ is a left $\mathfrak{A}(\mathscr{G}_1)$-module and $\mathfrak{Q}_1 \cap \mathfrak{Q}_2 = 0$. Therefore a projection T of $\mathfrak{Q}$ onto $\mathfrak{Q}_2$ exists such that T annihilates the elements of $\mathfrak{Q}_1$ and is the identity operator on $\mathfrak{Q}_2$ and such that T commutes with (the representations of) the elements of $\mathfrak{A}(\mathscr{G}_1)$. Now form the projection $T' = t^{-1} \sum H_i T H_i^{-1}$, summed over the t elements of $\mathscr{H}$. Then T' commutes with all the elements of $\mathscr{G}$ and hence the submodule $\mathfrak{Q}_1' = T'\mathfrak{Q}$ of $\mathfrak{Q}$ is a left $\mathfrak{A}(\mathscr{G})$-module. Furthermore $\mathfrak{Q} = \mathfrak{Q}_1 + \mathfrak{Q}_1'$ where $\mathfrak{Q}_1 \cap \mathfrak{Q}_1' = 0$. Thus $\mathfrak{Q}$ is a completely reducible left $\mathfrak{A}(\mathscr{G})$-module and so $\mathfrak{J}$ contains the radical of $\mathfrak{A}(\mathscr{G})$. This proves Theorem 3.

4. A related problem. An algebra having the property that only elements of the radical can be both left and right annihilators of the radical has been termed a *bound algebra* by M. Hall [3].

THEOREM 4. *If the group $\mathscr{G}$ contains an invariant p-subgroup $\mathscr{P}$, then the group algebra $\mathfrak{A}(\mathscr{G})$ of $\mathscr{G}$ over a field of characteristic p is a bound algebra.*

If $\mathscr{P}$ is of order $p^d = x$ and of index y, then the radical of $\mathfrak{A}(\mathscr{P})$ generates a nilpotent ideal $\mathfrak{J}$ of $\mathfrak{A}(\mathscr{G})$ of dimension $y(x-1)$. Now the element $P_1 + \cdots + P_x$, where P_i is in $\mathscr{P}$, annihilates $\mathfrak{J}$ and is also in the center of $\mathfrak{A}(\mathscr{G})$. Hence it generates an ideal J of order y which is contained in $\mathfrak{J}$ and $\mathfrak{J}J = J\mathfrak{J} = 0$. Since $\mathfrak{A}(\mathscr{G})$ is a Frobenius algebra, a result due to Nakayama [10] states that the set of all right annihilators of $\mathfrak{J}$ in $\mathfrak{A}(\mathscr{G})$ forms an ideal of dimension y. Hence $\mathfrak{J}$ contains all of the right annihilators of $\mathfrak{J}$. Since $\mathfrak{J} \subseteqq \mathfrak{R}$, $\mathfrak{J}$ contains the right annihilators of $\mathfrak{R}$, and so $\mathfrak{A}(\mathscr{G})$ is bound to $\mathfrak{R}$.

This raises the question: If $\mathfrak{A}(\mathscr{G})$ is bound to its radical $\mathfrak{R} \neq 0$, does $\mathscr{G}$ contain an invariant p-subgroup? A partial answer is provided by

THEOREM 5. *If the Sylow p-groups of $\mathscr{G}$ are cyclic and if the* *radical* $\mathfrak{R}$ *of* $\mathfrak{A}(\mathscr{G})$ *equals* $\mathfrak{R}'$ *then* $\mathscr{G}$ *contains an invariant p-subgroup if* $\mathfrak{A}(\mathscr{G})$ *is bound to* $\mathfrak{R}$.

Let $\mathscr{P}_1$ and $\mathscr{P}_2$ be two Sylow p-groups of $\mathscr{G}$ and let $\mathfrak{J}_1$ and $\mathfrak{J}_2$ be the two left ideals of $\mathfrak{A}(\mathscr{G})$ generated by the radicals of $\mathfrak{A}(\mathscr{P}_1)$ and $\mathfrak{A}(\mathscr{P}_2)$ respectively. Denote by $r(\mathfrak{J}_1)$ and $r(\mathfrak{J}_2)$ the right ideals of $\mathfrak{A}(\mathscr{G})$ consisting of all elements which annihilate $\mathfrak{J}_1$ and $\mathfrak{J}_2$, respectively, on the right. Then since $\mathfrak{R} \subseteqq \bigcap \mathfrak{J}_i$ and since $r(\mathfrak{R}) \subseteqq \mathfrak{R}$ it follows readily that $r(\mathfrak{J}_1)$ and $r(\mathfrak{J}_2)$ are contained in $\mathfrak{R} = \mathfrak{R}'$. In particular, the sum S of the elements of $\mathscr{P}_1$ is contained in $\mathfrak{J}_2$. Now the only elements of $\mathfrak{J}_2$ which involve 1, the identity of $\mathscr{G}$, also involve other elements of $\mathscr{P}_2$, so that the belonging of S to $\mathfrak{J}_2$ implies that $\mathscr{P}_1 \cap \mathscr{P}_2$ is a group containing more than one element. Then, since the $\mathscr{P}_i$ are all cyclic, it follows readily that the p-subgroup $\mathscr{P}_1 \cap \mathscr{P}_2$ is normal in $\mathscr{G}$.

Now $\mathfrak{A}(\mathscr{G})$ is bound to $\mathfrak{R}$ if and only if $\mathscr{G}$ possesses no representation of highest kind (see [1]). If $\mathscr{G}$ is S_5, the symmetric group of order 120 and if $p = 2$, then the table of ordinary characters readily demonstrates that $\mathscr{G}$ has no representation of highest kind. Yet S_5 has no invariant 2-subgroup. It may be noteworthy that this example is related to the one given by Lombardo-Radici [9] to show that $\mathfrak{R}$ is not always equal to $\mathfrak{R}'$.

REFERENCES

1. R. Brauer and C. Nesbitt, *On the modular characters of groups*, Ann. Math. **42** (1941), 556-590.
2. W. E. Deskins, *Finite Abelian groups with isomorphic group algebras*, Duke Math. J. **23** (1956), 35-40.
3. M. Hall, *The position of the radical in an algebra*, Trans. Amer. Math. Soc. **48** (1940), 391-404.
4. ———, *Group rings and extensions*, I, Ann. Math. **39** (1938), 220-234.
5. P. Hall, *A contribution to the theory of groups of prime power orders*, Proc. Lond. Math. Soc. **36** (1933-1934), 29-95.
6. S. A. Jennings, *The structure of the group ring of a p-group over a modular field*, Trans. Amer. Math. Soc. **50** (1941), 175-185.
7. A. G. Kurosch, *The Theory of Groups*, vol. 2, New York, Chelsea, 1956.
8. L. Lombardo-Radici, *Intorno alle algebre legate ai gruppi di ordine finito*, II, Rend. Sem. Mat. Roma **3** (1939), 239-256.
9. ———, *Sulle condizioni di appnrtenenza al radicale per gli elementi di un algebra legata a un gruppo finito*, Atti Accad. Naz. Lincei. Rend. Cl. Sci. Fis. Mat. Nat. (8) **4** (1948), 53-54.
10. T. Nakayama, *On Frobeniusean algebras*, I, Ann. Math. **40** (1939), 611-633.
11. A. Speiser, *Die Theorie der Gruppen*, New York, Dover.

MICHIGAN STATE UNIVERSITY

EQUIVALENCE AND PERPENDICULARITY OF GAUSSIAN PROCESSES

JACOB FELDMAN

1. Introduction. In [6] S. Kakutani showed that if one has equivalent probability measures μ_i and ν_i on the σ-field $\mathscr{S}_i$ of subsets of a set Ω_i, $i = 1, 2, \cdots$, and if μ and ν denote respectively the infinite product measures $\bigotimes_{i=1}^{\infty}\mu_i$ and $\bigotimes_{i=1}^{\infty}\nu_i$ on the infinite product σ-ring generated on the infinite product set Ω, then μ and ν are either equivalent or perpendicular, and he obtained necessary and sufficient conditions for equivalence to occur. The theorem here shown may be regarded as a generalization of a case of the Kakutani theorem.

Similar dichotomies have revealed themselves in the study of Gaussian stochastic processes. C. Cameron and W. T. Martin proved in [2] that if one considers the measures induced on path space by a Wiener process on the unit interval, then if the variances of the processes are different the measures are perpendicular. This sort of result was generalized by U. Grenander, starting from the viewpoint of statistical estimation, and utilizing a Karhunen representation for the processes involved. A wider sufficient condition for perpendicularity of the measures induced on path space by continuous Gaussian processes on the unit interval was obtained by G. Baxter in [1]. Cameron and Martin also examined the effect on the induced measure of taking certain types of affine transformations of a Wiener process (see [3], [4]). I. E. Segal extended their results in [8], and made the situation more transparent by use of his notion of "weak distributions", and in a large class of cases got conditions for equivalence.

In the present note it is shown that the equivalence-or-perpendicularity dichotomy holds in general for pairs of measures induced by Gaussian stochastic processes, and Segal's necesssary and sufficient conditions for equivalence are extended to cover the case of nonzero mean. It has been pointed out to the author by C. Stein that one could also give a proof, in the case of zero mean, by use of the techniques of statistical testing of hypotheses.

2. Several lemmas. All Hilbert spaces mentioned will be over the reals.

Definition 1. An operator T from Hilbert space $\mathbf{H}$ to Hilbert space $\mathbf{K}$ will be called on *equivalence* operator if

Received April 7, 1958.

(1) T is one-to-one onto, bounded, and has a bounded inverse.
(2) $\sqrt{T^*T} = I + H$, where H is Hilbert-Schmidt.

LEMMA 1. *If A is a self-adjoint bounded invertible operator on* **H** *then the following statements are equivalent*:
(a) *$A—I$ is Hilbert-Schmidt*;
(b) *$(A—I)^2$ is Hilbert-Schmidt*;
(c) *$A^{-1}—I$ is Hilbert-Schmidt.*
If A, B satisfy (a), *then so does ABA.*

Proof. The first part is clear from consideration of the eigenvalues of the operators. For the second part: write $A = I + K, B = I + H$. Then $ABA = (I + K)^2 + (I + K)H(I + K)$, and since the sum and product of two Hilbert-Schmidt operators is Hilbert-Schmidt, $ABA—I$ is Hilbert-Schmidt.

DEFINITION 1. An operator T from Hilbert space **H** to Hilbert space **K** will be called an *equivalence* operator if
(1) T is one-to-one onto, bounded, and has a bounded inverse;
(2) $\sqrt{T^*T} - I$ is Hilbert-Schmidt.

LEMMA 2. *Products, conjugates, and inverses of equivalence operators are again equivalence operators.*

Proof. That they are one-to-one onto, and bounded, is clear (in the case of the conjugate operator, use the fact that the nullspace of T^* is the orthogonal complement of the range of T).

Let T be an equivalence operator from **H** to **K**. Let $Q=\sqrt{T^*T}$. Then $V = TQ^{-1}$ is an isometry from **H** onto **K**, and $T = VQ$. Thus $T^{-1} = Q^{-1}V^*$, and $(T^{-1})^*(T^{-1}) = VQ^{-2}V^*$. Since Q is the type of operator occurring in Lemma 1, and $(T^{-1})^*T^{-1}$ is a unitary transform of Q^{-2}, we get the result. Similarly, $(T^*)^*T^* = TT^* = VQ^2V^*$. Finally, let S be an equivalence operator from **K** to **L**, and let $P = \sqrt{S^*S}$, $U = SP^{-1}$. Then

$$(ST)^*(ST) = (VQUP)^*(VQUP) = PU^*QV^*VQUP = PU^*QUP,$$

and again Lemma 1 tells us that $\sqrt{(ST)^*(ST)}$ is of the desired form.

DEFINITION 2. A function x on a measure space with measure μ of total mass 1 is called *Gaussian* if either
(1) x is almost everywhere a constant, γ.
or
(2) there are numbers $\sigma > 0$ and γ (depending on x) such that

$$\mu\{\omega \mid x(\omega) \leqq \lambda\} = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\lambda} exp\left\{\frac{1}{2}\left(\frac{t-\gamma}{\sigma}\right)^2\right\} dt \, .$$

Case 1 may be thought of as Case 2 with $\sigma = 0$. Then in either case we have

$$\int x d\mu = \gamma, \int (x-\gamma)^2 d\mu = \sigma^2$$

(the "mean" and "variance" of x).

LEMMA 3. *Let $\mathscr{S}_1 \subset \mathscr{S}_2 \subset \cdots$ be σ-fields of subsets of Ω, $\mathscr{S}$ the smallest σ-field containing their union. Let μ, ν be probability measures in $\mathscr{S}$ such that $\mu_i = \mu \mid \mathscr{S}_i$ is equivalent to $\nu_i = \nu \mid \mathscr{S}_i$. Let Λ, Λ_μ, Λ_ν be sets in $\mathscr{S}$ forming a Hahn decomposition of Ω; that is, μ is equivalent to ν when both are cut down to subsets of Λ, and $\mu(\Lambda_\nu) = \nu(\Lambda_\mu) = 0$. Then $d\mu_i/d\nu_i$ converges almost everywhere with respect to $\mu + \nu$ to $d\mu/d\nu$, if one makes the convention $d\mu/d\nu = 0$ on Λ_ν and $+\infty$ on Λ_μ.*

Proof. If $\Lambda_i \in \mathscr{S}_i$, then

$$\int_{\Lambda_i} \frac{d\mu_i}{d(\mu+\nu)_i} d(\mu+\nu) = \mu_i(\Lambda_i) = \mu(\Lambda) \, ,$$

so that $d\mu_i/d(\mu+\nu)_i$ is the conditional expectation of $d\mu_i/d(\mu+\nu)$ with respect to $\mathscr{S}_i$ and the measure $\mu+\nu$. Of course $\mu+\nu$ has total mass 2, but this is inessential; one can always normalize things if so inclined. The Martingale convergence theorem then tells us that

$$\frac{d\mu_i}{d(\mu+\nu)_i} \to \frac{d\mu}{d(\mu+\nu)}$$

almost everywhere with respect to $\mu+\nu$. Similarly

$$\frac{d\nu_i}{d(\mu+\nu)_i} \to \frac{d\nu}{d(\mu+\nu)} \, .$$

Now,

$$\frac{d\mu_i}{d\nu_i} = \frac{d\mu_i}{d(\mu+\nu)_i} \Big/ \frac{d\nu_i}{d(\mu+\nu)_i} \, ,$$

so

$$\frac{d\mu_i}{d\nu_i} \to \frac{d\mu}{d(\mu+\nu)} \Big/ \frac{d\nu}{d(\mu+\nu)} \, ,$$

where we understand the right hand side to be $+\infty$ when the denominator but not the numerator is zero. Now $d\mu/d(\mu+\nu)$ vanishes precisely on $\Lambda\mu$, and $d\nu/d(\mu+\nu)$ vanishes precisely on Λ_μ, all statements being up to $(\mu+\nu)$ — measure 0. Whence the lemma.

The following fact is known, and we list it for reference :

LEMMA 4. *If $z_1, z_2, \cdots$ are measurable functions with independent Gaussian distributions, mean* 0, *variance* 1, *then the product*

$$a_1 \cdots a_k \, exp\left\{\frac{1}{2}\sum_{i=1}^{h}(1-a_i^2)Z_i^2\right\}$$

converges to zero almost everywhere if

$$\sum_{i=1}^{\infty}|1-a_i^2|^2 = +\infty,$$

and converges to a finite non-zero limit almost everywhere if

$$\sum_{i=1}^{\infty}|1-a_i^2|^2 < \infty .$$

This can be proven, for example, by applying Kakutani's conditions in [6] for equivalence of product measures.

LEMMA 5. *Let R be a closed densely defined linear operator from the Hilbert space* **H** *to the Hilbert space* **K**. *Then there is an equivalence operator U from* **H** *onto* **H** *such that U^*R^*RU has pure point spectrum.*

Proof. Let $\sqrt{R^*R} = \int_0^\infty \lambda \, dF(\lambda)$. Let $F_i = F(2^i) - F(2^{i-1})$, $i = 0, \pm 1, \pm 2 \cdots$, and let $R_i = \sqrt{R^*R} \mid \mathbf{H}_i$, where $\mathbf{H}_i$ is the range of F_i. By a theorem of von Neumann in [7] there is a self-adjoint Hilbert-Schmidt operator H_i in $\mathbf{H}_i$ whose Hilbert-Schmidt norm $\| H_i \|_2 \leqq 2^{-2|i|-3}$, and such that $R_i + H_i$ has pure point spectrum. Now consider the equation $R_i(I+K_i) = R_i + H_i$, that is $R_iK_i = H_i$. Since R_i is invertible in $\mathbf{H}_i$, and, in fact, $\| R_i \|^{-1} \leqq 2^{-i+1}$, we get a solution $K_i = R_i^{-1}H_i$, and

$$\| K_i \|_2 \leqq \| R_i^{-1} \| \, \| H_i \|_2 \leqq 2^{-i+1} \cdot 2^{-2|i|-3} \leqq 2^{-|i|-2}.$$

Let K be defined on **H** by setting $K \mid \mathbf{H}_i = K_i$. Then

$$\| K \|_2^2 = \sum_{-\infty<i<+\infty} \| K_i \|_2^2 \leqq \sum_{-\infty<i<+\infty} 2^{-|i|-2} = 3/4 .$$

so K is Hilbert-Schmidt, and $U = I + K$ is an equivalence. Further, $(I+K_i)^*R_i^2(I+K_i)$ has a complete set of eigenvectors in $\mathbf{H}_i$. But this operator is precisely the restriction of U^*R^*RU to $\mathbf{H}_i$. Therefore U^*R^*RU has a complete set of eigenvectors in **H**.

We shall be considering linear spaces of Gaussian functions. In taking the closures of such linear spaces in the $L_2(\mu)$ norm, the functions obtained as limits will again be Gaussian, as is well known and easy to show, the means and variances of a limit being in fact limits of the means and variances of the approximating Gaussian functions. Furthermore, the topology of convergence in measure on Gaussian functions agrees with $L_2(\mu)$-topology. This is shown in the mean zero case in [8], and the general case can be reduced to this by showing the following :

LEMMA 6. *Let x_i be a net of μ-measurable functions with Gaussian distributions, converging in probability to zero. Then their means γ_i converge to zero.*

Proof. Suppose this does not occur. Then by cutting down to a subnet if necessary, and occasionally using $-x_i$ instead of x_i if necessary, we can assume that there is some $c > 0$ such that $\gamma_i \geqq c$ for all i. Now,

$$\mu\{\omega \mid x_i(\omega) - \gamma_i > 0\} = \mu\{\omega \mid x_i(\omega) - \gamma_i < 0\} ,$$

so that

$$\begin{aligned}\mu\{\omega \mid x_i(\omega) > c\} \geqq \mu\{\omega \mid x, (\omega) > \gamma_i\} &= \mu\{\omega \mid x_i(\omega) < \gamma_i\} \\ &\geqq \mu\{\omega \mid x_i(\omega) < c\} \geqq \mu\{\omega \mid \mid x_i(\omega) \mid < c\} .\end{aligned}$$

The sets on the two ends of the inequality are disjoint, and that on the small end has measure converging to 1, which gives the desired contradiction.

LEMMA 7. *Let μ, ν be nonperpendicular measures. Suppose x_i is Gaussian with respect to μ and ν, $x_i \to 0$ in μ-measure, and $x_i \to x$ in ν-measure. Then $x = 0$* a.e.(ν).

Proof. Since x is Gaussian under ν, the assumption that it is not zero a.e.(ν) implies that it is invertible a.e.(ν). Then $x_i x^{-1} \to 1$ in ν measure, whereas $x_j \to 0$ a.e.(μ), which implies $\mu \perp \nu$.

3. The theorem.

THEOREM *Let* **L** *be a linear space of real-valued functions on a set Ω. Let $\mathscr{S}$ be the smallest σ-field of subsets of Ω with respect to which all the functions in* **L** *are measurable. Let μ and ν be probability measures on $\mathscr{S}$. Suppose all the functions of* **L** *are Gaussian via both measures. Then either $\mu \sim \nu$ or $\mu \perp \nu$. Necessary and sufficient for equivalence is that if we let* **K** *be the linear space generated by* **L** *and the real-valued constant functions, then the μ-equivalence clusses of functions in* **K** *are* *the same as the ν-equivalence classes, and the identity correspondence between the two types of equivalence classes in* $\mathbf{K}$ *is induced by an equivalence operator between the* $\mathbf{L}_2(\mu)$*-closvre of the* μ*-equivalence classes and the* $\mathbf{L}_2(\nu)$*-closure of the* ν*-equivalence classes.*

Proof. First, assume μ not $\perp \nu$. Let $\mathbf{J} = \{x - \int x d\mu \mid x \varepsilon \mathbf{L}\}$. For any function x, let x^μ (respectively x^ν) denote the equivalence class of x modulo functions which are μ-null (respectively ν-null), and, for a set S of functions, let $\mathbf{S}_\nu, \mathbf{S}_\mu$ denote the corresponding set of equivalence classes. $\overline{\mathbf{S}}_\mu$ will mean the $\mathbf{L}_2(\mu)$ closure of S.

All elements in $\mathbf{K}$ are Gaussian under μ and ν, and the correspondence $x^\mu \leftrightarrow x^\nu$ between $\mathbf{K}_\mu$ and $\mathbf{K}_\nu$ is one-to-one and closable, by Lemma 5. So there is a one-to-one closed operator T from a dense subspace $\mathbf{D}_T$ of $\overline{\mathbf{K}}_\mu$ to a dense subspace $\mathbf{R}_T$ of $\overline{\mathbf{K}}_\nu$ such that $Tx^\mu = x^\nu$ for all x in $\mathbf{K}$. Further, given any ξ in $\mathbf{D}_T$, there is some $\mathscr{S}$-measurable x such that $\xi = x^\mu$ and $T\xi = x^\nu$. For choose x_i in $\mathbf{K}$ such that $x_i^\mu \rightarrow \xi$, $x_i^\nu \rightarrow T\xi$. By taking subsequences, the convergence can be made a.e.(μ) and a.e.(ν) respectively, so that ξ and $T\xi$ must agree a.e.($\mu \wedge \nu$).

Let $S = T \mid \mathbf{D}_T \cap \bar{J}_\mu$. Then S is closed, with dense domain in $\bar{J}_\mu$ and dense range in $\bar{J}_\nu$, by Lemma 6. Lemma 5 gives us an equivalence U in $\bar{J}_\mu$ such that U^*S^*SU has pure point spectrum. Choose $y_1, y_2, \cdots$ such that the y_i^μ are a complete orthonormal set of eigenvectors for $U^{-1*}U^{-1}$, with eigenvalues a_i^2. Then the vectors $U^{-1}y_i^\mu$ are again orthogonal, and $\| U^{-1}y_i^\mu \| = a_i^{-1}$. Let $\mathscr{S}_N$ be the sample space of $y_1, \cdots, y_N$. Put a new measure μ' on $\mathscr{S}$ by letting $y_1, y_2, \cdots$ be Gaussian, independent, mean 0, variance $1/a_1^2, 1/a_2^2, \cdots$ Then

$$\frac{d\mu' \mid \mathscr{S}_N}{d\mu \mid \mathscr{S}_N} = a_1 \cdots a_N \, exp\left\{\frac{1}{2}\sum_{i=1}^{N}(1 - a_i^2)y_i^2\right\},$$

which converges almost everywhere $(\mu + \nu)$ to a nonzero limit, so that $\mu' \sim \mu$.

Now we wish to show $\mu' \sim \nu$. We have $x^\nu = Sx^\mu$ for $x \in \mathbf{J}$, so $x^\nu = (SU)U^{-1}x^\mu = (SU)x^{\mu'}$. Let $S' = SU$. Then, taking a.e. limits on both sides, one can for every ξ in $\mathbf{D}_{S'}$ find some $\mathscr{S}$-measurable x such that $x^{\mu'} = \xi$ and $x^\nu = S'\xi$. Now choose functions $Z_1, Z_2, \cdots$ such that $Z_i^{\mu'}$ form a complete orthonormal set of eigenvalues for S'^*S'. The $S'Z_i^{\mu'} = Z_i^\nu$ are also orthogonal and span $\mathbf{J}_\nu$. Define $\int Z_i d_\nu = \gamma_i$ and $\int |Z_i|^2 d\nu = \alpha_i^2$. Then α_i is never zero, since S' is nonsingular, and $\int (Z_i - \gamma_i)(Z_j - \gamma_j)d\nu = \alpha_i^2\delta_{ij} - \gamma_i\gamma_j$, where δ_{ij} is the Kronecker delta function. So the covariance matrix of $Z_1, \cdots Z_N$ in ν-measure is given by $C_N = A_N^2 - \vec{\gamma}_N \otimes \vec{\gamma}_N$, where A_N is the matrix

$$\begin{pmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_N \end{pmatrix},$$

$\vec{\gamma}_N$ is the vector $(\gamma_1, \cdots, \gamma_N)$, and the notation $\vec{\gamma}_N \otimes \vec{\gamma}_N$ represents a dyadic operator. C_N is, of course, nonnegative definite. Let $\delta_i = \gamma_i/\alpha_i$, and $\vec{\delta}_N = (\delta_1, \cdots, \delta_N)$ So $\vec{\delta}_N = A_N^{-1}\vec{\gamma}_N$, and $A_N^{-1}C_N A_N^{-1} = I - \vec{\delta}_N \otimes \vec{\delta}_N$ is a nonnegative definite matrix. By conjugating with an orthogonal matrix, this can be transformed into the equivalent matrix $I - \| d_N \|^2 E$, where

$$E = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & & \cdots & 0 \end{pmatrix}.$$

The determinant of this, and hence of $I - \vec{\delta}_N \otimes \vec{\delta}_N$, is $I - \| \vec{\delta}_N \|^2$. Thus

$$| C_N | = | A_N |^2 | 1 - \| \vec{\delta}_N \|^2 | = \alpha_1^2 \cdots \alpha_N^2 (1 - \sum_{i=1}^{N} \delta_i^2) .$$

Observe that

$$\delta_i = \frac{\gamma_i}{\alpha_i} = \frac{(Z_i, 1)}{\sqrt{(Z_i, Z_i)}} ,$$

where $(\cdot\,, \cdot)$ denotes the inner product in $\mathbf{L}_2(\nu)$, so that, since $1 \notin \bar{J}\nu$, $\sum \delta^2 < (1, 1) = 1$. Thus $| C_N | \neq 0$, and C_N is nonsingular. The inverse matrix to C_N is

$$C_N^{-1} = (A_N^2 - \vec{\gamma}_N \otimes \vec{\gamma}_N)^{-1} = A_N^{-2} + \frac{A_N^{-2}\vec{\gamma}_N \otimes A_N^{-2}\vec{\gamma}_N}{1 - \| A_N^{-1}\vec{\gamma}_N \|^2} .$$

Now let $\mathscr{F}_N$ be the sample space of $Z_1, \cdots, Z_N$ and let $\mu_N = \mu' \,|\, \mathscr{F}_N$, $\nu_N = \nu \,|\, \mathscr{F}_N$. Then $d\nu_N/d\mu_N$ is precisely

$$\sqrt{| C_N |}^{-1} \exp \left\{ \frac{1}{2} [\langle \vec{Z}_N, \vec{Z}_N \rangle - \langle C_N^{-1}(\vec{Z}_N - \vec{\gamma}_N), (\vec{Z}_N - \vec{\gamma}_N) \rangle] \right\} ,$$

where $\vec{Z}_N = (Z_1, \cdots, Z_N)$. A calculation shows that the exponent can be written

$$\frac{1}{2} [\| \vec{Z}_N \|^2 - \| A_N^{-1}\vec{Z}_N \|^2 - \frac{1}{1 - \| \vec{\delta}_N \|^2} (\langle \vec{\delta}_N, A_N^{-1}\vec{Z}_N \rangle - 1)^2 - 1] .$$

Consider the convergence of $d\nu_N/d\mu_N$. At this stage one could already conclude from the zero-one law that $\{\omega \,|\, d\nu_N/d\mu_N(\omega) \to 0\}$ has measure 0 or 1, since the set is independent of $Z_1, \cdots, Z_N$ for each N. However, we wish to get precise conditions for when this occurs.

(a) Suppose

$$\sum_{i=1}^{\infty} \left| \frac{1}{\alpha_i^2} - 1 \right|^2 = \infty .$$

Consider the factor

$$\sqrt{|C_N|}^{-1} \exp\left\{\frac{1}{2}[\| \vec{Z}_N \|^2 - \| A_N^{-1}\vec{Z}_N \|^2\right\}$$
$$= (1 - \sum_{i=1}^{N} \delta_i^2)^{-1/2}\left(\frac{1}{\alpha_i} \cdots \frac{1}{\alpha_N}\right) \exp \frac{1}{2} \sum_{i=1}^{N} \left(1 - \frac{1}{\alpha_i^2}\right) Z_i^2 .$$

Applying Lemma 3, this factor converges to zero almost everywhere with respect to μ'. The other factor of $d\nu_N/d\mu_N$, namely

$$\exp \frac{1}{2}\left\{- \left(\frac{\langle\vec{\delta}_N, A_N^{-1}\vec{Z}_N\rangle - 1}{1 - \| \delta_N \|^2}\right)^2 - 1\right\} ,$$

is clearly bounded above, so in this case $d\nu_N/d\mu_N \to 0$.

(b) Suppose

$$\sum_{i=1}^{\infty} \left| \frac{1}{\alpha_i^2} - 1 \right|^2 < \infty .$$

Then, again by Lemma 3, the factor

$$\frac{1}{\alpha_1 \cdots \alpha_N} \exp \frac{1}{2} \sum_{i=1}^{N} \left(1 - \frac{1}{\alpha^2}\right) Z_i^2$$

converges almost everywhere μ' to a finite limit. The remaining factor is, except for a constant,

$$\frac{1}{1 - \| \vec{\delta}_N \|^2} = \exp \frac{1}{2}\left\{- \frac{1}{1 - \| \delta_N \|^2}\left(\sum_{i=1}^{N} \frac{\delta_i Z_i}{\alpha_i} - 1\right)^2\right\} .$$

Since $\sum_{i=1}^{\infty}\delta_i^2 < 1$, everything in sight converges, because $\sum_{i=1}^{\infty}(\delta_i/\alpha_i)^2 < \infty$. So $\mu' \sim \nu$, and S' is an equivalence from $\bar{\mathbf{J}}_\mu$ to $\bar{\mathbf{J}}_\nu$. Then S is likewise, and therefore T is an equivalence operator.

Conversely if $\mathbf{L}$ consists of Gaussian functions under μ and ν, and the correspondence $x_\mu + c \leftrightarrow x_\nu + c$, $x \in \mathbf{L}$, is the restriction of an equivalence operator T from $\mathbf{K}_\mu$ to $\mathbf{K}_\nu$, then again choosing a basis of eigenvectors for $T \mid \bar{\mathbf{J}}_\mu$, we get convergence of the Radon-Nikodym derivatives to a non-zero limit, because of Lemma 3, and therefore get equivalence of the induced measures.

4. An example. Let T be a set, Ω the set of all real-valued functions on T. Let $x_t(\omega) = \omega(t)$, and let $\mathscr{S}$ be the smallest σ-algebra with respect to which all the x_t are measurable. Let μ, ν be measures on $\mathscr{S}$, by each of which x_t becomes a Gaussian stochastic process. Let

$$m(t) = \int x_t d\mu, \ \rho(s, t) = \int x_s x_t d\mu - m(s)m(t),$$

$$n(t) = \int x_t d\nu, \ \sigma(s, t) = \int x_s x_t d\nu - n(s)n(t) \ .$$

Let τ be a measure on T such that ρ, σ, m, n, become measurable. Define, for τ-measurable f

$$[f, f]_\mu = \iint (\rho(s, t) + m(s)m(t))f(s)f(t)d\tau(s)d\tau(t) \ ,$$

$$[f, f]_\nu = \iint (\sigma(s, t) + n(s)n(t))f(s)f(t)d\tau(s)d\tau(t) \ ,$$

and $[f, f] = [f, f]_\mu + [f, f]_\nu$. Let $\mathbf{L}_0$ be those f for which $[f, f] < \infty$. Then we get inner products $[f, g]_\mu$, etc. on $\mathbf{L}_0$. Define $\int f(t)x_t d\tau(t)$ as an $\mathbf{L}_2(\mu + \nu)$ valued integral, for $f \in \mathbf{L}_0$. This can be done, and in fact

$$[f, g]_\mu = \int\Big(\int f(t)x_t d\tau(t)\Big)\Big(\int g(t)x_s d\tau(s)\Big)d\mu \ ,$$

$$[f, g]_\nu = \int\Big(\int f(t)x_t d\tau(t)\Big)\Big(\int g(s)x_s d\tau(s)\Big)d\nu \ .$$

Let $\mathbf{L} = \Big\{x \mid$ there is some f in $\mathbf{L}_0$ for which x has as its $\mu + \nu$ equivalence class $\int f(t)x_t d\tau(t)\Big\}$. $\mathbf{L}$ is a linear set of functions, all Gaussian with respect to either μ or ν. Let $\mu_0 = \mu \mid \mathscr{S}_0$ and $\nu_0 = \nu \mid \mathscr{S}_0$. We know from our theorem that μ_0 and ν_0 are either equivalent or perpendicular. Let $\mathbf{H}$ and $\mathbf{K}$ be the Hilbert spaces gotten by completing the μ-equivalence classes of $\mathbf{L}_0$ in $\mathbf{L}_2(\mu)$ and the ν-equivalence classes in $\mathbf{L}_2(\nu)$. The inner products then come from $[f, g]_\mu$ and $[f, g]_\nu$. Necessary for equivalence is that the identity map on $\mathbf{L}_0$ induce an equivalence operator from $\mathbf{H}$ to $\mathbf{K}$, and in order to get sufficiency we just have to be sure that if $\mathbf{L}_0$ is enlarged to include multiples of the identity, the identity map still induces an equivalence operator on the Hilbert spaces. This amounts to requiring that 1 be an $\mathbf{L}_2(\mu)$ limit of functions in $\mathbf{L}_0$ if and only if it is also an $\mathbf{L}_2(\nu)$ limit.

BIBLIOGRAPHY

1. Baxter : *A strong limit theorem for Gaussian processes*, Proc. Amer. Math. Soc., (1956).

2. Cameron and Martin: *The behavior of measurabilty under change of scale Wiener space*, Bull. Amer. Math., Soc., (1947).

3. Cameron and Martin: *Transformation of Wiener integrals under translation of Wiener integrals under translation*, Ann. Math., Stat. (1944).

4. Cameron and Martin: *Transformation of Wiener integrals under a general class of linear transformations*, Trans. Amer. Math. Soc., 1945.
5. Grenander: *Stochastic processes and statistical inference*, Arkiv für Matematik, Band 1, nr. 17.
6. S. Kakutani: *On equivalence of infinite product measures*, Ann. Math. Stat., 1948.
7. J. von Neumann: *Charakterisierung des Spektrums eines Integral-operatoren*, Actualités Sci. Ind. 229, Paris (1935).
8. L. E. Segal: *Distributions in Hilbert spaces and canonical systems of operators* (from privately distributed mimeographed notes).

UNIVERSITY OF CALIFORNIA, BERKELEY

MINIMAL COVERINGS OF PAIRS BY TRIPLES

M. K. FORT, JR. AND G. A. HEDLUND

1. Introduction. Let F be a finite set with n members, $n \geqq 3$. An *F-covering of pairs by triples*, which we abbreviate F-copt, is a set S of triples of distinct members of F which has the property that each pair of distinct members of F is contained in at least one member of S. If n is a positive integer, $n \geqq 3$, then an n-copt is an F-copt for the set $F = \{1, 2, \cdots, n\}$. We assume throughout that $n \geqq 3$.

For any finite set A, let $C(A)$ denote the number of members of A. An F-copt S is *minimal* if $C(S) \leqq C(S')$ for every F-copt S'. If $n \equiv 1 \pmod 6$ or $n \equiv 3 \pmod 6$, then a minimal n-copt S turns out to be *exact* in the sense that each pair is contained in exactly one member of S. Such exact coverings are called *Steiner triple systems*. The existence of Steiner triple systems for all n (of form $6h + 1$ or $6h + 3$) was proved by M. Reiss [2] in 1859.

Let S be a minimal n-copt and let $C(S) = \mu(n)$. The main result of this paper is obtained in §2, where we determine $\mu(n)$ explicitly for $n \geqq 3$. In § 3 we discuss certain properties of minimal n-copts, and give several methods for constructing minimal n-copts.

2. Determination of $\mu(n)$. Let S be a minimal n-copt. For each integer i, $1 \leqq i \leqq n$, we define $\alpha(i)$ to be the number of members of S that contain i. Then

$$\sum_{i=1}^{n} \alpha(i) = 3 \cdot C(S) .$$

Since i must appear in members of S with $n - 1$ other numbers we have $\alpha(i) \geqq [n/2]$. ($[x]$ is the largest integer which is not greater than x.) Thus,

$$\mu(n) = C(S) \geqq \frac{n}{3}\left[\frac{n}{2}\right] . \tag{1}$$

Since $(n/3)\,[n/2]$ may not be an integer, we define $\varphi(n)$ to be the least integer which is not less than $(n/3)\,[n/2]$. It is easy to compute

$$\varphi(n) = \begin{cases} n^2/6 \text{ if } n = 6k , \\ n(n-1)/6 \text{ if } n = 6k+1 \text{ or } n = 6k+3 , \\ (n^2+2)/6 \text{ if } n = 6k+2 \text{ or } n = 6k+4 , \\ (n^2-n+4)/6 \text{ if } n = 6k+5. \end{cases} \tag{2}$$

Received May 26, 1958. The preparation of this paper was sponsored in part by the Office of Naval Research.

We may clearly improve (1) to

$$\mu(n) = C(S) \geqq \varphi(n) . \tag{3}$$

Our main theorem proves that in (3) equality holds for every n.

Let A, B and C be pairwise disjoint sets, each having the same number n of members. A *tricover* for the system (A, B, C) is a set K of triples (x, y, z), $x \in A, y \in B, z \in C$ such that each pair uv, u and v in different ones of A, B, C, is contained in exactly one member of K.

LEMMA 1. *If n is a positive integer and A, B, C are pairwise disjoint sets each of which has n members, then a tricover K for (A, B, C) exists. Moreover, if $a \in A$, $b \in B$ and $c \in C$, then K may be chosen so that $(a, b, c) \in K$.*

Proof. Let the members of A, B, C be respectively

$$a_1, a_2, \cdots, a_n ; \qquad b_1, b_2, \cdots, b_n ; \qquad c_1, c_2 \cdots, c_n ,$$

where $a_1 = a$, $b_1 = b$, $c_1 = c$. We define K to be the set of all triples (a_i, b_j, c_k) for which $k \equiv i + j - 1 \pmod n$, $1 \leqq i, j, k \leqq n$. The set K obviously has the desired properties.

REMARK. Any tricover for (A, B, C) must have n^2 members.

LEMMA 2. *Let A, B, C be pairwise disjoint sets, each having n members. Let p be an integer such that $0 < p \leqq n/2$. Let $A^* \subset A$, $B^* \subset B$, $C^* \subset C$ be sets, each of which has p members and let K^* be a tricover for (A^*, B^*, C^*). Then there exists a tricover K for (A, B, C) such that $K^* \subset K$.*

Proof. Let

$$A = \{a_1, a_2, \cdots, a_n\} ,$$
$$B = \{b_1, b_2, \cdots, b_n\} ,$$
$$C = \{c_1, c_2, \cdots, c_n\} .$$

We can assume that

$$A^* = \{a_1, a_2, \cdots, a_p\} ,$$
$$B^* = \{b_1, b_2, \cdots, b_p\} ,$$
$$C^* = \{c_1, c_2, \cdots, c_p\} .$$

For $1 \leqq i, j \leqq p$, let m^*_{ij} be the unique integer k such that $(a_i, b_j, c_k) \in K^*$. Clearly $1 \leqq m^*_{ij} \leqq p$ and the square array (m^*_{ij}) is a Latin square of order p. It follows from a theorem of Marshall Hall [1] that there exists a Latin square (m_{ij}), $1 \leqq i, j \leqq n$, such that $m_{ij} = m^*_{ij}$,

$1 \leqq i, j \leqq p$. Let

$$K = \{(a_i, b_j, c_{m_{ij}}) \mid 1 \leqq i, j \leqq n\} .$$

The set K is the desired tricover.

In order to produce an inductive proof of our main theorem, it is convenient to restrict ourselves to a special type of minimal n-copt for the case $n \equiv 5 \pmod 6$. Also, for $n \equiv 3 \pmod 6$, there is a special type of minimal n-copt whose existence we wish to establish, and it is possible to include this result in our main theorem. For these reasons we introduce the notion of "admissible F-copt."

An F-copt S is *admissible* if $C(S) = \varphi(n)$, $n = C(F)$, and:
(1) $n \equiv 0, 1, 2,$ or $4 \pmod 6$;
(2) $n \equiv 3 \pmod 6$ and S contains a set of pairwise disjoint triples whose union is F; or
(3) $n \equiv 5 \pmod 6$ and S contains four elements of the form (a, b, x), (a, b, y), (a, b, z), (x, y, z).

THEOREM 1. *If n is a positive integer, $n \geqq 3$, then there exists an admissible n-copt.*

Proof. Our proof is by induction on n. However, it is necessary to prove independently that there are admissible n-copts for $n = 3, 5, 7, 9, 11, 13$, and 15. We accomplish this by exhibiting such admissible n-copts.

n = 3

(1,2,3)

n = 5

(1,2,3)
(1,2,4)
(1,2,5)
(3,4,5)

n = 7

(1,2,3)
(1,4,5)
(1,6,7)
(2,4,6)
(2,5,7)
(3,4,7)
(3,5,6)

n = 9

(1,2,3)	(2,4,9)
(4,5,6)	(2,5,8)
(7,8,9)	(2,6,7)
(1,4,7)	(3,4,8)
(1,5,9)	(3,5,7)
(1,6,8)	(3,6,9)

n = 11

(1, 2, 3)	(3, 6,10)
(1, 2, 4)	(3, 7, 9)
(1, 2, 5)	(3, 8,11)
(3, 4, 5)	(4, 6,11)
(1, 6, 7)	(4, 7, 8)
(1, 8, 9)	(4, 9,10)
(1,10,11)	(5, 6, 9)
(2, 6, 8)	(5, 7,11)
(2, 7,10)	(5, 8,10)
(2, 9,11)	

n = 13

(1, 2, 3)	(3, 6,12)
(1, 4, 5)	(3, 8,13)
(1, 6,13)	(3, 9,10)
(1, 7, 8)	(4, 6, 7)
(1, 9,12)	(4, 8, 9)
(1,10,11)	(4,12,13)
(2, 4,10)	(5, 8,11)
(2, 5, 6)	(5, 9,13)
(2, 7, 9)	(5,10,12)
(2. 8,12)	(6, 8,10)
(2,11,13)	(6, 9,11)
(3, 4,11)	(7,10,13)
(3, 5, 7)	(7,11,12)

$n = 15$

(1, 2, 3)	(2, 6, 8)	(3,12,14)	(6, 9,14)
(1, 4,14)	(2, 7,14)	(4, 5, 6)	(6,12,13)
(1, 5, 9)	(2, 9,11)	(4, 8,13)	(7, 8, 9)
(1, 6,10)	(2,10,15)	(4, 9,10)	(7,10,13)
(1, 7,12)	(3, 4, 7)	(4,11,15)	(8,11,14)
(1, 8,15)	(3, 5,11)	(5, 7,15)	(9,12,15)
(1,11,13)	(3, 6,15)	(5, 8,12)	(10,11,12)
(2, 4,12)	(3, 8,10)	(5,10,14)	(13,14,15)
(2, 5,13)	(3, 9,13)	(6, 7.11)	

Our proof now divides into six cases. In Case r, $0 \leqq r \leqq 5$, we assume that $n \equiv r \pmod 6$, that $n > 3$ and that there exist admissible m-copts for $3 \leqq m < n$. We then show that these assumptions imply that there exists an admissible n-copt.

Case 0. Let S_1 be an admissible $(n-1)$-copt having $(1, 2, 3), (1, 2, 4)$, and $(1, 2, 5)$ as three of its members. If we delete $(1, 2, 3)$ from S_1 and add

$$(1, 3, n), (2, 3, n), (4, 5, n), (6, 7, n), \cdots, (n-2, n-1, n) ,$$

we obtain a set S of triples which is an n-copt. Since S_1 has

$$[(n-1)^2 - (n-1) + 4]/6 = (n^2 - 3n + 6)/6$$

members, S has

$$(n^2 - 3n + 6)/6 - 1 + n/2 = n^2/6 = \varphi(n)$$

members.

Case 1. We have exhibited admissible n-copts for $n = 7$ and $n = 13$. Therefore we may assume $n = 6h + 1, h > 2$.

We consider two subcases.

Subcase i. Either $h \equiv 0$ or $h \equiv 1 \pmod 3$. Then there exists k such that $2h + 1 = 6k + 1$ or $2h + 1 = 6k + 3$.

Let

$$A_1 = \{1, \cdots, 2h, n\}$$
$$A_2 = \{2h + 1, \cdots, 4h, n\}$$
$$A_3 = \{4h + 1, \cdots, 6h, n\}$$

and let S_j be an admissible A_j-copt for $j = 1, 2, 3$. Let T be a tricover for $(\{1, \cdots, 2h\}, \{2h + 1, \cdots, 4h\}, \{4h + 1, \cdots, 6h\})$. We now define $S = S_1 \cup S_2 \cup S_3 \cup T$. It is easy to verify that S is an n-copt, and that S has

$$3 \cdot \frac{(2h+1)2h}{6} + (2h)^2 = \frac{n(n-1)}{6} = \varphi(n)$$

members.

Subcase ii. $h \equiv 2 \pmod 3$. In this case there exists k such that $2h+1 = 6k+5$. We define A_1, A_2, A_3 as above. Now, for $j = 0, 1, 2$, we let S_{j+1} be an admissible A_{j+1}-copt such that S_{j+1} contains a subset R_{j+1} whose members are:

$$\begin{array}{l}(2jh+1, 2jh+2, 2jh+3)\\(2jh+1, 2jh+2, 2jh+4)\\(2jh+1, 2jh+2, n)\\(2jh+3, 2jh+4, n)\,.\end{array}$$

Let T be a tricover for $(\{1, \cdots, 4\}, \{2h+1, \cdots, 2h+4\}, \{4h+1, \cdots, 4h+4\})$, and let T^* be a tricover for $(\{1, \cdots, 2h\}, \{2h+1, \cdots, 4h\}, \{4h+1, \cdots, 6h\})$ that is an extension of T. Since $h \geqq 5$, the existence of such a tricover follows from Lemma 2. We next take an admissible copt U for

$$\{1, \cdots, 4, 2h+1, \cdots, 2h+4, 4h+1, \cdots, 4h+4, n\}\,.$$

Finally, we define

$$S = (S_1 - R_1) \cup (S_2 - R_2) \cup (S_3 - R_3) \cup (T^* - T) \cup U\,.$$

It is easy to check that S is an n-copt. The number of member of S is

$$3 \cdot \left[\frac{(2h+1)^2 - (2h+1) + 4}{6} - 4\right] + \left[(2h)^2 - 16\right] + 26$$
$$= 6h^2 + h = \frac{n(n-1)}{6}.$$

Thus, S is admissible.

Case 2. Let S_1 be an admissible $(n-1)$-copt. We define S to be the set of triples obtained by adding to S_1 the triples

$$(1, 2, n), (3, 4, n), \cdots, (n-3, n-2, n), (n-2, n-1, n)\,.$$

Then, S is an n-copt and S has

$$\frac{(n-1)(n-2)}{6} + \frac{n}{2} = \frac{n^2+2}{6}$$

members. Thus S is admissible.

Case 3. There exists h such that $n = 6h+3$. Since we have listed admissible n-copts for $n = 3, 9, 15$, we may assume $h > 2$. We consider two subcases.

Subcase i. $h \equiv 0$ or $h \equiv 1 \pmod 3$. In this case there exists k such that $2h + 1 = 6k + 1$ or $2h + 1 = 6k + 3$. Let S_1 be an admissible $(2h+1)$—copt. For each triple $(a, b, c) \in S_1$ we choose a tricover for $(\{3a-2, 3a-1, 3a\}, \{3b-2, 3b-1, 3b\}, \{3c-2, 3c-1, 3c\})$. The union of all such tricovers, together with the triples $(1, 2, 3), (4, 5, 6), \cdots, (n-2, n-1, n)$ is an n-copt S. The number of members of S is

$$9 \cdot \frac{(2h+1) \cdot 2h}{6} + (2h+1) = (2h+1)(3h+1) = \frac{n(n-1)}{6} .$$

If follows that S is admissible.

Subcase ii. $h \equiv 2 \pmod 3$. In this case there exists k such that $2h + 1 = 6k + 5$. We choose an admissible $(2h + 1)$-copt S_1 that contains the triples $(1, 2, 3), (1, 2, 4), (1, 2, 5), (3, 4, 5)$. If (a, b, c) is any other member of S_1, we choose a tricover for $(\{3a-2, 3a-1, 3a\}, \{3b-2, 3b-1, 3b\}, \{3c-2, 3c-1, 3c\})$. Let S_2 be the 15-copt exhibited at the beginning of our proof. We now define S to be the set whose members are the members of S_2, the members of the chosen tricovers, and the triples $(16, 17, 18), \cdots, (n-2, n-1, n)$. S is an n-copt, and the number of members of S is

$$35 + 9\left[\frac{(2h+1)^2 - (2h+1) + 4}{6} - 4\right] + \frac{n-15}{3} = \frac{n(n-1)}{6} .$$

Since S has $(1, 2, 3), (4, 5, 6), \cdots, (n-2, n-1, n)$ as members, S is admissible.

Case 4. For this case, the construction is exactly the same as in Case 2.

Case 5. We first observe that numbers of the form $6h + 5$, h a non-negative integer, form the same set as numbers of the form $3s - 4$, s an odd integer and $s > 1$. We have listed an admissible 5-copt, and an admissible 11-copt. Thus, we may assume $n = 6h + 5 = 3s - 4$, $s > 5$. We consider two subcases.

Subcase i. There exists k such that $s = 6k + 1$ or $s = 6k + 3$. In this case, we let

$$A_1 = \{1, \cdots, s-2\}$$
$$A_2 = \{s-1, \cdots, 2s-4\}$$
$$A_3 = \{2s-3, \cdots, 3s-6\} .$$

There is a tricover K of (A_1, A_2, A_3) such that $(1, s-1, 2s-3) \in K$. For $i = 1, 2, 3$ we define

$$R_i = A_i \cup \{3s-5, 3s-4\}\ .$$

and let S_i be an admissible R_i-copt such that $(1, 3s-5, 3s-4) \in S_1$, $(s-1, 3s-5, 3s-4) \in S_2$ and $(2s-3, 3s-5, 3s-4) \in S_3$. We define $S = K \cup S_1 \cup S_2 \cup S_3$. It is easy to see that S is an n-copt, and that S has

$$(s-2)^2 + \frac{3s(s-1)}{6} = \frac{3s^2-9s+8}{2} = \frac{n^2-n+4}{6}$$

members. Since $(1, 3s-5, 3s-4)$, $(s-1, 3s-5, 3s-4)$, $(2s-3, 3s-5, 3s-4)$, $(1, s-1, 2s-3)$ are members of S, S is admissible.

Subcase ii. There exists k such that $s = 6k+5$. We define

$$A_1 = \{1, \cdots, s-2\}$$
$$A_2 = \{s-1, \cdots, 2s-4\}$$
$$A_3 = \{2s-3, \cdots, 3s-6\}$$

and let $R_i = A_i \cup \{3s-5, 3s-4\}$ for $i = 1, 2, 3$. By the inductive hypothesis, there exists an admissible R_i-copt S_i such that S_i contains the set B_i, where

$$B_1 = \{1, 2, 3), (1, 3s-5, 3s-4), (2, 3s-5, 3s-4), (3, 3s-5, 3s-4)\}\ ,$$
$$B_2 = \{(s-1, s, s+1), (s-1, 3s-5, 3s-4), (s, 3s-5, 3s-4), (s+1, 3s-5, 3s-4)\}\ .$$
$$B_3 = \{(2s-3, 2s-2, 2s-1), (2s-3, 3s-5, 3s-4), (2s-2, 3s-5, 3s-4), (2s-1, 3s-5, 3s-4)\}\ .$$

Let $G = \{1, 2, 3, s-1, s, s+1, 2s-3, 2s-2, 2s-1, 3s-5, 3s-4\}$. G has 11 members, and hence there exists an admissible G-copt M.

We choose a tricover T_1 for $(\{1, 2, 3\}, \{s-1, s, s+1\}, \{2s-3, 2s-2, 2s-1\})$ and extend T_1 to a tricover T for (A_1, A_2, A_3).

We now define

$$S = (S_1 - B_1) \cup (S_2 - B_2) \cup (S_3 - B_3) \cup M \cup (T - T_1)\ .$$

It is a routine matter to verify that S is an n-copt. The number of members of S is

$$3\left[\frac{s^2-s+4}{6} - 4\right] + 19 + \left[(s-2)^2 - 9\right] = \frac{3s^2-9s+8}{2} = \frac{n^2-n+4}{6}\ .$$

Since $S \supset M$ and M is admissible, it follows that S is admissible.

3. Properties of minimal n-copts. Let S be a minimal n-copt. If $n \equiv r \pmod 6$, for $r = 0, 2, 4, 5$, then the covering is not exact and some pairs must be contained in more than one member of S. However, it is possible to state precisely the way in which this sort of "multiple covering" takes place. Our results are contained in the next three theorems.

THEOREM 2. *Let $n = 6k$, and let S be an n-copt for which $C(S) = \varphi(n)$. There exists a partition of $\{1, 2, \cdots, n\}$ into $3k$ pairs $P_1, P_2, \cdots, P_{3k}$, each of which is contained in exactly two members of S. Every other pair (u, v), $1 \leqq u < v \leqq n$, is contained in exactly one member of S.*

Proof. For $1 \leqq j \leqq n$, let $f(j)$ be the number of members of S that contain j. It is clear that $f(j)$ is at least $n/2$, so that $f(j) = n/2 + g(j)$, $g(j) \geqq 0$. We obtain

$$\sum_{j=1}^{n} f(j) = 3\varphi(n) .$$

Thus

$$\sum_{j=1}^{n}\left[\frac{n}{2} + g(j)\right] = 3 \cdot \frac{n^2}{6} , \text{ and}$$

$$\frac{n^2}{2} + \sum_{j=1}^{n} g(j) = \frac{n^2}{2} .$$

We see that $g(j) = 0$ for $j = 1, \cdots, n$ and $f(j) = n/2$. Since for each $k \neq j$ there is at least one member of S which contains (j, k), there must exist $j^* \neq j$ such that (j, j^*) is contained in exactly two members of S, and (j, k) is contained in exactly one member of S for $j \neq k \neq j^*$. Moreover, $j^{**} = j$, and hence the pairs (j, j^*) are the $n/2$ pairs $P_1, P_2, \cdots, P_{3k}$.

THEOREM 3. *Let $n = 6k + 2$ or $n = 6k + 4$, and let S be an n-copt for which $C(S) = \varphi(n)$. There exist $n/2 + 1$ pairs $P_1, \cdots, P_{n/2+1}$ which are contained in exactly two members of S. Every other pair is contained in exactly one member of S. There exists an integer m which is contained in exactly three of the pairs $P_1, \cdots, P_{n/2+1}$. Every other integer is contained in exactly one of the pairs $P_1, \cdots, P_{n/2+1}$.*

Proof. Let $f(j)$ be the number of members of S that contain the integer j. Since $f(j) \geqq n/2$, we can write

$$f(j) = \frac{n}{2} + g(j) , \qquad g(j) \geqq 0 .$$

Then

$$\sum_{j=1}^{n} f(j) = \frac{n^2}{2} + \sum_{j=1}^{n} g(j) = 3 \cdot \varphi(n) = \frac{n^2+2}{2} .$$

Thus $\sum_{j=1}^{n} g(j) = 1$. There exists an integer m such that $g(m) = 1$ and $g(j) = 0$ for $j \neq m$.

Now suppose $j \neq m$. There must exist j^* such that (j, j^*) is contained in exactly two members of S, and (j, h) is contained in exactly one member of S for $j \neq h \neq j^*$.

Since there are $n/2 + 1$ members of S that contain m, and each pair (m, j) is contained in at least one and not more than two members of S, there exist a, b, c, such that $(m, a), (m, b), (m, c)$ are each contained in exactly two members of S, and (m, j) is contained in exactly one member of S if $j \neq a, j \neq b$, and $j \neq c$.

If j is a member of $T = \{1. \cdots, n\} - \{m, a, b, c\}$, then $j^{**} = j$. Hence T is partitioned into pairs $P_1, P_2, \cdots, P_{(n-4)/2}$, each of which is contained in exactly two members of S. These pairs, together with $(m, a), (m, b), (m, c)$ form the set $P_1, \cdots, P_{n/2+1}$.

THEOREM 4. *If $n = 6k + 5$ and S is a minimal n-copt for which $\varphi(n) = (n^2 - n + 4)/6$, then one pair is contained in three members of S and every other pair is contained in exactly one member of S.*

Proof. For $1 \leqq j \leqq n$, we define $f(j)$ to be the number of members of S that contain j. Clearly $f(j) \geqq (n-1)/2$. We define $g(j) = f(j) - (n-1)/2$. Since $\sum_{j=1}^{n} f(j) = 3\varphi(n) = (n^2 - n + 4)/2$, we obtain

$$\sum_{j=1}^{n} g(j) = 2 .$$

There exists j_1 such that $g(j_1) > 0$. Since there are more than $(n-1)/2$ triples of S that contain j_1, there exists j_2 such that the pair (j_1, j_2) is contained in at least two triples $(j_1, j_2, j_3), (j_1, j_2, j_4)$. The integer j_2 must be in triples with $n - 4$ integers other than j_1, j_3, j_4, and it requires at least $(n-3)/2$ triples to satisfy this condition. Thus $f(j_2) \geqq (n+1)/2$ and $g(j_2) > 0$. We now see that $g(j_1) = g(j_2) = 1$ and $g(j) = 0$ if $j_1 \neq j \neq j_2$.

It now follows that if (u, v) is a pair for which $g(u) = 0$ or $g(v) = 0$, then (u, v) is contained in exactly one member of S. Since $3\varphi(n) = n(n-1)/2 + 2$, the pair (j_1, j_2) must be contained in three members of S.

Our Theorem 1 is of a constructive nature, and indicates how minimal n-copts can be constructed out of minimal m-copts for $m < n$. There are other methods, however, of constructing minimal n-copts out of minimal m-copts for $m <$ n. We give a lemma and theorem due to Reiss [2] which are useful in this connection. Our final theorem is analogous to the Reiss Theorem.

REISS LEMMA. *Let n be a positive integer. Let*

$$P = \{(u, v) \mid 1 \leqq u < v \leqq 2n\} .$$

Then there exists a partition of P into sets $S_1, S_2, \cdots, S_{2n-1}$, each containing n elements, such that for each i, $i = 1, 2, \cdots, 2n-1$, the coordinates of the n pairs in S_i constitute the integers $1, 2, \cdots, 2n$.

Proof. Let j be an integer such that $1 \leqq j \leqq 2n - 1$. We define

$$T_j = \{(a, b) \mid 1 \leqq a < b \leqq j + 1 \text{ and } a + b = j + 2\}$$

and

$$R_j = \{(a, b) \mid j + 1 < a < b < 2n \text{ and } a + b = j + 2n + 1\} .$$

Let $S_{2n-1} = T_{2n-1}$. For j even, $1 \leqq j \leqq 2n - 2$, let

$$S_j = T_j \cup R_j \cup \left\{\left(\frac{j+2}{2}, 2n\right)\right\} .$$

For j odd, $1 \leqq j \leqq 2n - 3$, let

$$S_j = T_j \cup R_j \cup \left\{\left(\frac{j+1+2n}{2}, 2n\right)\right\} .$$

It may be verified that the sets S_j have the desired properties.

REISS THEOREM. *Let m be odd and let S be an m-copt for which $C(S) = \varphi(m)$. Then there exists a $(2m+1)$-copt T such that $T \supset S$ and $C(T) = \varphi(2m+1)$.*

Proof. Let $P = \{(u, v) \mid m < u < v \leqq 2m + 1\}$. We use the Reiss lemma to partition P into sets $S_1, \cdots, S_m$, each containing $(m+1)/2$ elements, such that for each i, $i = 1, 2, \cdots, m$, the coordinates of the $(m+1)/2$ pairs in S_i constitute the integers $m+1, m+2, \cdots, 2m+1$. We now define

$$T = S \cup \{(i, j, k) \mid 1 \leqq i \leqq m \text{ and } (j, k) \in S_i\} .$$

It is easily verified that T is a $(2m+1)$-copt. If $m \equiv 1$ or $m \equiv 3 \pmod 6$, then

$$C(S) = \frac{m(m-1)}{6} + \frac{m(m+1)}{2} = \frac{4m^2 + 2m}{6} = \frac{(2m+1)(2m)}{6} = \varphi(2m+1) .$$

If $m \equiv 5 \pmod 6$, then

$$C(S) = \frac{m^2 - m + 1}{6} + \frac{m(m+1)}{2} = \frac{4m^2 + 2m + 4}{6}$$
$$= \frac{(2m+1)^2 - (2m+1) + 4}{6} = \varphi(2m+1) .$$

THEOREM 5. *Let n be an even integer and let S be an n-copt for which $C(S) = \varphi(n)$. Then there exists a $2n$-copt T such that $C(T) = \varphi(2n)$ and $S \subset T$.*

Proof. According to the Reiss Lemma there exists a partition of the set

$$P = \{(u, v) \mid n + 1 \leqq u < v \leqq 2n\}$$

into $n - 1$ sets $A_1, A_2, \cdots, A_{n-1}$ such that for each $i, i = 1, 2, \cdots, n - 1$, the coordinates of the $n/2$ pairs in A_i constitute the integers $\{n + 1, \cdots, 2n\}$. Let $A_n = A_{n-1}$, and let

$$T = S \cup \{(i, j, k) \mid i = 1, 2, \cdots, n ; \ (j, k) \in A_i\} .$$

It is easy to prove that T satisfies the desired conditions.

REFERENCES

1. Marshall Hall, *An existence theorem for Latin squares*, Bull. Amer. Math Soc. **51** (1945), 387.

2. M. Reiss, *Ueber eine Steinersche combinatorische Aufgabe welche im* 45 *sten* *Bande dieses Journal, Seite* 181, *gestellt worden ist.* J. reine angew. Math. **56** (1859), 326–344.

UNIVERSITY OF GEORGIA AND YALE UNIVERSITY

ON THE THEORY OF (m, n)-COMPACT TOPOLOGICAL SPACES

I. S. GÁL

In a recent paper I introduced the following generalization of the notion of compactness:

A topological space X is (m, n)-compact if from every open covering $\{O_i\}$ $(i \in I)$ of X whose cardinality card I is at most n one can select a subcovering $\{O_{i_j}\}$ $(j \in J)$ of X whose cardinality card *J is at most m.*

A similar definition was introduced earlier by P. Alexandroff and P. Urysohn [1]. If no inaccessible cardinals exist between m and n the two definitions are equivalent. The present definition has the advantage that in applications the question of the existence of inaccessible cardinals does not generally come up. The basic results on (m, n)-compact spaces were published by me in [8] and a detailed study of generalized compactness in the Alexandroff-Urysohn sense was made by Yu. M. Smirnov in [14] and [15]. The special case $m = \omega$ and $n = \infty$ was first studied much earlier by C. Kuratowski and W. Sierpinski in [13] and [10]. These spaces are generally known as Lindelöf spaces.

The present paper contains four types of results on (m, n)-compactness which were obtained since the publication of [8]. The problems and the principal results are stated in the beginnings of the individual Sections 1, 2, 3, and 4.

The following notations will be used: $\overline{A}$ and A^i denote the closure and the interior of the set A. The symbols O and C stand for open and closed sets, respectively. ϕ denotes the empty set. N_x is an arbitrary neighborhood of the point x and O_x denotes any open set containing x. Filters are denoted by $\mathscr{F}$, nets by (x_d) $(d \in D)$ where D stands for the directed set on which the net is formed. The set of adherence points of $\mathscr{F}$ is denoted by adh $\mathscr{F}$. Similarly the set of adherence points of a net is denoted by adh(x_d). The set of limit points is denoted by $\lim \mathscr{F}$ and $\lim (x_d)$, respectively. A topological space X is called normal if for any pair of disjoint closed sets A and B there exist disjoint open sets O_A and O_B such that $A \subseteqq O_A$ and $B \subseteqq O_B$.

Uniform structures for a set X will be denoted by $\mathscr{U}$. The symbol $U[x]$ stands for "the vicinity $U \in \mathscr{U}$ evaluated at $x \in X$" so that $U[x] = [y : (x, y) \in U]$. The composition operator is denoted by $\circ$ and so $U \circ V$ consists of those ordered pairs $(x, z) \in X \times X$ for which there is a $y \in X$ with $(x, y) \in U$ and $(y, z) \in V$.

Received April 8, 1958.

1. Characterization of (m, n)-compactness by filters and nets. A topological space X is compact if and only if every filter $\mathscr{F}$ in X has a non-void adherence. A similar characterization of compactness can be given also in terms of nets (x_d) $(d \in D)$ with values in X. As a matter of fact it is sufficient to prove only one of these propositions. For one can associate with every filter $\mathscr{F}$ in X a net $(x_d)\,(d \in D)$ with values in X such that adh $\mathscr{F}$ = adh (x_d) and lim $\mathscr{F}$ = lim (x_d) and conversely given any net with values in X there is a filter $\mathscr{F}$ in X having the same adherence and limit as $(x_d)\,(d \in D)$. The equivalence of filters and nets relative to adherence properties is due to R. G. Bartle [3] and the equivalence relative to both adherence and limit properties is discussed in [9].

It is natural to ask whether (m, n)-compactness can be characterized in term of filters and nets. We shall prove here that such characterization can be given both in terms of filters and nets. Namely for every pair of cardinals $m < n$ a class of filters called (m, n)-filters can be selected such that X is (m, n)-compact if and only if each of these filters has adherence points. Similarly we can define the class of (m, n)-nets with values in X such that X is (m, n)-compact if and only if adh (x_d) is not void for every one of these nets $(x_d)\,(d \in D)$. This indicates that there is a natural correspondence between the class of (m, n)-filters and the class of (m, n)-nets and one can expect that these two classes exhibit the same adherence and limit phenomena. However it will be seen that this is not the case. Hence if we consider filters and nets in a topological space X not as whole classes but in their finer classification then their behavior relative to convergence is not the same.

In the next definition we use the concept of "m-intersection property". A family $\{C_i\}(i \in I)$ of subsets of a set X is said to have the *m-intersection property* if every subfamily of cardinality at most m has a non-void intersection. If every finite subfamily of $\{C_i\}(i \in I)$ has a non-void intersection we say that the family has the *finite intersection property* or *1-intersection property*.

DEFINITION 1.1. A filter $\mathscr{F}$ is called an (m, n)-filter if it has the m-intersection property and if it has a base $\mathscr{B}$ of cardinality card $\mathscr{B} \leqq n$.

If $\mathscr{F}$ is a filter which has a base of cardinality at most n then $\mathscr{F}$ is called an n-filter or an $(1, n)$-filter. If the filter $\mathscr{F}$ has the m-intersection property we say that $\mathscr{F}$ is an (m, ∞)-filter. A $(1, \infty)$-filter means a filter in the usual sense.

DEFINITION 1.2. A directed set D is called an (m, n)-directed set if every subset $S \subseteqq D$ of cardinality card $S \leqq m$ has an upper bound in

D and if card $D \leqq n$.

If every subset $S \subseteqq D$ of cardinality card $S \leqq m$ has an upper bound in D or in other words if for every S with card $S \leqq m$ there is a $d \in D$ such that $s \leqq d$ for every $s \in S$ then D will be called an m-directed set or an (m, ∞)-directed set. If card $D \leqq n$ we speak about a $(1, n)$-directed set. A $(1, \infty)$-directed set means a directed set in the usual sense.

DEFINITION 1.3. An (m, n)-net $(x_d)(d \in D)$ with values in a set X is a function x defined on an (m, n)-directed set D whose function values x_d belong to the set X.

If the directed set D is linearly ordered we call $(x_d)(d \in D)$ a linearly ordered (m, n)-net.

It is known that filters and nets exhibit the same convergence and adherence phenomena. The following lemmas show that the same holds for the more restricted class of (m, ∞)-filters and (m, ∞)-nets:

LEMMA 1.1. *Let X be a topological space and let $(x_d)(d \in D)$ be an (m, n)-net in X. Then there exists an (m, n)-filter $\mathscr{F}$ in X having the property that* adh $\mathscr{F}$ = adh (x_d) *and* lim $\mathscr{F}$ = lim (x_d).

Proof. For every $d \in D$ we define $B_d = [x_\delta : d \leqq \delta]$. Since D is an (m, n)-directed set the family $\mathscr{B} = \{B_d\}(d \in D)$ has the m-intersection property and card $\mathscr{B} \leqq n$. Let $\mathscr{F}$ be the (m, n)-filter generated by the filter base $\mathscr{B}$. One shows that $\mathscr{F}$ satisfies the requirements.

LEMMA 1.2. *Let $\mathscr{F}$ be an (m, ∞)-filter in a topological space X. Then there is an (m, ∞)-net $(x_d)(d \in D)$ with values in X and having the property that* adh (x_d) = adh $\mathscr{F}$ *and* lim (x_d) = lim $\mathscr{F}$.

Proof. Let us consider the set D of all ordered pairs $d = (x, F)$ where $x \in F \in \mathscr{F}$. We say that $d_1 \leqq d_2$ if $F_1 \supseteqq F_2$. Under this ordering D becomes an (m, ∞)-directed set. In fact if $d_i = (x_i, F_i)$ for $i \in I$ and card $I \leqq m$ then $d_i \leqq d$ for every $d = (x, F)$ where $x \in F = \bigcap F_i \in \mathscr{F}$. An (m, ∞)-net can be defined on D with values in X by choosing $x_d = x$ for every $d = (x, F) \in D$. Let $x \in \lim \mathscr{F}$ and let N_x be arbitrary. Then there is an $F \in \mathscr{F}$ such that $F \subseteqq N_x$. Hence if $\delta = (\xi, \Phi)$ satisfies $d \leqq \delta$, or in other words if $\Phi \subseteqq F$ then $x_\delta = \xi \in \Phi \subseteqq F \subseteqq N_x$ and so x is a limit point of $(x_d)(d \in D)$. Conversely let $x \in \lim (x_d)$ and let N_x be given. Then there is a $d = (x, F)$ such that $x_\delta \in N_x$ for every δ satisfying $d \leqq \delta$. Using this for every $\delta = (\xi, F)(\xi \in F)$ we see that

$x_\delta = \xi \in N_x$ for every $\xi \in F$ and so $F \subseteqq N_x$. This shows that $x \in \lim \mathscr{F}$ and $\lim (x_d) = \lim \mathscr{F}$.

Now we suppose that $x \in \operatorname{adh} \mathscr{F}$ so that $N_x \cap F \neq \phi$ for every neighborhood N_x and for every $F \in \mathscr{F}$. Given N_x and $d = (x, F)$ we choose ξ in $N_x \cap F$ and consider $\delta = (\xi, F)$. Then $d \leqq \delta$ and $x_\delta = \xi \in N_x$ and so $x \in \operatorname{adh} (x_d)$. On the other hand if $x \in \operatorname{adh} (x_d)$ then given $F \in \mathscr{F}$ and N_x there is a $\delta \in D$ such that $d = (x, F) \leqq \delta = (\xi, \Phi)$ and $x_\delta \in N_x$. In other words $x_\delta = \xi \in \Phi \cap N_x \subseteqq F \cap N_x$ and so F and N_x intersect for every $F \in \mathscr{F}$ and for every $N_x \in \mathscr{N}(x)$. This proves that $x \in \operatorname{adh} \mathscr{F}$ and $\operatorname{adh} (x_d) = \operatorname{adh} \mathscr{F}$.

Using the same reasoning similar results can be derived for (m, n)-filters. For instance we can easily prove that if X is an (m, n)-filter in a space X and if card $F \leqq n$ for some $F \in \mathscr{F}$ then there is an (m, n)-net $(x_d)(d \in D)$ with values in X and having the property that $\operatorname{adh} (x_d) = \operatorname{adh} \mathscr{F}$ and $\lim (x_d) = \lim \mathscr{F}$. If the hypothesis card $F \leqq n$ is dropped we can find only an (m, n)-net satisfying $\operatorname{adh} \mathscr{F} \supseteqq \operatorname{adh} (x_d)$ and $\lim \mathscr{F} \subseteqq \lim (x_d)$. None of these results will be used in the sequel.

We can easily find examples where only the strict inclusion $\operatorname{adh} \mathscr{F} \supset \operatorname{adh} (x_d)$ can be realized. For instance let X be a non-countable set and let X be topologized by the discrete topology. If $\mathscr{F}$ consists of the single element X then $\mathscr{F}$ is an (m, n)-filter for any pair of cardinals m and n. Moreover $\operatorname{adh} \mathscr{F} = X$ and so the cardinality of $\operatorname{adh} \mathscr{F}$ is greater than that of $\mathscr{F}$. On the other hand if $(x_d)(d \in D)$ is a $(1, \omega)$-net with values in X then the cardinality of $\operatorname{adh} (x_d)$ is at most ω. Hence $\operatorname{adh} \mathscr{F} \supset \operatorname{adh} (x_d)$ for every $(1, \omega)$-net in X.

This example shows that (m, n)-filters and (m, n)-nets in arbitrary topological spaces have different adherence properties. Nevertheless the following theorems show that both (m, n)-filters and (m, n)-nets can be used to characterize (m, n)-compactness.

THEOREM 1.1. *A topological space X is (m, n)-compact if and only if every (m, n)-filter in X has a non-void adherence.*

Proof. In [8] we proved that X is (m, n)-compact if and only if every family $\{C_i\}$ of closed sets $C_i \subseteqq X$ having the m-intersection property also has the n-intersection property. We apply this result: Let X be (m, n)-compact and let $\mathscr{B}$ with card $\mathscr{B} \leqq n$ be a filter base for an (m, n)-filter $\mathscr{F}$ in X. Then the family $\{\overline{B}\}$ $(B \in \mathscr{B})$ has the m-intersection property and so it has the n-intersection property. Since card $\mathscr{B} \leqq n$ this implies that $\cap \overline{B} = \operatorname{adh} \mathscr{B}$ is not void. Conversely if X is not (m, n)-compact then there is a family $\mathscr{B}$ of closed sets with card $\mathscr{B} \leqq n$ and having the m-intersection property but with total intersection void. Thus $\mathscr{B}$ is a filter base for an (m, n)-filter $\mathscr{F}$ and $\operatorname{adh} \mathscr{F} = \Phi$.

THEOREM 1.2. *A topological space X is (m, n)-compact if and only if every (m, n)-net with values in X has a non-void adherence.*

Proof. If there is an (m, n)-net with values in X whose adherence is void then by Lemma 1.1 there is an (m, n)-filter without adherence points and so by Theorem 1.1 the space X is not (m, n)-compact. Next we prove that if every (m, n)-net with values in X has a non-void adherence then the same is true for every (m, n)-filter in X. By Theorem 1.1 this will prove that X is (m, n)-compact. Let $\mathscr{B} = \{B_d\}(d \in D)$ be a filter base for an (m, n)-filter $\mathscr{F}$ in X and let card $D \leqq n$. We order D by using inverse inclusion of $\mathscr{B}$: $d_1 \leqq d_2$ if $B_{d_1} \supseteqq B_{d_2}$. Under this ordering D becomes an (m, n)-directed set. We form a net $(x_d)(d \in D)$ by choosing x_d in $\overline{B}_d$. By hypothesis $(x_d)(d \in D)$ has an adherence point x. Given any neighborhood N_x and any $d \in D$ there is a $\delta \geqq d$ such that $x_\delta \in N_x$. Hence $N_x \cap B_\delta \neq \phi$ and so by $B_\delta \subseteqq B_d$ also $N_x \cap B_d \neq \phi$. Consequently $x \in \overline{B}_d$ for every $d \in D$ and so $x \in \operatorname{adh} \mathscr{F}$.

2. Uniformizability and (m, n)-compactness. This section contains the generalization to infinite cardinals of the following results:

A space X is countably compact if and only if every infinite set $S \subseteqq X$ has an accumulation point in X.

If X is a metric space such that every infinite set $S \subseteqq X$ has an accumulation point in X then the open sets of X have a countable base and so X is a Lindelöf space.

Countable compactness will be replaced by (m, m')-compactness where m is an infinite cardinal and m' denotes the first cardinal succeeding m. If m denotes the symbol 1 then $1'$ is defined to be ω. Instead of accumulation points we must consider m-accumulation points:

DEFINITION 2.1. *A point x of a topological space X is called an m-accumulation point of a set $S \subseteqq X$ if for every open set O_x containing x we have* card $(O_x \cap S) > m$.

If m is 0, 1 or ω then the relation card $(O_x \cap S) > m$ means that $O_x \cap S$ is not void, not finite or not countable. If card $S \leqq m$ the set of its m-accumulation points is void. In particular if S is countable then it has no ω-accumulation points and if S is finite then it has no 1-accumulation points. The notion of an m-accumulation point is related to Fréchet's "point d'accumulation maximé" (see [7]).

The metrizability condition can be rephrased as follows: There is a uniform structure $\mathscr{U}$ which is compatible with the topology of X and has a countable structure base. This hypothesis will be replaced by another which requires the existence of a structure base of cardinality at most m.

DEFINITION 2.2. A uniformizable space X is said to be of uniform cardinality u if there is a base $\mathscr{U}_B$ for a uniform structure $\mathscr{U}$ compatible with the topology of X whose cardinality card $\mathscr{U}_B$ is at most u.

Every pseudo-metric space is of uniform cardinality ω. If for every uniform structure $\mathscr{U}$ compatible with the topology of X and for every base $\mathscr{U}_B$ of $\mathscr{U}$ we have card $\mathscr{U}_B \geqq u$ where u is a uniform cardinality of X then we say that X is of uniform cardinality exactly u. The exact uniform cardinality of a pseudo-metric-space is at most ω; it can also be 1.

The first result which we mentioned in the beginning is a special case of

THEOREM 2.1. *Let m be an infinite cardinal and let m' denote the next cardinal. Then a topological space X is (m, m')-compact if and only if every set $S \subseteqq X$ of cardinality* card $S > m$ *has an m-accumulation point in X.*

Proof. First we prove the necessity of the condition. If X contains sets of cardinality greater than m which have no m-accumulation points in X then we can select a set S of cardinality exactly m' such that it has no m-accumulation points in X. Let $\mathscr{F}$ denote the set of all those sets $F \subseteqq S$ whose cardinality is at most m. For every $x \in X$ there is an open set O_x such that $O_x \cap S \in \mathscr{F}$. We define O_F for every $F \in \mathscr{F}$ as $O_F = \cup [O_x : O_x \cap S = F]$. The family $\{O_F\}(F \in \mathscr{F})$ is an open cover of X whose cardinality card $\mathscr{F}$ is m'. Since every subfamily of cardinality at most m would cover at most m points of S the family $\{O_F\}(F \in \mathscr{F})$ cannot contain such subfamilies. Hence X is not (m, m')-compact.

The deeper part of the theorem is the sufficiency of the condition. Here we need the axiom of choice both in the form of Zorn's lemma and also in the form of the well-ordering theorem. Suppose that X is not (m, m')-compact. Let $\{O_i\}(i \in I)$ be an open cover of cardinality card $I = m'$ which contains no subcovers of cardinality at most m. Let the index set I be well ordered. Since X is not (m, m')-compact there is a point $x_2 \in X$ such that $x_2 \notin O_1$. More generally for every positive integer $n > 1$ there are points $x_2, \cdots, x_n$ such that $x_j \notin O_1 \cup \cdots \cup O_{j-1}$ for every $j \leqq n$. In general we consider segments J of I such that a segment (or net) of points (x_j) can be selected so that $x_j \notin \cup [O_j : i < j]$ for every $j \in J$. Let $\mathscr{A}$ denote the family of ordered pairs $(J, (x_j))$ where J denotes a segment of I and (x_j) a segment of points associated with J. We order $\mathscr{A}$ as follows: $(J_1, (x_j)) \leqq (J_2, (y_j))$ if $J_1 \subseteqq J_2$ and $x_j = y_j$ for every $j \in J_1$. Clearly every linearly ordered subfamily of $\mathscr{A}$ has an upper bound in $\mathscr{A}$ and so by Zorn's lemma $\mathscr{A}$ has a maximal element, say $(J, (x_j))$. The maximality of $(J, (x_j))$ implies that J is a limit ordinal and $\{O_j\}(j \in J)$ is a cover of X. We claim that the set $S = \{x_j\}(j \in J)$ has no m-accumulation points in X. For if $x \in X$ then $x \in O_j$ for some $j \in J$ and so

$$\operatorname{card}(O_j \cap S) \leqq \operatorname{card}[x_i : i < j] \leqq m .$$

Finally $\operatorname{card} S = m'$ because $\operatorname{card} S = \operatorname{card} J$ and $\{O_j\}(j \in J)$ is a cover of the not (m, m')-compact space X.

THEOREM 2.2. *If the uniformizable space X is of uniform cardinality u and if there is an m such that every set $S \subseteqq X$ of cardinality* $\operatorname{card} X > m$ *has a non-void derived set then the open sets of X have a base of cardinality at most* $\max(m, u)$.

Proof. Let $\mathscr{U}_B = \{U\}$ be a base of a uniform structure $\mathscr{U}$ for X and let $\operatorname{card} \mathscr{U}_B \leqq u$. We may suppose that every $U \in \mathscr{U}_B$ is symmetric. We fix a vicinity $U \in \mathscr{U}_B$ and consider systems of points $\{x_i\}$ $(i \in I)$ having the property that $U[x_i] \cap U[x_j]$ is void for every $i \neq j$. Let $\mathscr{A}$ be the set of all such systems $\{x_i\}(i \in I)$. The set $\mathscr{A}$ is not void for such systems exist at least in the case when the index set I consists of a single element. We order $\mathscr{A}$ by inclusion : $\{x_i\} \leqq \{y_j\}$ if $\{x_i\} \subseteqq \{y_j\}$. Every linearly ordered subset $\mathscr{L}$ of $\mathscr{A}$ has an upper bound in $\mathscr{A}$, namely $\cup [\{x_i\} : \{x_i\} \in \mathscr{L}]$ is in $\mathscr{A}$ and it majorizes every $\{x_i\} \in \mathscr{A}$. Hence Zorn's lemma can be applied to show the existence of a maximal system which we denote by $\{x_i\}(i \in I)$. If $y \in \cup [U[x_i] : i \in I]$ then by the maximality $U[y] \cap U[x_i]$ is non-void for some $i \in I$. Hence by the symmetry of U we have $y \in (U \circ U)[x_i]$. Therefore the family $\{(U \circ U)[x_i]\}(i \in I)$ is a cover of the uniform space X.

Let $S = \{x_i\}(i \in I)$ so that $\operatorname{card} S = \operatorname{card} I$. We show that the derived set of S is void and so $\operatorname{card} I \leqq m$: Let V be a symmetric vicinity in $\mathscr{U}$ such that $V \circ V \subseteqq U$ and let x be an arbitrary point in X. If $x \in V[x_i]$ for some $i \in I$ then $V[x] \subseteqq (V \circ V)[x_i] \subseteqq U[x_i]$ and so $V[x] \cap S$ is void or contains at most the point x_i. If $x \notin V[x_i]$ for every $i \in I$ then by the symmetry $x_i \notin V[x]$ for every $i \in I$ and so $V[x] \cap S$ is void. It follows that $\operatorname{card} I \leqq m$.

The family $\{(U \circ U)[x_i]\}(i \in I)$ is a cover of X and so the interiors of the sets $(U \circ U \circ U)[x_i](i \in I)$ form an open cover of X. Its cardinality is at most m. Hence the cardinality of the union of these families for every choice of $U \in \mathscr{U}_B$ is of cardinality at most $\max(m, u)$. Since for every vicinity $V \in \mathscr{U}$ there is a $U \in \mathscr{U}_B$ such that $U \circ U \circ U \subseteqq V$ these sets form a base for the open sets of X.

The results of this section can be combined to obtain the following

THEOREM 2.3. *If X is a uniformizable space of uniform cardinality u which is (m, n)-compact for some cardinals m and n where $m < n$ then X is (m, ∞)-compact or (u, ∞)-compact according as $m \geqq u$ or $u > m$.*

Proof. By Theorem 2.1 every set S of cardinality greater than m has an m-accumulation point and so its derived set is not void. Theorem 2.2 implies the existence of a base of cardinality at most $\max(m, u)$ for the family of open sets of X. Hence the space X is $(\max(m, u), \infty)$-compact.

This proof did not make use of the full force of Theorem 2.1. It is sufficient to know for instance that every set of cardinality $\operatorname{card} S > m$ has a 1-accumulation point whenever the space X is (m, n)-compact for some $n > m$. This weaker statement can be proved without using the axiom of choice or the well ordering theorem. Nevertheless the axiom of choice is used in the proof of Theorem 2.2.

3. **Dense sets, (m, n)-compact spaces and complete structures.** It is knownthat if X is a compact topological space then every net with values in X has a non-void adherence and conversely if the adherence of every net with values in X is not void then X is compact. We can raise the following question: Suppose A is a dense subset of X and that $\operatorname{adh}(x_d)$ is not void for every net $(x_d)(d \in D)$ with values in A. Does it follow that X is compact? We shall prove a theorem a special case of which states that for regular spaces the answer is affirmative. The result can be formulated also in terms of filters: Every filter $\mathscr{F}$ in A is a filter base in X. If the adherence af the filter generated by the base $\mathscr{F}$ is not void we say that the filter $\mathscr{F}$ has a non-void adherence in X. It was proved earlier that if X is regular and if every filter in the dense set A has a non-void adherence in X then X is compact. (See [4] p. 109 Ex. 1 a.)

The same type of question can be raised when the net $(x_d)(d \in D)$ is subject to additional restrictions: For instance we can assume that every countable net with values in A has a non-void adherence in X and ask whether this implies that X is countably compact. It will be proved that the conclusion holds under the assumption of normality and countable compactness.

As is known a family $\mathscr{L}$ of sets $S_i \subseteqq X$ is called a locally finite system if every $x \in X$ has a neighborhood N_x which meets only finitely many sets of the family $\mathscr{L}$. We shall deal only with locally finite systems which consist of open sets.

DEFINITION 3.1. A topological space X is called n-paracompact if every open cover $\{O_i\}(i \in I)$ satisfying $\operatorname{card} I \leqq n$ admits a refinement $\{Q_j\}(j \in J)$ which is a locally finite system.

Clearly every topological space is 1-paracompact and we agree that ∞-paracompactness means paracompactness in the usual sense. Using this definition we can state the following

THEOREM 3.1. *Let X be a normal n-paracompact space which contains a dense set A such that every (m, m')-net with values in A (or every (m, m')-filter in A) has a non-void adherence in X. Then X is (m, n)-compact.*

Since every regular paracompact space is normal in the special case when $n = +\infty$ normality can be replaced by the formally weaker requirement of regularity. However this is not a real improvement of the result. If $m = 1$ then by our agreement $m' = \omega$ and so if X is countably compact then every (m, m')-net with values in X has a non-void adherence. Hence as a corollary we have the following result due to R. Arens and J. Dugundji [2]:

COROLLARY. *If X is regular, paracompact and countably compact then X is compact.*

Since every pseudo-metric space is paracompact (see [17]) the corollary is a generalization of the following known result: If the pseudo-metric space X is countably compact then it is compact. A weaker form of the corollary was obtained by Miss A. Dickinson who proved in [5] that every uniformizable space with a unique structure is countably compact and a paracompact space with a unique structure is compact.

In the proof of the theorem we shall use the following known lemmas:

LEMMA 3.1. *If $\{S_i\}(i \in I)$ is a locally finite system of sets then $\overline{\cup S_i} = \cup \overline{S_i}$.*

A short proof can be found for instance in [16].

LEMMA 3.2. *Let $\{O_i\}(i \in I)$ be a locally finite open cover of the normal space X. Then there is an open cover $\{Q_i\}(i \in I)$ of X such that $\overline{Q}_i \subseteqq O_i$ for every $i \in I$.*

Proofs of this lemma can be found in [12], [9], [6] or [11],

LEMMA 3.3. *If* $\{\omega_i\}(i \in I)$ *is a set of ordinals such that* $\omega_i < m'$ *for every* $i \in I$ *and if* $\operatorname{card} I \leqq m$ *then* $\operatorname{lub}\{\omega_i\} < m'$.

Proof. Since m' is the first ordinal of cardinality m' we have $\omega_i < m'$, that is, $\operatorname{card}\omega_i \leqq m$ for every $i \in I$. Hence by $\operatorname{card} I \leqq m$ the cardinality of $\operatorname{lub}\{\omega_i\}$ is m.

Proof of Theorem 3.1. We assume that X is normal and n-paracompact but it is not (m, n)-compact. We shall construct a linear (m, m')-net $(x_d)(d \in D)$ with values in A and such that it has no adherence points in X. Then the sets $[x_\delta : d \leqq \delta](d \in D)$ form a filter base for an (m, m') filter in A which has no adherence points in X.

Let $\{O_i\}(i \in I)$ be an open cover of cardinality at most n which contains no subcover of cardinality at most m. Since X is n-paracompact $\{O_i\}(i \in I)$ admits a refinement $\{Q_j\}(j \in J)$ which is a locally finite system of open sets. The space being normal by Lemma 3.2 we may assume that $\overline{Q}_j \subseteqq O_i$ for every $j \in J$ and for a suitable $i = i(j) \in I$. Since $\{Q_j\}(j \in J)$ is locally finite Lemma 3.1. can be applied to any subfamily of this cover.

Let the index set J be well ordered; for the sake of simplicity we assume that the elements of J are ordinals. Denote by S_k the open set

$$S_k = Q_k - \bigcup[\overline{Q}_j : j < k] .$$

Let D be the set of those indices $k \in J$ for which S_k is not void. We prove that $\kappa = \operatorname{card} D \geqq m'$.

For let $\mathscr{C}$ be the class of those initial segments $K \subseteqq J$ for which $\bigcup[\overline{Q}_j : j \in K] \subseteqq \bigcup[\overline{Q}_j : j \in D]$. Then $\mathscr{C}$ is not void because $(1, \cdots, \kappa) \in \mathscr{C}$. It can be ordered by inclusion: $K_1 \leqq K_2$ if $K_1 \subseteqq K_2$. There is a maximal element in $\mathscr{C}$ namely $K_m = \bigcup[K : K \in \mathscr{C}]$ itself is an element of $\mathscr{C}$. We prove that $K_m = J$. For let $K \in \mathscr{C}$ be a proper subset of J which contains κ and let k' be the first index not in K. We set $K' = K \cup \{k'\}$ and obtain by $\kappa < k'$

$$\begin{aligned}\bigcup[\overline{Q}_j : j \in K'] &= \bigcup[\overline{Q}_j : j \in K] \cup \overline{Q}_{k'} = \bigcup[\overline{Q}_j : j \in K] \\ &\subseteqq \bigcup[\overline{Q}_j : j \in D] .\end{aligned}$$

Hence $K' \in \mathscr{C}$ and K is not maximal. Consequently $K_m = J$ and this implies that

$$\bigcup[Q_j : j \in J] \subseteqq \bigcup[\overline{Q}_j : j \in D] .$$

However on the one hand $\{Q_j\}(j \in J)$ is a cover of X and so $X =$

$\cup[\overline{Q}_j : j \in D]$. On the other hand $\{\overline{Q}_j\}(j \in J)$ is a refinement of $\{O_i\}$ $(i \in I)$ and by hypothesis $\{O_i\}(i \in I)$ does not contain a subcover of cardinality at most m. Hence we have card $D \geqq m'$.

The well-ordered set D is order isomorphic to an initial segment of the ordinals which segment contains every ordinal preceding m'. If we discard from D every element corresponding to m' and to the ordinals succeeding m' we obtain a subset of D of cardinality at least m'. We denote this subset again by D. By Lemma 3.3 the new D is an (m, m')-directed set.

The open sets S_d are not void for every $d \in D$ and A is dense in X. Hence we can choose a point $a_d \in A$ in each of the sets $S_d(d \in D)$. The linear net $(a_d)(d \in D)$ is an (m, m')-net with values in A and it has no adherence points in X: In fact $\{Q_j\}(j \in J)$ being locally finite for every point $x \in X$ we can find an open set O_x such that $O_x \cap Q_d$ is not void only for finitely many indices $d \in D$. If d is larger than any of these finitely many indices then $O_x \cap Q_\delta = \phi$ for every $\delta \geqq d$ and so $a_\delta \notin O_x$ for every $\delta \geqq d$. This, however, shows that x is not an adherence point of the net $(a_d)(d \in D)$. This completes the proof of Theorem 3.1.

Now we turn to uniformizable spaces:

THEOREM 3.2. *Let X be a uniformizable space of uniform cardinality u. Suppose that X contains a dense subset A such that every (m, n)-filter in A has a non-void adherence in X. Then X is (m, ∞)-compact.*

It is sufficient to prove that X is (m, n)-compact. The (m, ∞)-compactness follows from Theorem 2.3. The proof of the (m, u)-compactness can be modified such that we obtain the following known result (see [7] p. 150, Proposition 7):

Let A be a dense subset of a uniform space X with uniform structure $\mathscr{U}$. If every Cauchy filter in A is convergent to some point of X then the structure $\mathscr{U}$ is complete.

Proof. Let $\mathscr{F}$ be a filter (or an (m, u)-filter) in X. Consider the family $\mathscr{B} = \{U[F] \cap A\}$ $(U \in \mathscr{U}$ and $F \in \mathscr{F})$. Since A is dense in X every set $U[F] \cap A$ is non-void and

$$(U[F_1] \cap A) \cap (U[F_2] \cap A) \supseteqq U[F_1 \cap F_2] \cap A .$$

Hence $\mathscr{B}$ is a filter base in X. (Moreover if $\mathscr{F}$ is an (m, u)-filter and $\mathscr{U}$ is of uniform cardinality u then $\mathscr{B}$ is a base for an (m, u)-filter in A.) If $\mathscr{F}$ is a Cauchy filter then $\mathscr{B}$ is a base for a Cauchy filter because if $F \times F \subseteqq V$ where V is symmetric then $V[F] \times V[F] \subseteqq$

$V \circ V \circ V$: In fact if $x \in V[F]$ and $y \in V[F]$ then $(x, a) \in V$ and $(b, y) \in V$ for some $a, b \in F$. Thus by $(a, b) \in F \times F \subseteqq V$ we have $(x, y) = (x, a) \circ (a, b) \circ (b, y) \in V \circ V \circ V$. By hypothesis $\mathscr{B}$ as a Cauchy filter base in X (as a base for an (m, u)-filter in X) is convergent to some point $x \in X$. We show that $x \in \operatorname{adh} \mathscr{F}$ which is equivalent of saying that $x \in \lim \mathscr{F}$. Given any an open neighborhood O_x of x there is a $U \in \mathscr{U}$ such that $U[x] \subseteqq O_x$. We determine the symmetric $V \in \mathscr{U}$ such that $V \circ V \subseteqq U$. Since $x \in \operatorname{adh} \mathscr{B}$ and $V[F] \cap A$ is an element of $\mathscr{B}$ we have $V[x] \cap V[F] \neq \phi$ for every $F \in \mathscr{F}$. Hence there is an $a \in A$ and an $f \in F$ such that $(x, a) \in V$ and $(a, f) \in V$. Therefore $(x, f) \in V \circ V \subseteqq U$ and $f \in U[x] \cap F$. This shows that $U[x]$ and F intersect for every $U \in \mathscr{U}$ and for every $F \in \mathscr{F}$. Thus $x \in \operatorname{adh} \mathscr{F}$.

4. **Additional results and notes.** In my first paper on (m, n)-compactness I introduced the notion of a hereditary or completely (m, n)-compact topological space: X is completely (m, n)-compact if every subspace Y of X is (m, n)-compact. It can be easily proved that if every open set Y is an (m, n)-compact subspace then X is completely (m, n)-compact. In the same paper I gave a number of equivalent characterizations of complete (m, n)-compactness. At that time time I did not notice that one of these criteria (Theorem 4, condition (ii) in [8]) involves n only in a formal way.[1] I should have added as a corollary the following.

THEOREM 4.1. *If X is completely (m, n)-compact for some cardinals $m < n$ then X is completely (m, ∞)-compact.*

Proof. Suppose that X is not completely (m, ∞)-compact. Then there is a family of open sets O_i $(i \in I)$ in X such that $\bigcup O_{i_j}$ is a proper subset of $\bigcup O_i$ whenever $\operatorname{card} J \leqq m$. Let the index set I be well ordered. Let O_{i_1} be the first non-void O_i and let O_{i_2} be the first O_i such that $O_{i_2} \nsubseteq O_{i_1}$. In general we consider initial segments J of the ordinals $1, 2, \cdots, j, \cdots$ and sets $O_{i_j} (j \in J)$ such that for every $j \in J$ the set O_{i_j} is the first O_i set which is not a subset of $\bigcup [O_{i_k} : k < j]$. By hypothesis $\{O_i\}(i \in I)$ does not admit a subfamily $\{O_{i_j}\}(j \in J)$ satisfying $\bigcup O_i = \bigcup O_{i_j}$ with $\operatorname{card} J \leqq m$. Hence using Zorn's lemma we can find initial segments J and corresponding sets O_{i_j} such that $\operatorname{card} J \geqq m'$ where m' is the first cardinal greater than m. We restrict ourselves to ordinals preceding m' so that $J = [j : j < m']$ and $O_{i_j} \nsubseteq \bigcup [O_{i_k} : k < j]$ for every $j \in J$. The family $\{O_{i_j}\}(j \in J)$ is of cardinality $\operatorname{card} J = m'$ and if $\operatorname{card} K < m'$ where $K \subset J$ then by Lemma 3.3

[1] This was first noticed by Mr. R. D. Joseph.

$$\cup [O_{i_k} : k \in K] \subset \cup [O_{i_j} : j \in J] .$$

This shows that X is not completely (m, m')-compact and so the theorem is proved.

Let $\mathscr{B} = \{B\}$ be a base for the open sets of a space X and let $Y \subseteqq X$. Then $\mathscr{B}_Y = \{B \cap Y\}(B \in \mathscr{B}\}$ is a base for the subspace Y. Hence if $\mathscr{B}$ is a base for the topology of the space X then every subspace of X is (card $\mathscr{B}$, ∞)-compact. Applying this remark to the situation described in Theorem 2.3 we obtain

THEOREM 4.2. *Let X be a uniformizable space of uniform cardinality u. If X is (m, n)-compact for some $m < n$ then X is completely (m, ∞)-compact or completely (u, ∞)-compact according as $m \geqq u$ or $m \leqq u$.*

If $n = \infty$ this result can be obtained directly by using the definition of (m, ∞)-compactness and of hereditary (m, ∞)-compactness.

The product of a $(1, \infty)$-compact space with an (m, n)-compact space is (m, n)-compact. This was proved a few years ago by Yu. M. Smirnov [14]. Not knowing the existence of this paper, I proved in [8] (Theorem 8) the result in the special case when $n = \infty$, but a slight modification in my reasoning gives a new proof of Smirnov's theorem: Start again by replacing the open cover $\{O_i\}(i \in I)$ where card $I \leqq n$ by a family of sets $O_x^y \times O_y^x$. However instead of forming the intersection $O_y^{x_1} \cap \cdots \cap O_y^{x_m}$ form the intersection of those sets $O_{i_1}^{(y)}, \cdots, O_{i_m}^{(y)}$ of the given family which have the property that

$$O_{x_1}^y \times O_y^{x_1} \subseteqq O_{i_1}^{(y)}, \cdots, O_{x_m}^y \times O_y^{x_m} \subseteqq O_{i_m}^{(y)} .$$

Since card $I \leqq n$ there are at most n distinct ones among the finite intersections $Q_y = O_{i_1}^{(y)} \cap \cdots \cap O_{i_m}^{(y)}$. The rest of the reasoning then is the same as in [8].

We end by stating two unsolved problems: Professor Erdös mentioned to me that he was thinking without success of the following problem: Let m be an infinite cardinal. We say that X is $[m]$-compact if from every open covering of X one can select a subcovering having fewer than m elements. Is there an infinite cardinal m such that the product of any two $[m]$-compact spaces is again $[m]$-compact?

It is known that given any filter $\mathscr{F}$ in a set X there exists an ultrafilter $\mathscr{M}$ such that $\mathscr{F} \subseteqq \mathscr{M}$. Let $\mathscr{F}$ be an (m, ∞)-filter. The corresponding ultrafilter $\mathscr{M}$ need not be an (m, ∞)-filter and in general there is no (m, ∞)-ultrafilter $\mathscr{M}$ satisfying the requirement $\mathscr{F} \subseteqq \mathscr{M}$. We can ask the following question: Is there any infinite cardinal m such that for every (m, ∞)-filter $\mathscr{F}$ the ultrafilter $\mathscr{M}$ can be chosen such that $\mathscr{M}$ is an (m, ∞)-filter and $\mathscr{F} \subseteqq \mathscr{M}$?

I do not know to what extent n-paracompactness is necessary in the hypothesis of Theorem 3.1. The only example that I know of shows that there exists a non-compact space X which contains a dense set A such that every filter in A has a non-void adherence in X: We choose X to be the interval $[-1, 1]$ and call O open if it can be obtained from an open set in the usual sense by omitting points of the form $x = \pm 1, \pm \frac{1}{2}, \cdots$. We can choose $A = X - \{\pm 1, \pm \frac{1}{2}, \cdots\}$. The space X is neither regular nor compact. It can be proved that X is not countably paracompact.

BIBLIOGRAPHY

1. P. Alexandroff, et. P. Urysohn, *Mémoire sur les espaces topologiques compact dédie à Monsieur D. Egoroff*, Verh. Amsterdam **14** No. 1 (1929). A revised second edition of this article is available in Russian: (Trudy Mat. Inst. Steklov **31**, (1950) 95).

2. R. Arens, and J. Dugundji, *Remark on the conpcept of compactness*, Portugaliae Math. (1950) 141–143.

3. R. G. Bartle, *Nets and filters in topology*, Amer. Math. Monthly **62** (1955), 551–557.

4. N. Bourbaki, *Topologie générale*, Chap. I, II; 2nd edition, Hermann & Cie, Editeurs Paris (1951).

5. Alice Dickinson, *Compactness conditions and uniform structures*, Amer. J. Math. **75** (1953), 224-227.

6. J. Dieudonné, *Une généralisation des espaces compacts*, J. Math. Pures Appl. (9) **23** (1944), 65–76.

7. M. Fréchet, *Les espaces abstraits*, Gauthier Villars, Paris (1928).

8. István S. Gál, *On a generalized notion of compactness* I-II, Proc. Nederl. Akad. Wetensch. **60** (1957), 421–435.

9. ———, *Lecture notes on the foundations of analysis and geometry*, Cornell University (1957–1958).

10. C. Kuratowski, et W. Sierpinski, *Le théorème de Borel-Lebesgue dans la théorie des ensembles abstraits*, Fund. Math. **2** (1921), 172–178.

11. Solomon Lefschetz, *Algebraic topology*, Amer. Math. Soc. Coll. Publ. **27** New York (1942).

12. G. Nöbeling, *Grundlagen der analytischen Topologie*, Springer Verlag, Berlin (1954).

13. W. Sierpinski, *Sur l'équivalence de trois propriété des ensembles abstraits*, Fund. Math. **2** (1921), 179–188.

14. Yu. M. Smirnov, *On topological spaces compact in a given interval of powers*, Izvestiya Akad. Nauk SSSR, Ser. Mat. **14** (1950), 155-178.

15. ———, *On the theory of finally compact spaces*, Ukrain. Mat. Žurnal **3** (1951), 52–60.

16. ———, *On metrization of topological spaces*, Uspehi Mat. Nauk (N. S.) **6** (1951) 100–111 and Amer. Math. Soc. Translation No. **91** (1953).

17. A. H. Stone, *Paracompactness and product spaces*, Bull. Amer. Math. Soc. **54** (1948), 977–982.

CORNELL UNIVERSITY,
ITHACA, NEW YORK

A NOTE ON POLYNOMIAL AND SEPARABLE GAMES

DAVID GALE AND OLIVER GROSS

1. Introduction. A two-person zero-sum game Γ is called polynomial-like or *separable* if its payoff function is of the form

$$M(x, y) = \sum_{i=1}^{n} f_i(x) g_i(y),$$

where x and y are elements of any strategy sets X and Y. Important special cases of separable games are those in which X and Y are bounded (usually compact) subsets of Euclidean spaces and M is a polynomial in the coordinates of x and y. These latter are called *polynomial games*.

It is a basic and fairly elementary fact concerning separable games [1], that, if optimal strategies exist, then these can always be chosen to be finite mixed strategies. We consider here the inverse question: Given a pair of finite mixed strategies, does there exist a separable (respectively, polynomial) game whose unique optimal strategies are the given pair? In case either X or Y is *finite* the answer is known to be in the negative. We here show, however, that.

THEOREM 1. *If X and Y are metric spaces containing infinitely many points and μ and ν are any finite mixed strategies on X and Y respectively, then there is a payoff M, bounded continuous and separable on $X \times Y$, such that the associated game has μ and ν as unique optimal strategies.*

COROLLARY. *If X is a metric space containing infinitely many points and μ is any finite mixed strategy on X, then there is a skew-symmetric payoff M, bounded continuous and separable on $X \times X$ such that the associated symmetric game has μ as the unique optimal strategy.*

For the case of polynomial games we show:

THEOREM 2. *If X and Y are bounded subsets of Euclidean spaces whose closures contain infinitely many cluster points, then for any finite mixed strategies μ and ν there exists a polynomial payoff function M such that the associated game has μ and ν as its unique optimal strategies.*

(An analogous corollary holds here, also.)

Concerning Theorem 2, we remark that Glicksberg and Gross, [2],

Received April 30, 1958.

have shown that any pair of mixed strategies can be the unique solution of a continuous game on the unit square. For finite mixtures, however, their construction is complicated, involving consideration of four special cases, and the payoff function is not a polynomial, nor even separable. The rather simple construction involved in our proof of Theorem 2 shows that their result still holds under the much stronger requirement that the payoff be a polynomial.

Finally, we credit Dresher, Karlin and Shapley, [1], for their rather exhaustive study of the structure of solutions of separable and polynomial games. However, their results do not include the theorems proved in this note. Indeed, one of the above authors has pointed out that the construction of the next section provides a *counter-example* to one of the conclusions of a structure theorem in [1], and fortunately (for mathematics) an error in the proof of that part of the theorem[1] was subsequently uncovered.

2. Polynomial games with prescribed unique solutions. This section contains the proof of Theorem 2. Let X and Y be sets satisfying the hypothesis of the theorem (We pause to note that boundedness of X and Y is required to insure integrability, since polynomials may otherwise be unbounded.). Let μ be the mixed strategy which assigns the weight μ_i to the point x_i of X, $i = 1, \cdots, m$, where $\sum \mu_i = 1$. Similarly, let ν assign the weight ν_j to the point y_j in Y, $j = 1, \cdots, n$ where $\sum \nu_j = 1$.

The set of points $\{x_1, \cdots, x_m\}$, the *spectrum* of μ, will be denoted by $\sigma(\mu)$. Similarly, $\sigma(\nu)$ will denote the spectrum of ν.

We now define the following set of polynomials:

$$f_0(x) = \prod_{x' \in \sigma(\mu)} |x - x'|^2 ,$$

$$f_i(x) = \prod_{x' \in \sigma(\mu) - \{x_i\}} \frac{|x - x'|^2}{|x_i - x'|^2} , \qquad i = 1, \cdots, m,$$

where $|x - x'|$ is the usual Euclidean distance from x to x'.

It is clear that the above functions are polynomials; however, aside from continuity, the only properties of them which we shall use are the following:

$$f_0(x) \geq 0 \text{ for all } x \in X \text{ and } f_0(x) = 0$$

if and only if $x \in \sigma(\mu)$.

$$f_i(x) \geq 0 \text{ for all } x \in X \text{ and } f_i(x) = 0$$

if and only if $x \in \sigma(\mu) - \{x_i\}$, $(i = 1, \cdots, m)$.

[1] Theorem 6, fourth inequality pp. 175–176 of [**1**].

$$f_i(x_i) = 1, i = 1, \cdots, m.$$

In a precisely analogous manner we define the polynomials g_0 and g_j, $j = 1, \cdots, n$, on the set Y.

Next, let $\alpha_0, \alpha_1, \cdots, \alpha_n$ be $n+1$ distinct cluster points of $\bar{X}$ (the closure of X) which do not meet $\sigma(\mu)$ (these exist by hypothesis), and define polynomials ϕ and ϕ_j, $j = 0, \cdots, n$ on X via

$$\phi(x) = \prod_{k=0}^{n} |x - \alpha_k|^2 ,$$

$$\phi_j(x) = \prod_{\substack{k \neq j \\ 0 \leq k \leq n}} |x - \alpha_k|^2, \qquad j = 0, \cdots, n.$$

The only properties of these functions we shall use are that they are all non-negative, that ϕ vanishes only on the α_k, and that ϕ_j vanishes only on α_k with $k \neq j$.

Finally, let $\beta_0, \cdots, \beta_m$ be $m+1$ distinct cluster points of $\bar{Y}$ which do not meet $\sigma(\nu)$, and define polynomials ψ and ψ_i on Y analogous to the functions ϕ and ϕ_j above.

We now define the desired payoff M by

$$\begin{aligned} M(x, y) = {} & f_0(x)\phi(x)(g_0(y)\phi_0(x) + \sum_{j=1}^{n} (g_j(y) - \nu_j)\phi_j(x)) \\ & - g_0(y)\psi(y)(f_0(x)\psi_0(y) + \sum_{i=1}^{m} (f_i(x) - \mu_i)\psi_i(y)) \\ & - (f_0(x)\phi(x))^2 + (g_0(y)\psi(y))^2 . \end{aligned} \tag{1}$$

We show first that μ and ν are optimal strategies. If we compute $M(x, \nu)$ (in the usual extension), we obtain

$$M(x, \nu) = -(f_0(x)\phi(x))^2 \leq 0. \tag{2}$$

To see this, it is sufficient to observe that, according to the properties noted above, $\int g_j d\nu = \nu_j$ and g_0 and ψ vanish on $\sigma(\nu)$. Similarly, we obtain

$$M(\mu, y) = (g_0(y)\psi(y))^2 \geq 0.$$

Thus μ and ν are optimal and 0 is the value of the game.

It follows also from (2) above that if μ' is any optimal strategy for player I, then the spectrum of μ' is contained in the zeros of $f_0\phi$. Thus any optimal μ' has weight only on the pure strategies x_i and α_j, and similarly any optimal ν for player II restricts its weight to $\{y_j\} \cup \{\beta_i\}$.

We now show that ν is the only optimal strategy for player II. For suppose ν' is optimal. Then, in the expression for $M(x, \nu)$, the second and fourth terms in (1) drop out in view of the remark of the preceding paragraph and the payoff becomes

$$M(x, \nu') = f_0(x)\varphi(x)\Big[\varphi_0(x)\int g_0 d\nu' + \sum_{j=1}^{n} \varphi_j(x)\int g_j(y) - \nu_j) d\nu' - f_0(x)\varphi(x)\Big] .$$

For x close to α_0 the expression in brackets above approaches $\varphi_0(\alpha_0)\int g_0 d\nu'$, and since $\varphi_0(\alpha_0)$ is positive we must have $\int g_0 d\nu' = 0$. Otherwise the x-player by choosing x sufficiently close to α_0 could achieve a positive payoff, contradicting the optimality of ν'. Next, since $\int g_0 d\nu' = 0$, ν' must concentrate all of its mass on the zeros of g_0, that is, on the points y_j. Finally, if $\nu_k' \neq \nu_k$ for some index k then the x-player could again achieve a positive payoff by choosing x sufficiently close to x_k. It follows that $\nu' = \nu$ as asserted.

Thus, Theorem 2 is established.

3. Metric space games—construction of payoff. This section is dedicated to the construction of the payoff required for the establishment of Theorem 1 and its corollary, which will be proved in the final section. The construction and method of proof are quite similar to those used in proving Theorem 2; however, to preserve continuity of presentation, we shall paraphrase identical details.

Therefore, let X and Y be the respective spaces according to hypothesis, μ and ν the respective finite probability measures on them. μ and ν will be described with the same notations used previously. Finally, let ρ and ρ' denote the associated metrics of X and Y respectively. Then, without further ado, we initiate our construction.

The basis of our construction hinges on the fact that any infinite metric space contains a sequence of disjoint neighborhoods. To see this for X, say, there is no loss in generality in assuming that X has a cluster point, for otherwise we are guaranteed a sequence by the discrete topology induced by ρ and the infiniteness of X. Therefore, let x^* denote a cluster point of X. First, choose $\alpha_1 \neq x^*$, and, for $i > 1$ choose α_i so that $0 < \rho(x^*, \alpha_i) < \rho(x^*, \alpha_{i-1})/2$. Then, as our sequence of neighborhoods, $\{N_{\alpha_i}\}$, we set

$$N_{\alpha_i} = \{x \mid \rho(x, \alpha_i) < r_i\} , \qquad i = 1, 2, \cdots,$$

where $r_i = \rho(x^*, \alpha_i)/3$. It is easy to verify, using the triangle inequality, that these neighborhoods are disjoint.

Therefore, let $\{N_{\alpha_i}\}$ denote a sequence of disjoint neighborhoods contained in X (spheres of radius r_i centered at α_i). Define functions ϕ_j, $j = 0, \cdots, n$, as follows:

$$\phi_j(x) = \begin{cases} \dfrac{r_i - \rho(x, \alpha_i)}{r_i} \cdot \dfrac{1}{i} \text{ if } x \in N_{\alpha_i} \text{ for some } i \text{ (at most one) and} \\ \qquad\qquad i \equiv j \pmod{(n+1)} \\ 0 \text{ otherwise.} \end{cases}$$

One verifies that ϕ_j is a bounded continuous function on X into the non-negative reals, and which, moreover, satisfies

$$\phi_j(\alpha_i) = \begin{cases} \dfrac{1}{i} \text{ if } i \equiv j \pmod{(n+1)} \\ 0 \text{ otherwise.} \end{cases} \tag{6}$$

Next, let the function ϕ be given by

$$\phi(x) = \prod_{i=1}^{m} \rho(x, x_i), \qquad x \in X, \tag{7}$$

(where, as previously, $\{x_i\} = \sigma(\mu)$). There is no question about continuity here. We note merely that

$$\phi(x) \begin{cases} = 0 \text{ if } x \in \sigma(\mu) \\ > 0 \text{ otherwise.} \end{cases} \tag{8}$$

Finally, we define functions f_j, $j = 0, \cdots, m$, as follows:

$$f_0(x) = \phi(x) , \tag{9}$$

and, for $j \in \{1, \cdots, m\}$, set

$$f_j(x) = \prod_{i \neq j} \frac{\rho(x, x_i)}{\rho(x_j, x_i)} . \tag{10}$$

Here, again, continuity is immediate, and we note merely that

$$f_j(x_i) = \delta_{ij}, \qquad i, j = 1, \cdots, m, \tag{11}$$

where δ is Kronecker's delta. Moreover, to insure boundedness of these functions, if such is not the case, we need only replace ρ by the function $\rho/(1+\rho)$ in the formulas (7) and (10) without affecting subsequent arguments.

The remainder of our construction involves defining certain bounded continuous functions on Y into the non-negative reals. To accomplish this we merely repeat the foregoing construction with the replacements:

"X" → "Y",	"m" → "n",
"ρ" → "ρ'",	"n" → "m",
"x" → "y",	"ϕ" → "ψ",
"α" → "β",	"f" → "g",
"r" → "r'",	"μ" → "ν".

In terms of these functions, then, and using the convention $\mu_0 = \nu_0 = 0$, we define our bounded continuous polynomial-like payoff M as follows :

$$M(x, y) = -\psi(y)\sum_{j=0}^{m}(f_j(x) - \mu_j)\psi_j(y) + \phi(x)\sum_{j=0}^{n}(g_i(y) - \nu_j)\phi_j(x) - \phi(x)\sum_{j=0}^{n}\phi_j(x)^2 + \psi(y)\sum_{j=0}^{m}\psi_j(y)^2 ,$$

$(x, y) \in X \times Y$. This completes our construction.

4. Verification of solution and proof of uniqueness. To verify that (μ, ν) is a solution, we calculate first the expectation $M(\mu, y)$:

$$M(\mu, y) = \psi(y)\sum_{j=0}^{m}\psi_j(y)^2 \geq 0, \qquad \text{all } y \in Y.$$

To see this, we note that the remaining sums vanish by virtue of (8), (9), and (11), i.e. ϕ vanishes on $\sigma(\mu)$ and $\int f_j d\mu = \mu_j$, $j = 0. \cdots, m$. Similarly,

$$M(x, \nu) = -\phi(x)\sum_{j=0}^{n}\phi_j(x)^2 \leq 0, \qquad \text{all } x \in X. \tag{12}$$

Thus, (μ, ν) is a solution and 0 is the value of the game.

To show uniqueness for the first player, let μ' denote an optimal strategy for him. From the non-negativity of the functions ϕ, ϕ_j in (12), we see that $\int \phi\phi_j^2 d\mu' = 0$ for all $j \in \{0, \cdots, n\}$ and hence that $\int \phi\phi_j d\mu' = 0$; for otherwise, by (12), a counter strategy is provided by ν. Thus, if μ' is optimal, we have

$$M(\mu', y) = -\psi(y)\sum_{j=0}^{m}(\mu - \mu_j)\psi_j(y) + \psi(y)\sum_{j=0}^{m}\psi_j(y)^2 , \tag{13}$$

where we have written $\mu'_j = \int f_j d\mu'$, $j = 0, \cdots, m$. Next, suppose $\mu'_0 = \int f_0 d\mu' \neq 0$ (and hence, positive). Choose as possible counters a subsequence of the β's, $\{\beta_{\alpha_i}\}$ such that $n_i \equiv 0 \pmod{(m + 1)}$. Then, by virtue of the minimizer's counterpart of (6), (13) becomes

$$M(\mu', \beta_{n_i}) = -\psi(\beta_{n_i})\mu'_0 \cdot \frac{1}{n_i} + \psi(\beta_{n_i})\frac{1}{n_i^2} = \frac{\psi(\beta_{n_i})}{n_i}\Big(-\mu'_0 + \frac{1}{n_i}\Big), \qquad i = 1, 2, \cdots. \tag{14}$$

Since ψ vanishes only on a finite set and is positive elsewhere, we see that the expression above can be made negative for i sufficiently large. Hence $\mu'_0 = 0$, and if follows from (8) and (9) that $\sigma(\mu') \subseteq \sigma(\mu)$, i.e. any optimal μ' must restrict its spectrum to the set $\{x_1, \cdots, x_m\}$. Thus, finally, to establish uniqueness, we need only show that the corresponding weights are equal. Let μ'_i denote the weight on x_i placed by μ'. Substituting in our payoff M we obtain (noting $\mu'_0 = 0$),

$$M(\mu', y) = -\psi(y) \sum_{j=1}^{m} (\mu'_j - \mu_j)\psi_j(y) + \psi(y) \sum_{j=0}^{m} \psi_j(y)^2. \tag{15}$$

Now suppose $\mu'_k \neq \mu_k$ for some $k \in \{1, \cdots, m\}$. Then, since

$$\sum_{j=1}^{m} \mu'_j = \sum_{j=1}^{m} \mu_j = 1 ,$$

we would have some $j = j_0 \in \{1, \cdots, m\}$ such that $\mu'_{j0} > \mu_{j0}$. But, by choosing the subsequence $\{\beta_{n_i}\}$ with $n_i \equiv j_0 \pmod{(m+1)}$, by the identical argument used before, we would find a counter rendering the expectation (15) negative. Hence $\mu'_j = \mu_j$ and thus $\mu' = \mu$. Uniqueness for the minimizer can be established in a similar manner, as is clear. So Theorem 1 is proved.

Finally, to establish the corollary, we need only make the appropriate identifications in our payoff to ensure that $M(x, y) = -M(y, x)$.

The authors would like to thank Dr. Irving Glicksberg for his valuable comments on this paper. As a matter of fact, Dr. Glicksberg suggested an alternate proof for Theorem 1 which extends it to completely regular spaces X, Y. The gist of his proof involves obtaining the extended theorem by making it a corollary of Theorem 2 via a mapping : $X \to R^n, Y \to R^m$.

REFERENCES

1. M. Dresher, S. Karlin, and L. S. Shapley, *Polynomial games*, Ann. of Math. Study No. **24** (Princeton 1950), 161–180.

2. I. Glicksberg and 0. Gross, *Notes on games over the square*, Ann. of Math. Study No. **28** (Princeton 1953), 173–182.

RAND CORPORATION AND BROWN UNIVERSITY

ON THE NUMBER OF BI-COLORED GRAPHS

FRANK HARARY

1. Introduction. This is an extension of papers [2, 3, 4] whose notation and terminology will be used. The main result is a formulation of the generating function or counting polynomial of bicolored graphs, obtained by the enumeration methods of Pólya [6]. A modification of the method yields the number of balanced signed graphs, solving a problem proposed in [5]. In the process of enumerating bicolored graphs, we consider two binary operations on permutation groups called "cartesian product" and "exponentiation" which are abstractly but not permutationally equivalent to the direct product and Pólya's "Gruppenkranz" [6], respectively.

A *graph* consists of a finite set of *points* together with a prescribed subset of the collection of all *lines*, i.e., unordered pairs of distinct points. Two points are *adjacent* if there is a line joining them. A graph is *k-chromatic*[1] if each of the points can be assigned one of k given colors so that any two adjacent points have different colors. A graph is k-colored if it is k-chromatic and its points are colored so that all k colors are used. More precisely, a k-colored graph is a pair (G,f) where G is a graph and f is a function from the set of points of G *onto* the set of numbers $1, 2, \cdots, k$ such that if a and b are adjacent points, then $f(a) \neq f(b)$. Two graphs are *isomorphic* if there exists a one-to-one adjacency preserving transformation between their sets of points. Two k-colored graphs are *chromaticallg isomorphic* if there is a color preserving isomorphism between them. Thus (G_1, f_1) is chromatically isomorphic with (G_2, f_2) if there is an isomorphism $\theta\colon G_1 \to G_2$ and a permutation $\omega\colon \{1, \cdots, k\} \to \{1, \cdots, k\}$ such that $\omega(f_1(a)) = f_2(\theta(a))$ for every point a in G_1. Let $g_{p,q}^{(k)}$ be the number of chromatically nonisomorphic k-colored graphs with p points and q lines, and let the corresponding generating function be

$$g_p^{(k)}(x) = \sum_{q=0}^{p(p-1)/2} g_{p,q}^{(k)} x^q . \tag{1}$$

We first derive the number of bicolored graphs, $k = 2$, and then discuss the formula for $k = 3$. The problem remains open for $k > 2$.

Received November 4, 1957, and in revised forms March 28, 1958, and June 16, 1958. This work was supported by a grant from the National Science Foundation. The author is deeply grateful to the referee for making several insightful comments of clarification.

[1] This definition is different from that of Dirac [1]. According to Dirac, a graph has *chromatic number k* if it is k-chromatic but not $(k-1)$-chromatic as defined here.

In precisely the form in which we require it, Pólya's enumeration theorem is reviewed briefly in §2 of [2]. Therefore, we shall not repeat here the definitions leading up to it, but shall only restate the theorem itself.

PÓLYA'S THEOREM. *The configuration counting series $F(x)$ is obtained by substituting the figure counting series $\varphi(x)$ into the cycle index $Z(\Gamma)$ of the configuration group Γ. Symbolically,*

$$F(x) = Z(\Gamma, \varphi(x)) . \tag{2}$$

This theorem reduces the problem of finding the configuration counting series to the determination of the figure counting series and the cycle index of the configuration group.

2. Bicolored graphs; the cartesian product of permutation groups. Let K_n be the complete graph of n points, in which any two points are adjacent. Let K_{mn} be the bicolored graph whose $m + n$ points are $a_1, a_2, \cdots, a_m, b_1, b_2, \cdots, b_n$ and whose mn lines are all those of the form a_ib_j.

Clearly if a graph is k-colored then its point set is partitioned into k disjoint non-empty subsets such that no two points in the same subset are adjacent. Hence a bicolored graph with p points is a "line-subgraph" (as in [3]) or a spanning subgraph[2] of a graph K_{mn} for which $m + n = p$. Let $g_{mn,p}$ be the number of chromatically nonisomorphic spanning subgraphs of K_{mn} having q lines, and let

$$g_{mn}(x) = \sum_{q=0}^{mn} g_{mn,q}x^q . \tag{3}$$

Then

$$g_p^{(2)}(x) = \sum_{1 \leqq m \leqq n} g_{mn}(x) , \tag{4}$$

where the sum is taken over all m and n such that $m + n = p$. Therefore in order to obtain a formula for the counting polynomial (4), it is sufficient to find that for (3). In this section, we find $g_{mn}(x)$ for the case $m \neq n$ using the "cartesian product" of two permutation groups. In the next section we see that this combinatorial technique is not valid for $m = n$ and formulate $g_{nn}(x)$ in terms of the "exponentiation" of the appropriate two permutation groups.

By Theorem 1 of [3], the counting polynomial $g_{mn}(x)$ for the

---

[2] A *spanning subgraph* of a graph G is one whose set of points coincides with that of G.

number of spanning subgraphs of K_{mn} is obtained by substituting $1+x$ into the cycle index of the line-group[3] of K_{mn}:

$$g_{mn}(x) = Z(\Gamma_1(K_{mn}), 1+x) . \tag{5}$$

We note that this equation can also be obtained from the main result, equation (5), of [4]. For the subgraphs of K_{mn} correspond to the different supergraphs of the union $K_m \cup K_n$ of two complete graphs on disjoint point sets. The derivation of $Z(\Gamma_1(K_{mn}))$ for the case $m \neq n$ is parallel to and algebraically simpler than that of $Z(\Gamma_1(K_p))$, which appears in §3 of [2]. Throughout the rest of this section we assume $m \neq n$.

The line group of K_{mn} may be described as an appropriately formulated product of the symmetric groups $\mathbf{S}_m$ and $\mathbf{S}_n$. This product can be generally defined for any two permutation groups in the following way. Let **A** and **B** be any two permutation groups with object sets X and Y, degrees d and e and orders m and n respectively. The *cartesian product* $\mathbf{A} \times \mathbf{B}$ of these two permutation groups has degree de and order mn. Its object set is the cartesian product of X and Y and each of its permutations (α, β) is the cartesian product of permutations α and β from **A** and **B** defined by $(\alpha, \beta)(x, y) = (\alpha x, \beta y)$. As an abstract group, the cartesian product is isomorphic to the direct product **AB**, but they are not permutationally equivalent. For the degree of the direct product is $d+e$ since the group **AB** has $X \cup Y$ as its object set.

There is a precise method for finding the cycle index of a cartesian product in terms of the cycle indices of the two permutation groups. We first illustrate the method by finding $Z(\Gamma_1(K_{23}))$. The line group of K_{23} is the cartesian product of $\mathbf{S}_2$ and $\mathbf{S}_3$ which is a permutation group of degree 6 and order 12, written $\Gamma_1(K_{23}) = \mathbf{S}_2 \times \mathbf{S}_3$. Let a_1, a_2 be the indeterminates occurring in $Z(\mathbf{S}_2)$ and b_1, b_2, b_3 be those in $Z(\mathbf{S}_3)$ so that

$$Z(\mathbf{S}_2) = \frac{1}{2}(a_1^2 + a_2) \text{ and } Z(\mathbf{S}_3) = \frac{1}{6}(b_1^3 + 3b_1b_2 + 2b_3) .$$

Then we write[4]

$$\begin{aligned} Z(\Gamma_1(K_{23})) &= Z(\mathbf{S}_2 \times \mathbf{S}_3) = Z(\mathbf{S}_2) \times Z(\mathbf{S}_3) \\ &= \frac{1}{12}(a_1^2 \times b_1^3 + 3a_1^2 \times b_1b_2 + 2a_1^2 \times b_3 + a_2 \times b_1^3 + 3a_2 \times b_1b_2 + 2a_2 \times b_3) , \end{aligned}$$

---

[3] The *line group* $\Gamma_1(G)$ of a graph G is the collection of all permutations on the set of lines of G consistent with the automorphism group $\Gamma(G)$ of G; see [**3**].

[4] By the following formulas we mean that the cartesian product of two permutation groups can be extended to the cartesian product of their cycle indices in the indicated manner.

and give each of these six terms in Table 1, in which c_1 to c_6 denote the indeterminates in $Z(\Gamma_1(K_{23}))$.

Table 1

Term of $Z(S_2 \times S_3)$	$a_1^2 \times b_1^3$	$a_1^2 \times b_1 b_2$	$a_1^2 \times b_3$	$a_2 \times b_1^3$	$a_2 \times b_1 b_2$	$a_2 \times b_3$
Term of $Z(\Gamma_1(K_{23}))$	c_1^6	$c_1^2 c_2^2$	c_3^2	c_2^3	c_2^3	c_3

We illustrate Table 1 for the term $a_2 \times b_1 b_2$. Let the 2-cycle $(p_1 p_2)$ stand for a_2 and the 1-cycle and 2-cycle $(q_1)(q_2 q_3)$ for $b_1 b_2$. The admissible lines of K_{23} are only those of the form $p_i q_j$. The pair $p_1 q_1$ is transformed into $p_2 q_1$, and then back again to give the cycle of length 2 in the corresponding permutation of $\Gamma_1(K_{23})$ of the form $(p_1 q_1 \; p_2 q_1)$. Similarly the transpositions $(p_1 q_2 \; p_2 q_3)$ and $(p_1 q_3 \; p_2 q_2)$ are factors of this element of $\Gamma_1(K_{23})$. Altogether there are three transpositions, so the corresponding term of $Z(\Gamma_1(K_{23}))$ is c_2^3.

In general, we have

$$a_1^{i_1} a_2^{i_2} \cdots a_m^{i_m} \times b_1^{j_1} b_2^{j_2} \cdots b_n^{j_n} = \prod_{\alpha, \beta} (a_\alpha^{i_\alpha} \times b_\beta^{j_\beta}) \tag{6}$$

and

$$a_\alpha^{i_\alpha} \times b_\beta^{j_\beta} = c_{m(\alpha,\beta)}^{i_\alpha j_\beta d(\alpha,\beta)} \tag{7}$$

where $d(\alpha, \beta)$ and $m(\alpha, \beta)$ are the greatest common divisor and least common multiple.

The cycle index of S_p is

$$Z(S_p) = \frac{1}{p!} \sum_{(j)} \frac{p!}{1^{j_1} j_1! \cdots p^{j_p} j_p!} f_1^{j_1} \cdots f_p^{j_p} \tag{8}$$

where the sum is taken over all partitions $(j) = (j_1, j_2, \cdots, j_p)$ of p such that

$$1j_1 + 2j_2 + \cdots + pj_p = p \, .$$

The last four equations together with

$$Z(\Gamma_1(K_{mn})) = Z(S_m) \times Z(S_n) \tag{9}$$

provide a formula for $g_{mn}(x)$ when $m \neq n$.

We use Table 1 to illustrate equation (9) by finding $g_{23}(x)$. Here

$$Z(\Gamma_1(K_{23})) = \frac{1}{12}(c_1^6 + 3c_1^2 c_2^2 + 2c_3^2 + 4c_2^3 + 2c_6) \, ,$$

so that

$$g_{23}(x) = 1 + x + 3x^2 + 3x^3 + 3x^4 + x^5 + x^6 \, .$$

The bicolored graphs with two points of one color and three points of the other color which correspond to the coefficients in the preceding counting polynomial are shown in Figure 1.

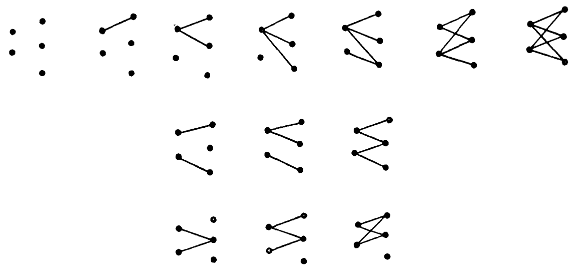

Fig. 1.

3. Bicolored graphs; exponentiation of permutation groups. We now turn to the enumeration of bicolored graphs for the case $m = n$. As in the preceding section, we again have equation (5) holding for this special case:

$$g_{nn}(x) = Z(\Gamma_1(K_{nn}), 1 + x) .$$

However, it is not true that $\Gamma_1(K_{nn}) = \mathrm{S}_n \times \mathrm{S}_n$ since $\mathrm{S}_n \times \mathrm{S}_n$ is a proper subgroup of $\Gamma_1(K_{nn})$. The remaining $(n!)^2$ permutations in $\Gamma_1(K_{nn})$ are obtained on interchanging the two colors in accordance with the definition of chromatic isomorphism.

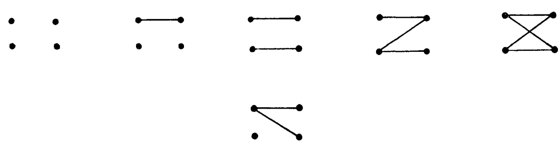

Fig. 2.

For example, all the (chromatically nonisomorphic) bicolored graphs for $m = n = 2$ are shown in Figure 2, so that $g_{22}(x) = 1 + x + 2x^2 + x^3 + x^4$. However, the formulas in the preceding section give

$$Z(\mathrm{S}_2 \times \mathrm{S}_2, 1 + x) = 1 + x + 3x^2 + x^3 + x^4 ,$$

since the permutations in $\mathrm{S}_2 \times \mathrm{S}_2$ distinguish between the two bicolored graphs in Figure 3, in which the color assigned to each point is indicated by one of the integers 1 or 2.

With the appropriate definition of group exponentiation, we will express $\Gamma_1(K_{nn})$ as S_n raised to the power S_2. We first review the definition of the "composition" of two permutation groups (the "Gruppenkranz" of Pólya [6]). Let **A** and **B** be any two permutation groups as in the preceding section. Then using the notation of Pólya [6] the *composition* **A**[**B**] of **A** with **B** has object set $X \times Y$ (as for the cartesian product). However, it is more convenient to regard the object set here as a d by e matrix $M = (x_{ij})$. Then the elements of **A** [**B**] are the permutations of the entries of M constructed as follows. First permute the rows of M in accordance with an element of **A**. Then permute the column indices in each row separately using one element

Fig. 3.

of **B** for each row, repetitions permitted. Hence the degree of **A** [**B**] is de and the order is mn^d.

The *exponentiation* $\mathbf{B}^{\mathbf{A}}$ of **A** with B is that permutation group whose object set is Y^X, the collection of all functions from X into Y, and whose elements are constructed as follows. It is assumed that the objects $x_1, x_2, \cdots, x_d$ in X are indexed. First permute the objects in X in accordance with an element α of **A**. Then for each object x_1 in X, permute the e objects of Y into which it can be mapped, using a permutation β_i from **B**. More precisely each selection of $\alpha \in \mathbf{A}$ and $\beta_1, \beta_2, \cdots, \beta_d \in \mathbf{B}$ (not necessarily distinct) determines a permutation of Y^X which takes the function f into the function f^* defined by:

$$f^*(x_i) = \beta_i f(\alpha x_i) \text{ for all } x_i \in X; \qquad i = 1, 2, \cdots, d .$$

It can easily be shown that distinct selections of $\alpha, \beta_1, \cdots, \beta_d$ lead to distinct permutations of Y^X and that these permutations form a group.

The degree of $\mathbf{B}^{\mathbf{A}}$ is e^d and the order is mn^d. It follows at once from their constructions that the group $\mathbf{B}^{\mathbf{A}}$ and **A**[**B**] are isomorphic as abstract groups. But they are not equivalent as permutation groups since they have different degrees.

With this definition of exponentiation, it follows at once that the line group of K_{nn} is given by

$$\Gamma_1(K_{nn}) = \mathbf{S}_{n}^{\mathbf{S}_2} . \tag{10}$$

Before calculating the cycle index of $\mathbf{S}_{n}^{\mathbf{S}_2}$, we illustrate for $n = 2$ and 3. Since $\mathbf{S}_2^{\mathbf{S}_2} = \mathbf{S}_2[\mathbf{S}_2] = \mathbf{D}_4$, the dihedral group of degree 4 and order 8, its cycle index in terms of the indeterminates c_1, c_2, c_3, c_4 is given by

$$Z(\mathbf{D}_4) = \frac{1}{8}(c_1^4 + 3c_2^2 + 2c_1^2 c_2 + 2c_4) .$$

The correct polynomial $g_{22}(x)$ which verifies Figure 2 follows at once from this cycle index.

For $n = 3$, let the object set of $\Gamma_1(K_{33})$ be denoted:

$$X = \{11', 12', 13', 21', 22', 23', 31', 32', 33'\} .$$

Then $\Gamma_1(K_{33}) = S_3^{S_2}$ contains the $(3!)^2$ permutations in $S_3 \times S_3$ and also the $(3!)^2$ permutations obtained from these on multiplying each of them by the following reflection ρ which interchanges primed and unprimed digits in the objects in X:

$$\rho = (11')(22')(33')(12'\,21')(23'\,32')(31'\,13') .$$

Symbolically, we write

$$\Gamma_1(K_{33}) = S_3^{S_2} = (S_3 \times S_3) \cup \rho(S_3 \times S_3) .$$

Then

$$Z(S_3 \times S_3) = \frac{1}{(3!)^2}(c_1^9 + 6c_1^3c_2^3 + 8c_3^3 + 9c_1c_2^4 + 12c_3c_6) ,$$

and a straightforward calculation gives (using not quite proper notation since cycle index is defined for groups rather than cosets):

$$Z(\rho(S_3 \times S_3)) = \frac{1}{(3!)^2}(6c_1^3c_2^3 + 18c_1c_4^2 + 12c_3c_6) .$$

Combining these, we have

$$Z(S_3^{S_2}) = \frac{1}{2\cdot(3!)^2}(c_1^9 + 12c_1^3c_2^3 + 8c_3^3 + 9c_1c_2^4 + 18c_1c_4^2 + 24c_3c_6) ,$$

from which one readily calculates using (10) and (5),

$$g_{33}(x) = 1 + x + 2x^2 + 4x^3 + 5x^4 + 5x^5 + 4x^6 + 2x^7 + x^8 + x^9 .$$

We now proceed to obtain a closed formula for $Z(S_n^{S_2})$, thereby completing the explicit solution of the enumeration of bicolored graphs. The process of finding this cycle index is also analogous to the calculation of of $Z(\Gamma_1(K_p))$ which appears in §3 of [2]. Clearly, the automorphism group of K_{nn} is $S_2[S_n]$. For the complement K'_{nn} consists of two disjoint copies of K_n. By a result in Pólya [6], the cycle index of the composition of two permutation groups is the composition of their cycle indices. For example,

$$Z(S_2[S_3]) = \frac{1}{2}\left[\left(\frac{1}{6}(a_1^3 + 3a_1a_2 + 2a_3)\right)^2 + \frac{1}{6}\left(a_2^3 + 3a_2a_4 + 2a_6\right)\right]$$

But we require here the cycle index of the line group of K_{nn}. There is a one-to-one correspondence between the terms of the cycle indices

$Z(S_n^{S_2})$ and $Z(S_2[S_n])$ with the same integral coefficients. Analogous to the terms in the above illustration of $Z(S_2[S_3])$, let us write

$$Z(S_2[S_n]) = \frac{1}{2}[(Z(S_n))^2 + Z(S_n(2))] . \tag{11}$$

Thus $Z(S_n(2))$ is obtained from $Z(S_n)$ on replacing each indeterminate f_k by f_{2k}.

The term of $Z(S_n^{S_2})$ corresponding to the first term of the right hand member of equation (11) is $Z(S_n \times S_n)$. For bicolored graphs with $m \neq n$, this is the result of the preceding section. The term of $Z(S_n^{S_2})$ corresponding to the term $Z(S_n(2))$ of (11) is derived as follows. Let the general term of $Z(S_n(2))$ be given by

$$f_2^{j_1} f_4^{j_2} \cdots f_{2n}^{j_n} . \tag{12}$$

This term (12) occurs in the cycle index of the point group of K_{nn}. We require the corresponding term in the cycle index of the line group obtained by calculating the induced permutation on pairs of points from two disjoint sets. Let the letters c_i be the indeterminates in the cycle index $Z(S_n^{S_2})$. There are two contributions to $Z(S_n^{S_2})$ arising from (12): those from each of the n factors $f_{2k}^{j_k}$ separately, and those from pairs of factors $f_{2r}^{j_r} f_{2s}^{j_s}$, $r \neq s$.

The contribution to the cycle index due to each factor in (12) is

$$f_2^{j_1} \to c_1^{j_1} c_2^{\binom{j_1}{2}}, \quad f_4^{j_2} \to c_4^{j_2} c_4^{2\binom{j_2}{2}}, \quad f_6^{j_3} \to (c_3 c_6)^{j_3} c_6^{3\binom{j_3}{2}}, \quad \cdots$$

It is convenient to express the contribution of $f_{2k}^{j_k}$ separately for k even and k odd:

$$f_{2k}^{j_k} \to \begin{cases} (c_{2k}^{k/2})^{j_k} c_{2k}^{k\binom{j_k}{2}}, & k \text{ even} \\ (c_k c_{2k}^{(k-1)/2})^{j_k} c_{2k}^{k\binom{j_k}{2}}, & k \text{ odd} . \end{cases}$$

Similarly, the contribution from pairs is given by

$$f_{2r}^{j_r} f_{2s}^{j_s} \to c_{2m(r,s)}^{j_r j_s d(r,s)} , \qquad r < s ,$$

where $m(r, s)$ and $d(r, s)$ are the least common multiple and greatest common divisior respectively.

Collecting these observations, we find

$$Z(S_n^{S_2}) = \frac{1}{2}[Z(S_n \times S_n) + Z'] , \qquad \text{where}$$

$$Z' = \frac{1}{n!} \sum_{(j)} \frac{n!}{\prod_k k^{j_k} j_k!} \prod_{k \text{ even}} c_{2k}^{k j_k/2 + k\binom{j_k}{2}} \prod_{k \text{ odd}} \left(c_k c_{2k}^{(k-1)/2}\right)^{j_k} c_{2k}^{k\binom{j_k}{2}} \prod_{r<s} c_{2m(r,s)}^{j_r j_s d(r,s)} .$$

This formula for $Z(\mathbf{S}_n^{\mathbf{S}_2})$ together with equations (10) and (5) give the number of bicolored graphs for $m = n$. For $n = 3$, this expression for Z' specializes to that for $Z(\rho(\mathbf{S}_3 \times \mathbf{S}_3))$ in the above example.

The only other known cycle index of the exponentiation of two permutation groups also involves complex combinatorial calculations and is worked out in Slepian [8]. Consider the counting polynomial $b_n(x) = \sum b_{nm} x^m$, where b_{nm} is the number of symmetry types of boolean functions of n variables having m nonzero terms when written in disjunctive normal form. Pólya [7] showed that

$$b_n(x) = Z(\mathbf{Q}_n,\ 1 + x) ,$$

where $\mathbf{Q}_n$ is the automorphism group of the n-cube. It is easily seen that $\mathbf{Q}_n = \mathbf{S}_2^{\mathbf{S}_n}$ and in fact Pólya [7, footnote 7] comments that $\mathbf{Q}_n$ and $\mathbf{S}_n[\mathbf{S}_2]$ are isomorphic as abstract groups. Slepian [8] completed the enumeration problem for $b_n(x)$ by providing a calculus for an explicit formulation of $Z(\mathbf{S}_2^{\mathbf{S}_n})$, although using different terminology and notation.

It would be interesting to solve the general problem of obtaining an expression for $Z(\mathbf{B}^{\mathbf{A}})$ in terms of $Z(\mathbf{A})$ and $Z(\mathbf{B})$. This would be analogous to equations (6) and (7) which give $Z(\mathbf{A} \times \mathbf{B})$ in terms of $Z(\mathbf{A})$ and $Z(\mathbf{B})$.

To summarize, the counting polynomial $g_{mn}(x)$ for bicolored graphs is given by

$$g_{mn}(x) = \begin{cases} Z(\mathbf{S}_m \times \mathbf{S}_n, 1 + x) & \text{when } m \neq n \\ Z(\mathbf{S}_n^{\mathbf{S}_2}, 1 + x) & \text{when } m = n . \end{cases} \tag{14}$$

4. **k-colored graphs.** We illustrate the general problem for $k = 3$. Here we have, analogous to equations (3), (4) and (5), and with similar notation:

$$g_{mnt}(x) = \sum_{q=0}^{mnt} g_{mnt,q} x^q , \tag{3'}$$

$$g_p^{(3)}(x) = \sum_{1 \leq m \leq n \leq t} g_{mnt}(x) , \tag{4'}$$

$$g_{mnt}(x) = Z(\Gamma_1(K_{mnt}), 1 + x) . \tag{5'}$$

Thus K_{mnt} is the complete tricolored graph with $m + n + t$ points $p_1, p_2, \cdots, p_m, q_1, q_2, \cdots, q_n, r_1, r_2, \cdots, r_t$ and all $mn + nt + tm$ lines of the form p_iq_j, q_jr_k, and r_kp_i. Similarly $g_{mnt,q}$ is the number of spanning subgraphs of K_{mnt} having q lines, etc. We distinguish between three cases: (a) m, n, t distinct, (b) $m = n \neq t$, and (c) $m = n = t$. These are illustrated in Figures 4 (a), (b), and (c).

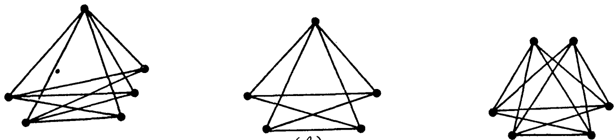

Fig. 4.

Only in case (a) have we obtained an algorithm for $Z(\Gamma_1(K_{mnt}))$ in closed form. The result analogous to (9), derived in same manner, is as follows. Let a_1 to a_m, b_1 to b_n, and c_1 to c_t be the indeterminates in $Z(S_m)$, $Z(S_n)$, and $Z(S_t)$ respectively. Let the indeterminates in $(\Gamma_1(K_{mnt}))$ be $d_1, d_2, \cdots$. Let A, B, and C denote arbitrary terms of $Z(S_m)$, $Z(S_n)$, and $Z(S_t)$ respectively. Then in this notation the left-hand member of equation (6) is $A \times B$ and the term of $Z(\Gamma_1(K_{mnt}))$ obtained from A, B, and C is,

$$(A \times B)(B \times C)(C \times A)\,, \tag{15}$$

where each of the three factors in the expression (15) is a product of the indeterminates d_i using equations (6) and (7). For example,

$$Z(\Gamma_1(K_{123})) = \frac{1}{12}(d_1^{11} + 3d_2^3d_1^5 + 2d_3^3d_1^2 + d_2^4d_1^3 + 3d_2^5d_1 + 2d_6d_3d_2)$$

is the cycle index of the line group of the tricolored graph K_{123} shown in Figure 4 (a).

Referring to Figure 4(b), one can find

$$Z(\Gamma_1(K_{122})) = \frac{1}{8}(d_1^8 + 2d_4^2 + 4d_1^2d_2^3 + d_2^4)\,.$$

The group $\Gamma_1(K_{122})$ appears to be irreducible by any of the operations of direct product, cartesian product, composition, or exponentiation. However, it is abstractly isomorphic to D_4 and can be obtained from two copies of D_4 defined on disjoint object sets $\{1, 2, 3, 4\}$ and $\{5, 6, 7, 8\}$ by the following operation.

Let the set X be the union of the disjoint sets X_1 and X_2. Let $\mathbf{A}_1$ and $\mathbf{A}_2$ be permutation groups defined on X_1 and X_2 respectively, such that h is an abstract isomorphism of $\mathbf{A}_1$ onto $\mathbf{A}_2$. Then the permutation group $\mathbf{A}_1 \oplus_h \mathbf{A}_2$ can be defined as follows: The function f from X onto X belongs to $\mathbf{A}_1 \oplus_h \mathbf{A}_2$ if and only if there exist $\alpha_1 \in \mathbf{A}_2$ and $\alpha_2 \in \mathbf{A}_2$

with $\alpha_2 = h\alpha_1$ such that $f(x) = \alpha_1(x)$ if $x \in X_1$ and $f(x) = \alpha_2(x)$ if $x \in X_2$. Clearly $\mathbf{A} = \mathbf{A}_1 \oplus_h \mathbf{A}_2$ is abstractly isomorphic to $\mathbf{A}_1$.

Now let $D_{4,1}$ be the dihedral group of degree 4 generated by the permutations (1234) and (12)(34), let $D_{4,2}$ be generated by (5678) and (57)(6)(8), and let h be the isomorphism between them which preserves respective generators. Then

$$\Gamma_1(K_{122}) = D_{4,1} \oplus_h D_{4,2} .$$

Finally, it is easy to see that $\Gamma_1(K_{222})$ is abstractly isomorphic to $S_3[S_2]$ and that

$$Z(\Gamma_1(K_{222})) = \frac{1}{3!\,2^3}[(c_1^{12} + 3c_1^4c_2^4 + 4c_2^6) + 3\cdot 2(2c_1^2c_2^5 + 2c_4^3) + 2\cdot 2^2(c_3^4 + c_6^2)]$$

$$= \frac{1}{48}(c_1^{12} + 3c_1^4c_2^4 + 4c_2^6 + 8c_3^4 + 8c_6^2 + 12c_1^2c_2^5 + 12c_4^3) .$$

It is clear that the line group of K_{nnn} is abstractly isomorphic to the automorphism group of K_{nnn}. Its complement K'_{nnn} consists of three disjoint copies of K_n, so that the group of K_{nnn} is $S_3[S_n]$. But an explicit expression for $Z(\Gamma_1(K_{nnn}))$ does not appear to be obvious. (For the particular case $n = 3$, it can be shown that $\Gamma_1(K_{333})$ is permutationally equivalent to $S_3^{S_3}$.) It does not appear that the operations considered here will suffice to enumerate even the tricolored graphs.

5. Connected k-colored graphs.

Let

$$g(x, y) = \sum_p g_p^{(k)}(x)y^p$$

be the generating function for *all* (connected or not) k-colored graphs, and let $c(x, y)$ be that for the connected ones only. Then to find the number of connected k-colored graphs, we substitute into equation (33) of [2] to get

$$1 + g(x, y) = \exp\left(\sum_{n=1}^{\infty} \frac{1}{n} c(x^n, y^n)\right) \tag{16}$$

or equivalently,

$$\sum_{n=1}^{\infty} \frac{1}{n} c(x^n, y^n) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} g^n(x,y) . \tag{16'}$$

From either of these last two equations, the number of connected k-colored graphs of p points can be computed recursively in terms of the total number of k-colored graphs.

6. Balanced signed graphs. *Signed graphs* are obtained by assigning either a positive or a negative sign to each line of a graphs. It was indicated in [5] how one could enumerate all signed graphs by a variation in one of the formulas of [2]. The *sign of a cycle* is the product of the signs of its lines, and a signed graph is *balanced* if all its cycles are positive. The problem of enumerating balanced signed graphs was proposed in [5]. The result is derivable by an appropriate modification of the generating function for bicolored graphs.

It was shown in [5] that a signed graph is balanced if and only if its set of points can be partitioned into two disjoint subsets such that each positive line and no negative line joins two points in the same subset. In view of this characterization, called the "structure theorem for balance", on deleting all the positive lines of a balanced signed graph one obtains a bicolored graph. Let G_1 and G_2 be arbitrary graphs with m and n points respectively, $m \leq n$, and let $p = m + n$. Let $\Gamma(G_1)$ and $\Gamma(G_2)$ be the groups of G_1 and G_2 respectively. Let $b_q(G_1, G_2)$ be the number of nonisomorphic balanced signed graphs with q negative lines, whose positive lines generate the (disjoint) graphs G_1 and G_2 in accordance with the structure theorem for balance. Let

$$b(G_1, G_2, x) = \sum_{q=0}^{mn} b_q(G_1, G_2)x^q$$

be the desired configuration counting series. Then the figure counting series is $1 + x$. For the figures are the mn pairs of points (c, d) where $c \in G_1$ and $d \in G_2$. The content of a figure (c, d) is 0 if c and d are not joined by a negative line and is 1 if they are.

Analogously to the situation for bicolored graphs there are two possibilities. If G_1 and G_2 are not isomorphic, then the configuration group is $\Gamma(G_1) \times \Gamma(G_2)$. But if they are isomorphic, the configuration group is $\Gamma(G_1)^{S_2}$. Hence an application of Pólya's Theorem yields

$$b(G_1, G_2, x) = \begin{cases} Z(\Gamma(G_1) \times \Gamma(G_2), 1 + x) & \text{when } G_1 \not\cong G_2 \\ Z(\Gamma(G_1)^{S_2} 1, + x) & \text{when } G_1 \cong G_2 \end{cases} \tag{17}$$

It is clear for the special case where G_1 and G_2 are the totally disconnected graphs of m and n points that $b(G_1, G_2, x) = g_{mn}(x)$, $\Gamma(G_1) = S_m$, and $\Gamma(G_2) = S_n$. Thus the formula (17) is a generalization of that for bicolored graphs.

References

1. G. A. Dirac, *The structure of k-chromatic graphs*, Fund. Math. **40** (1953), 42-55.
2. F. Harary, *The number of linear, directed rooted and connected graphs*. Trans. Amer. Math. Soc. **78** (1955), 445-463.

3. ——, *On the number of dissimilar line-subgraphs of a given graph*, Pacific. J. Math., **6** (1956), 57-64.
4. ——, *The number of dissimilar supergraphs of a linear graph*, Pacific J. Math, **7** (1957), 903-911.
5. ——, *On the notion of balance of a signed graph*, Michigan Math. J. **2** (1953-54), 143-146.
6. G. Pôlya, *Kombinatorische Anzahlbestimmungen für Gruppen, Graphen, und chemishe, Verbindungen.* Acta Math. **68** (1937), 145-254.
7. ——, *Sur les types des propositions composees*, J. Symbolic Logic **5** (1940), 98-103.
8. D. Slepian, *On the number of symmetry types of boolean functions of n variables*, Canadion J. Math. **5** (1953), 185-193.

UNIVERSITY OF MICHIGAN AND THE INSTITUTE FOR ADVANCED STUDY

CENTRALIZERS IN JORDAN ALGEBRAS

BRUNO HARRIS

Introduction. The aim of this paper is to prove for Jordan algebras some theorems on centralizers of subalgebras analogous to known results in the theory of associative algebras (contained in [6, Chapter 3] and [7, Chapter 6], for instance).

The definition of the centralizer of a subalgebra in a Jordan algebra is based on the concept of "operator commutativity" introduced by Jordan, von Neumann and Wigner in [17]: two elements x, y of the Jordan algebra J operator commute if the operators $R_x: a \to ax$ and $R_y: a \to ay$, acting on J, commute, that is $(ax)y = (ay)x$ for all elements a of J. In §1 we study this concept, extend the results of [8] to algebras over fields of characteristic not two, and show that for many types of Jordan algebras obtained from associative algebras by introducing the Jordan product $a \circ b = ab + ba$ (ab the associative product), the centralizer of a subalgebra is just the set of elements commuting in the associative multiplication with the elements of the subalgebra. Thus some of our later results can be regarded as generalizations of the associative algebra results if we convert the associative algebras into Jordan algebras by means of the Jordan product.

In §2 we generalize some of the theory of a single linear transformation in a finite dimensional vector space (see [6, Chapter 3] and [13]) to the subalgebra generated by a single element in a simple finite dimensional Jordan algebra. We show that such a subalgebra is equal to the centralizer of its centralizer, and we also generalize to any central simple Jordan algebra a formula of Frobenius giving the dimensionality of the centralizer of a single linear transformation in terms of the degrees of its invariant factors. A special case of this formula—namely, the formula for the central simple Jordan algebra of all symmetric matrices—was proved earlier, and by a different method, by H. Osborn (to appear in these Transactions).

In §3 we study the centralizer theory of a simple subalgebra in a central simple Jordan algebra. We show that the analogues of the centralizer and double centralizer theorems for simple finite dimensional subalgebras of the associative algebra of all linear transformations on a vector space ([15]) also hold for simple finite dimensional Jordan sub-

This paper was originally accepted by the Transactions of the American Mathematical Society, Received July 10, 1957, by Trans. Amer. Math. Soc. Most of the material of this paper is contained in the author's doctoral dissertation, Yale University, 1956. Some of the results were presented to the Society in October, 1955 and April, 1957. The author is a National Science Foundation Postdoctoral Fellow, 1956–57.

algebras of the Jordan algebra of all self-adjoint linear transformations on a vector space with inner product. Incidentally we show that some of the results of [15] can be generalized from the class of rings of all continuous linear transformations on a vector space to the larger class of primitive rings with minimal ideals. In the same way the Galois theory of automorphisms of [16] can be generalized to primitive rings with minimal ideals.

In conclusion we would like to express our gratitude to Professor Nathan Jacobson who suggested these problems and gave much stimulating advice.

Preliminary Notions. A Jordan algebra is a linear algebra, whose multiplication we shall denote by $x \circ y$, satisfying the following identities

$$(1) \qquad x \circ y = y \circ x$$

$$(2) \qquad ((x \circ x) \circ y) \circ x = (x \circ x) \circ (y \circ x) \, .$$

We shall always assume that the base field has characteristic different from 2.

A *special* Jordan algebra is a subspace of an associative algebra (with associative multiplication xy) closed under the composition $x \circ y = xy + yx$. The special Jordan algebra whose underlying vector space coincides with that of the associative algebra $\mathfrak{A}$ and whose multiplication is $x \circ y = xy + yx$ (xy the multiplication in $\mathfrak{A}$) will be denoted by $\mathfrak{A}_J$. If $\mathfrak{A}$ is an associative algebra with an involution, the subset of elements left fixed by the involution is also a special Jordan algebra, which will be denoted by $H(\mathfrak{A})$. The same notation will be used for the set of elements left fixed by an involution in a possibly non-associative algebra $\mathfrak{A}$: this set may or may not be a Jordan algebra.

We shall have to consider sometimes matrix algebras with coefficients in a (possibly non-associative) algebra with identity element. A set of matrix units in an algebra of all $n \times n$ matrices ($n \geqq 2$) will mean a set of elements e_{ij}, $i, j = 1, \cdots, n$ which associate with every pair of elements of the algebra (i.e. lie in the nucleus) and satisfy

$$(3) \qquad e_{ij}e_{kl} = \delta_{jk}e_{il} \ (\delta_{jk} \text{ the Kronecker delta})$$
$$e_{11} + \cdots + e_{nn} = 1, \text{ the identity element.}$$

If we consider Jordan algebras (with identity) of all hermitian matrices with coefficients in an involutorial algebra we are led to consider elements (which we shall also call matrix units) e_{ii}, u_{ij} with $i < j$, $i, j = 1, \ldots, n$, $n \geqq 3$, such that

$$(4) \qquad e_{ii} \circ e_{ii} = 2e_{ii} \, , \quad e_{ii} \circ u_{ij} = u_{ij} \, , \quad u_{ij} \circ u_{ij} = 2(e_{ii} + e_{jj})$$
$$u_{ij} \circ u_{jk} = u_{ik} \text{ if } i, j, k \text{ are distinct}, \ \textstyle\sum e_{ii} = 1$$

and all other products are zero. As shown in [9], Th. 9.1, any set of elements e_{ii}, u_{ij} satisfying (4) leads to a representation of the Jordan algebra as the subalgebra of $n \times n$ hermitian matrices of an algebra $\mathfrak{A}_j$ where $\mathfrak{A}$ is the algebra of all $n \times n$ matrices with coefficients in an involutorial algebra, and, if f_{ij}; $i, j = 1, \ldots, n$ are the matrix units in $\mathfrak{A}$, then

$$f_{ii} = e_{ii} \ , \quad u_{ij} = f_{ij} + f_{ji} \qquad (i < j) \ . \tag{5}$$

If the base field is algebraically closed and $\mathfrak{J}$ is the exceptional simple Jordan algebra of all 3×3 hermitian matrices with coefficients in the Cayley algebra, then given elements e_{ii}, $i = 1, 2, 3$, in $\mathfrak{J}$, satisfying $e_{ii} \circ e_{jj} = 2\delta_{ij}e_{ij}$ we can find elements u_{12}, u_{13} (in many different ways) satisfying $u_{12} \circ u_{12} = 2(e_{11} + e_{22})$, $u_{13} \circ u_{13} = 2(e_{11} + e_{33})$ such that the e_{ii}, u_{12}, u_{13} and $u_{23} = u_{12} \circ u_{13}$ satisfy the conditions (4) and hence are the "matrix units" of another representation of $\mathfrak{J}$ as 3×3 hermitian matrices with Cayley number coefficients.

Finally we shall summarize briefly the classification of the finite dimensional central simple Jordan algebras. For further references about classification, structure, or representation theory of Jordan algebras, one should consult [9].

First, assume the base field algebraically closed. Then the algebra has an identity element; if the identity element can be written as a sum of n, but not more, mutually orthogonal idempotents, then the algebra is said to have degree n. (e is an idempotent if $e \circ e = e$).

If $n = 1$, then the algebra is one-dimensional, [10]. If $n = 2$, the algebra is a vector space direct sum of the subspace generated by 1 and of a vector space V of dimension at least 2 with non-degenerate symmetric scalar product. The multiplication is

$$(\alpha 1 + x) \circ (\beta 1 + y) = [\alpha\beta + (x, y)]1 + \alpha y + \beta x$$

α, β scalars, x, y in V and (x, y) their scalar product. Such an algebra is said to be of type D.

If $n \geqq 3$, there are 4 types: A, B, C, E. Types A, B, C are special, while type E is the exceptional algebra described above. To each of the types A, B, C (and also to D) corresponds an associative algebra $\mathfrak{U}$ such that if the corresponding Jordan algebra is contained in an algebra $\mathfrak{A}_j$, $\mathfrak{A}$ associative, then the associative subalgebra of $\mathfrak{A}$ generated by $\mathfrak{J}$ (enveloping algebra of $\mathfrak{J}$) is a homomorphic image of $\mathfrak{U}$.

Type A: $\mathfrak{J} = \mathfrak{A}_j$, $\mathfrak{A}$ the associative algebra of all $n \times n$ matrices over the base field, $n \geqq 3$. $\mathfrak{U} = \mathfrak{A} \oplus \mathfrak{A}$

Type B: $\mathfrak{J}$ is the algebra of all $n \times n$ symmetric matrices, $n \geqq 3$. $\mathfrak{U} = \mathfrak{A}$ ($\mathfrak{A}$ as for type A).

Type C: $\mathfrak{J}$ is the algebra of all $2n \times 2n$ symplectic-symmetric matrices, isomorphic to the set of all self-adjoint linear transformations on a vector space V with non-degenerate skew-symmetric scalar product. $\mathfrak{U}$ is the algebra of all linear transformations on V.

Type D: $\mathfrak{U}$ is the Clifford algebra determined by the space V and the inner product.

If the base field is not algebraically closed, then the algebras which become of type A on extension of the base field are of two subtypes:

A_1: $\mathfrak{J} = H(\mathfrak{A})$, $\mathfrak{A}$ a simple algebra with involution such that the involution is not the identity automorphism on the center of $\mathfrak{A}$. $\mathfrak{U} = \mathfrak{A}$.

A_2: $\mathfrak{J} = \mathfrak{A}_j$, $\mathfrak{A}$ a central simple associative algebra.
$\mathfrak{U} = \mathfrak{A} \oplus \mathfrak{A}'$, $\mathfrak{A}'$ anti-isomorphic to $\mathfrak{A}$.

The algebras that become types B or C are of the form $\mathfrak{J} = H(\mathfrak{A})$, $\mathfrak{A}$ simple involutorial with the involution acting as the identity automorphism on the center. $\mathfrak{U} = \mathfrak{A}$.

Algebras of type D over an arbitrary base field are as described above for an algebraically closed base field. $\mathfrak{U} = C$, the Clifford algebra.

Algebras of type E need not be algebras of all 3×3 hermitian matrices if the base field is not algebraically closed, according to recent unpublished work of A. A. Albert.

Section 1. Operator Commutativity. We will consider Jordan algebras over fields of characteristic different from 2; this assumption on characteristic will be made throughout this paper.

The concept of Operator-Commutativity, introduced by Jordan, von Neumann and Wigner in [**17**], is the natural analogue of the concept of commutativity of two elements in an associative algebra, as some of the following propositions will show. Some of these results were proved for characteristic zero in [**8**].

Let $\mathfrak{J}$ be a Jordan algebra, x and y elements of $\mathfrak{J}$. Let $x \circ y$ denote their product and write x^2 for $\frac{1}{2}(x \circ x)$. We denote by R_x the linear transformation $a \rightarrow a \circ x$ acting in $\mathfrak{J}$.

DEFINITION. Two elements x, y of $\mathfrak{J}$ operator-commute (we will write also: o-commute) if $R_x R_y = R_y R_x$.

The set of elements of $\mathfrak{J}$ o-commuting with x will be denoted by $\mathfrak{C}_{\mathfrak{J}}(x)$. If $\mathfrak{K}$ is a subalgebra of $\mathfrak{J}$, the set of elements of $\mathfrak{J}$ o-commuting with all elements of $\mathfrak{K}$ will be denoted by $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{K})$.

If $\mathfrak{A}$ is an associative algebra and $\mathfrak{B}$ an associative subalgebra, we will write $\mathfrak{C}_{\mathfrak{A}}(\mathfrak{B})$ for the subalgebra of elements of $\mathfrak{A}$ commuting with the elements of $\mathfrak{B}$. $\mathfrak{C}_{\mathfrak{A}}(\mathfrak{A})$ is the center of $\mathfrak{A}$—its elements will be called "central".

The following example, due to McCoy [**14**], and Jacobson [**8**], shows that the set $\mathfrak{C}_{\mathfrak{J}}(x)$ of elements of $\mathfrak{J}$ o-commuting with an element x of

$\mathfrak{J}$ is not necessarily a subalgebra of $\mathfrak{J}$: $\mathfrak{J}$ will be a special Jordan algebra (product $x \circ y = xy + yx$) consisting of 6×6 matrices whose coefficients are rational numbers, namely the Jordan algebra over the field of rational numbers generated by the following matrices $a, b, 1$ (using e_{kl} to denote the matrix with 1 in the k, l position and zeros elsewhere, $k, l = 1, \ldots, 6$): $a = e_{14} + e_{25} + 2e_{16}$, $b = e_{42} + e_{53} + e_{65}$, and $1 = e_{11} + \cdots + e_{66}$. Let $c = ab - ba$, then $c = e_{12} + e_{23} + e_{45}$ and $c^2 = e_{13} \neq 0$. Also $ac = ca$, $bc = cb$, so that c commutes with every polynomial in a and b. Consider now $\mathfrak{C}_{\mathfrak{J}}(a)$: we claim $\mathfrak{C}_{\mathfrak{J}}(a)$ contains b but not b^2. The equation $(a \circ x) \circ b = a \circ (x \circ b)$ becomes, on replacing $y \circ z$ by $yz + zy$, $[[a, b], x] = 0$ where $[y, z]$ denotes $yz - zy$. But $[a, b] = c$ and c clearly commutes with every element of the Jordan algebra generated by a and b, so b o-commutes with a. We compute $[a, b^2]$ and show this element does not commute with every x in $\mathfrak{J}$:

$$[a, b^2] = [a, b]b + b[a, b] = bc + cb = 2bc = 2[a, b]b$$
$$[[a, b^2], a] = 2[bc, a] = 2[b, a]c + 2b[c, a] = -2[a, b]c = -2c^2 \text{ as } [a, c] = 0$$

but $c^2 \neq 0$, and $2c^2 \neq 0$, thus b^2 is not in $\mathfrak{C}_{\mathfrak{J}}(a)$. We note also that b does not o-commute with a^2 and that $[a, b]^2 \neq 0$.

In the preceding example, b did not o-commute with all elements of the subalgebra generated by a and 1. One may ask whether $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{K})$ is a subalgebra if $\mathfrak{K}$ is a subalgebra. This is so in many cases, as the following propositions will show, and we conjecture that it is true in general. Some of the following results were proved in [8] for characteristic zero.

We say a finite dimensional Jordan algebra $\mathfrak{J}$ over a field F is separable if it is semi-simple and the algebra $\mathfrak{J} \otimes_F E$ obtained by extending the base field to E is also semi-simple for any extension field E of F. We note at this time a few simple facts about the effect of field extension on centralizers: if x and y are elements of $\mathfrak{J}$, they o-commute in $\mathfrak{J}$ if and only if they o-commute in $\mathfrak{J} \otimes E$; also, since the equation $R_a R_b = R_b R_a$ expressing that two elements a and b o-commute is linear in each, it follows that if $\mathfrak{K}$ is a subalgebra of $\mathfrak{J}$, then $\mathfrak{C}_{\mathfrak{J} \otimes E}(\mathfrak{K} \otimes E) = \mathfrak{C}_{\mathfrak{J}}(\mathfrak{K}) \otimes E$, and this allows us to extend the base field in many of our arguments.

PROPOSITION 1.1. Let $\mathfrak{J}$ be any Jordan algebra (possibly infinite dimensional), $\mathfrak{K}$ a separable subalgebra. Then $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{K})$ is a subalgebra of $\mathfrak{J}$.

Proof. First we show that if e is an idempotent in $\mathfrak{J}$, then $\mathfrak{C}_{\mathfrak{J}}(e)$ is a subalgebra. Let $\mathfrak{J} = \mathfrak{J}_0 + \mathfrak{J}_{1/2} + \mathfrak{J}_1$ be the Peirce decomposition of $\mathfrak{J}$ relative to e, i.e. $\mathfrak{J}_i = \{x \in J \mid e \circ x = ix\}$, $i = 0, 1/2, 1$. (We will also write $\mathfrak{J}_e(i)$ for $\mathfrak{J}_i$.) $\mathfrak{J}_0$ and $\mathfrak{J}_1$ are subalgebras, and $\mathfrak{J}_0\mathfrak{J}_1 = 0$, $\mathfrak{J}_0\mathfrak{J}_{1/2} \subseteqq \mathfrak{J}_{1/2}$, $\mathfrak{J}_1\mathfrak{J}_{1/2} \subseteqq \mathfrak{J}_{1/2}$. From these relations it follows that $\mathfrak{C}_\mathfrak{J}(e) = \mathfrak{J}_0 + \mathfrak{J}_1$: let x be in $\mathfrak{C}_\mathfrak{J}(e)$, and let $x = x_0 + x_{1/2} + x_1$, x_i in $\mathfrak{J}_i$. Then $e \circ (e \circ x) = x \circ (e \circ e) = x \circ e$, but $x \circ e = (1/2)x_{1/2} + x_1$, and $e \circ (x \circ e) = (1/4)x_{1/2} + x_1$, so $x_{1/2} = 0$ and x is in $\mathfrak{J}_0 + \mathfrak{J}_1$. Conversely, let x be in $\mathfrak{J}_0$, $z = z_0 + z_{1/2} + z_1$ in $\mathfrak{J}$. $x \circ (e \circ z) = x \circ [(1/2)z_{1/2} + z_1] = 1/2(x \circ z_{1/2})$, whereas $e \circ (x \circ z) = e \circ (x \circ z_0 + x \circ z_{1/2}) = 1/2(x \circ z_{1/2})$ since $x \circ z_0$ is in $\mathfrak{J}_0$, $x \circ z_{1/2}$ is in $\mathfrak{J}_{1/2}$; thus x o-commutes with e, and similarly if x is in $\mathfrak{J}_1$ it o-commutes with e. On extending the base field F of $\mathfrak{J}$ to its algebraic closure, $\mathfrak{K}$ remains semi-simple, and we will conclude the proof by showing that a semi-simple Jordan algebra over an algebraically closed base field has a basis consisting of idempotents—for if $e_1, \ldots, e_n$ are this basis, then $\mathfrak{C}_\mathfrak{J}(\mathfrak{K})$ is the intersection of all the $\mathfrak{C}_\mathfrak{J}(e_i)$ and the latter are subalgebras. From the structure theory, it is known that $\mathfrak{K}$ is a direct sum of simple algebras, and each simple algebra is either of degree one, i.e. of the form $F \cdot e$, e an idempotent, or else is a vector space sum of algebras of degree two. An algebra of degree two has a basis of elements $e_1, e_2, x_1, \ldots, x_n$ where the e_i are idempotents with $e_1 + e_2 = e$ being the identity element of the algebra, and $x_i \circ x_i = e$. Then $1/2(e + x_i)$ is also an idempotent, and $e_1, e_2, 1/2(e + x_1), \ldots, 1/2(e + x_n)$ is a basis consisting of idempotents, which we had to show.

The next proposition shows that for a large class of special Jordan algebras, including all that we will be concerned with in later sections, the set of elements o-commuting with the elements of a subalgebra is the same as the set of elements commuting with them in the associative multiplication. In particular, if $\mathfrak{J} = \mathfrak{A}_j$ is such an algebra, where $\mathfrak{A}$ is associative, and $\mathfrak{K} = \mathfrak{B}_j$, where $\mathfrak{B}$ is an associative subalgebra of $\mathfrak{A}$, then $\mathfrak{C}_\mathfrak{J}(\mathfrak{K}) = [\mathfrak{C}_\mathfrak{A}(\mathfrak{B})]_j$.

PROPOSITION 1.2. *Let $\mathfrak{J}$ be a special Jordan algebra with enveloping associative algebra $\mathfrak{A}$, and assume $\mathfrak{A}$ has no central nilpotent elements (e.g. $\mathfrak{A}$ any semi-simple algebra). Let $\mathfrak{K}$ be a Jordan subalgebra of $\mathfrak{J}$, y an element of $\mathfrak{J}$. Then y is in $\mathfrak{C}_\mathfrak{J}(\mathfrak{K})$ if and only if $xy = yx$ for all $x \in \mathfrak{K}$. Thus $\mathfrak{C}_\mathfrak{J}(\mathfrak{K})$ consists of all elements of $\mathfrak{J}$ commuting in the associative multiplication with the elements of $\mathfrak{K}$, and $\mathfrak{C}_\mathfrak{J}(\mathfrak{K})$ is a subalgebra of $\mathfrak{J}$.*

Proof. We make extensive use of the assumption that $2 \neq 0$. Also, we note that the equation $z(R_xR_y - R_yR_x) = 0$ is equivalent to $[[xy]z] = 0$, where $[ab] = ab - ba$. Let now $x \in \mathfrak{K}$, $y \in \mathfrak{C}_\mathfrak{J}(\mathfrak{K})$. Since $\mathfrak{K}$ is a subalgebra, $x \circ x = 2x^2 \in \mathfrak{K}$, and $x^2 \in \mathfrak{K}$ so that $[[xy]z] = 0$ and $[[x^2y]z] = 0$ for all $z \in \mathfrak{J}$. Since $\mathfrak{A}$ is the enveloping algebra of $\mathfrak{J}$, $[xy]$ and $[x^2y]$ are in the center of $\mathfrak{A}$. But $[x^2y] = [xy]x + x[xy] = 2x[xy]$ since $[xy]$ is in the center of $\mathfrak{A}$. $[x^2y] \in$ the center of $\mathfrak{A}$ implies that

$$0 = [y[x^2y]] = 2[y, (x[xy])] = 2[yx][xy] + 0 = -2[xy]^2 .$$

Thus $[xy]^2 = 0$, so by hypothesis on $\mathfrak{A}$, $[xy] = 0$. Conversely, if $[xy] = 0$, then $[[xy]z] = 0$ so that x o-commutes with y. The statements on $\mathfrak{C}_{\mathfrak{Z}}(\mathfrak{K})$ now follow.

Let $\mathfrak{A}$ be a possibly non-associative algebra with multiplication denote by xy. We again introduce the new multiplication $x \circ y = xy + yx$. Let $\mathfrak{B}$ be a subspace of $\mathfrak{A}$ closed under $x \circ y$, then we denote by R_y the operation $x \rightarrow x \circ y = xR_y$ acting in $\mathfrak{B}$, where x, y are elements of $\mathfrak{B}$. As before we say x and y o-commute if $R_xR_y = R_yR_x$. We can make a remark in case $\mathfrak{A}$ is a matrix algebra with canonical involution and $\mathfrak{B}$ is the set of self-adjoint elements (see [9]), i.e. $\mathfrak{A}$ is a matrix algebra D_n over an algebra D with identity 1 and involution $d \rightarrow \bar{d}$ in D, and $\mathfrak{A}$ has involution $a = \sum_{i,j} d_{ij}e_{ij} \rightarrow \sum_{i,j} \gamma_j^{-1}\bar{d}_{ij}\gamma_i e_{ji}$ where the e_{ij} are matrix units, $\gamma_1 = 1, \gamma_2, \ldots, \gamma_n$ are self-adjoint elements of the nucleus of D having inverses, and $n \geqq 2$; $\mathfrak{B}$ is the set of self-adjoint matrices, denoted by $H(D_n)$. Such algebras have been studied in [9], and include all simple Jordan algebras of degree greater than two over algebraically closed fields.

LEMMA. *Let $\mathfrak{A} = D_n$ be a matrix algebra with canonical involution, $\mathfrak{B} = H(D_n)$, x, y elements of $\mathfrak{B}$. If x and y o-commute, then $[xy] = d \cdot I$, where d is a skew element of D, and I is the unit matrix.*

Proof. Since x and y are self-adjoint, $[xy]$ is skew. $\mathfrak{B}$ contains the elements e_{jj} and $d[i, j] = de_{ij} + \gamma_j^{-1}\bar{d}\gamma_i e_{ji}$ for $i \neq j$. Since x and y o-commute, we have $(e_{ii} \circ x) \circ y = (e_{ii} \circ y) \circ x$, which is equivalent to $[e_{ii}[xy]] = 0$, $i = 1, \ldots, n$, since e_{ii} is in the nucleus of $\mathfrak{A}$. The matrix $[xy]$ thus has zeros off the main diagonal. The elements $1[ij] = e_{ij} + \gamma_j^{-1}\gamma_i e_{ji}$ are in the nucleus since the γ_i and e_{ij} are, so that

$$[(e_{ij} + \gamma_j^{-1}\gamma_i e_{ji}), [xy]] = 0 , \quad i \neq j$$

or

$$e_{ij}[xy] + \gamma_j^{-1}\gamma_i e_{ji}[xy] = [xy]e_{ij} + [xy]\gamma_j^{-1}\gamma_i e_{ji} .$$

Denote by $[xy]_{ij}$ the i, j entry of $[xy]$ for $i, j = 1, \ldots, n$. Then

$$[xy]_{jj}e_{ij} + \gamma_j^{-1}\gamma_i[xy]_{ii}e_{ji} = [xy]_{ii}e_{ij} + [xy]_{jj}\gamma_j^{-1}\gamma_i e_{ji} \quad \text{for } i \neq j .$$

Since the coefficients of e_{ij} in the above equation must be equal,

$$[xy]_{jj} = [xy]_{ii} = d \text{ for all } i, j .$$

Thus $[xy] = dI$, $\bar{d} = -d$, which completes the proof.

From this lemma we can derive conditions for the centralizer of a subalgebra of $H(D_n)$ to consist of the matrices commuting with the matrices of the subalgebra. We denote by Z the set of elements of D commuting with every element of D, and by N (nucleus) the set of elements of D associating with every pair of elements of D.

PROPOSITION 1.3. Let $\mathfrak{B} = H(D_n)$, $n \geqq 2$.

(a) Let n not be divisible by the characteristic of F. If x and y are two matrices of $\mathfrak{B}$ such that the coefficients of x commute with those of y and if x and y o-commute, then they commute. In particular if x has coefficients in Z and y o-commutes with x, then $xy = yx$.

(b) Let $\mathfrak{K}$ be a subalgebra of $H(D_n)$ such that every element of $\mathfrak{K}$ has coefficients in N, the nucleus of D. If Z contains no skew-elements whose squares are zero then y is in $\mathfrak{C}_{\mathfrak{B}}(\mathfrak{K})$ if and only if $xy = yx$ for all x in $\mathfrak{K}$, so $\mathfrak{C}_{\mathfrak{B}}(\mathfrak{K})$ is a subalgebra. In particular, if D is associative with no central skew elements of square zero, the conclusion holds for any subalgebra $\mathfrak{K}$ of $H(D_n)$. If $\mathfrak{B}$ is an exceptional simple Jordan algebra, then D is a Cayley algebra and $Z = N = F \cdot 1$, F the base field, so that the conclusion holds for subalgebras $\mathfrak{K}$ whose elements have coefficients in F.

Proof. (a) By the lemma, if y o-commutes with x, then $[xy] = dI$, $d \in D$, $I =$ identity matrix. Let us now take the trace of the elements on each side of this equation. If the coefficients of x commute with those of y, then $tr(xy) = tr(yx)$ so $tr([xy]) = 0 = tr(dI) = nd$. Since $nd = 0$ implies $d = 0$, $xy = yx$.

(b) Let every x in $\mathfrak{K}$ have coefficients in N. Then $(xy)z = x(yz)$ for all $y, z \in D_n$. Let now $y \in \mathfrak{C}_{\mathfrak{B}}(\mathfrak{K})$. Then $(x \circ z) \circ y = (y \circ z) \circ x$ for all z in $\mathfrak{B}$, which is equivalent to: $[[xy]z] = 0$ for all z in $\mathfrak{B}$. $[xy] = dI$, so $dz_{ij} = z_{ij}d$ for each coefficient z_{ij} of z. Since $n \geqq 2$, for any $a \in D$, there is a z in $\mathfrak{B}$ with $z_{12} = a$. Thus d is in Z. Since x^2 is in $\mathfrak{K}$ also, $[x^2y] = fI$ where f is a skew element of Z. The calculation of Proposition 1.2 is still valid since x and x^2 associate with any two elements of $\mathfrak{B}$, and we conclude that $d^2 = 0$ and so $d = 0$, or $xy = yx$. The remaining statements of the proposition are now obvious.

PROPOSITION 1.4. Let $\mathfrak{M}$ be an exceptional central simple Jordan algebra, and $\mathfrak{K}$ a separable subalgebra of $\mathfrak{M}$ containing the identity. Then $\mathfrak{C}_{\mathfrak{M}}(\mathfrak{K})$ is separable, and, if the base field is algebraically closed, $\mathfrak{M}$ can be represented as an $H(C_3)$ such that $\mathfrak{C}_{\mathfrak{M}}(\mathfrak{K})$ consists of the matrices in $\mathfrak{M}$ commuting with those in $\mathfrak{K}$.

Proof. We first assume the base field F algebraically closed, and $\mathfrak{K}$ semi-simple. $\mathfrak{M}$ is then of degree 3, i.e. if $1 = e_1 + \cdots + e_r$, e_i primitive mutually orthogonal idempotents in $\mathfrak{M}$, then $r = 3$, and $\mathfrak{M}$ is a vector space sum $\sum_{i,j}\mathfrak{M}_{ij},\ i \leqq j,\ i,j = 1, 2, 3$ where

$$\mathfrak{M}_{ij} = \{x \in M \mid e_i \circ x = (1/2)x = e_j \circ x\} \text{ for } i \neq j, \text{ and}$$
$$\mathfrak{M}_{ii} = \{x \in M \mid e_i \circ x = x\} = Fe_i\ .$$

These facts limit $\mathfrak{K}$ to a few possibilities:

1. $\mathfrak{K}$ not simple: then $\mathfrak{K} = \mathfrak{K}_1 \oplus \cdots \oplus \mathfrak{K}_r$, $\mathfrak{K}_i$ simple and $r = 2$ or 3, since the identity element 1 of $\mathfrak{M}$ is also in $\mathfrak{K}$, and $1 = u_1 + \cdots + u_r$, u_i the identities of $\mathfrak{K}_i$.

(a) $r = 3$. Then u_1, u_2, u_3 are primitive (in $\mathfrak{M}$) orthogonal idempotents and each $\mathfrak{K}_i$ is of degree 1 so that we may take $e_i = u_i$. We now represent $\mathfrak{M}$ as $H(C_3)$ with $e_i = (1/2)e_{ii}$. We have $\mathfrak{K}_i = Fe_i$ and $\mathfrak{C}_{\mathfrak{M}}(\mathfrak{K}) = \cap_i \mathfrak{C}_{\mathfrak{M}}(e_i)$. But $\mathfrak{C}_{\mathfrak{M}}(e_i) = \mathfrak{M}_{ii} + \mathfrak{M}_{jj} + \mathfrak{M}_{jk} + \mathfrak{M}_{kk}$, where $j, k \neq i$ since with respect to e_i, $\mathfrak{M}_1 = \mathfrak{M}_{ii}$ and $\mathfrak{M}_0 = \sum_{j,k\neq i}\mathfrak{M}_{jk}$. Hence $\mathfrak{C}_{\mathfrak{M}}(\mathfrak{K}) = \mathfrak{M}_{11} + \mathfrak{M}_{22} + \mathfrak{M}_{33} = \mathfrak{K}$. It is clear that the matrices of $\mathfrak{C}_{\mathfrak{M}}(\mathfrak{K})$ commute with those of $\mathfrak{K}$. Conversely, any matrix commuting with e_1, e_2, e_3 is a linear combination of e_1, e_2, e_3, and thus in $\mathfrak{C}_{\mathfrak{M}}(\mathfrak{K})$.

(b) $r = 2$. Then $\mathfrak{K} = \mathfrak{K}_1 \oplus \mathfrak{K}_2$ and $\mathfrak{K}_i$ has identity u_i. Since $1 = u_1 + u_2$, one of the u_i must be primitive in $\mathfrak{M}$, say u_1, and the other one not: $u_1 = e_1$, $u_2 = e_2 + e_3$, e_i primitive orthogonal idempotents. Again write $\mathfrak{M} = H(C_3)$, $(1/2)e_{ii} = e_i$. Here there are two cases

(i) $\mathfrak{K}_2$ is of degree one: $\mathfrak{K}_2 = Fu_2$. Since $\mathfrak{K}_1 = Fu_1$,

$$\begin{aligned}\mathfrak{C}_{\mathfrak{M}}(\mathfrak{K}) &= \mathfrak{C}_{\mathfrak{M}}(u_1) \cap \mathfrak{C}_{\mathfrak{M}}(u_2)\\ &= (\mathfrak{M}_{11} + \mathfrak{M}_{22} + \mathfrak{M}_{23} + \mathfrak{M}_{33}) \cap (\mathfrak{M}_{11} + \mathfrak{M}_{22} + \mathfrak{M}_{23} + \mathfrak{M}_{33})\\ &= \mathfrak{M}_{11} + \mathfrak{M}_{22} + \mathfrak{M}_{23} + \mathfrak{M}_{33}\ .\end{aligned}$$

But $\mathfrak{M}_{22} + \mathfrak{M}_{23} + \mathfrak{M}_{33}$ is a simple Jordan algebra $\mathfrak{M}'_{11}$ of degree two, so that $\mathfrak{C}_{\mathfrak{M}}(\mathfrak{K}) = \mathfrak{M}_{11} \oplus \mathfrak{M}'_{11}$ is semi-simple. The matrices of $\mathfrak{C}_{\mathfrak{M}}(\mathfrak{K})$ are evidently just those matrices commuting with e_1 and $e_2 + e_3$ and therefore with the matrices of $\mathfrak{K}$.

(ii) $\mathfrak{K}_2$ is of degree two: $u_2 = e_2 + e_3$ where both e_2 and e_3 are in $\mathfrak{K}_2$. Then $\mathfrak{K}_2$ is a simple Jordan algebra of degree two, and so contains an element a with $a \circ a = 4(e_2 + e_3)$, $a \in \mathfrak{M}_{23}$. Since e_1, e_2, e_3 all belong to $\mathfrak{K}$, $\mathfrak{C}_{\mathfrak{M}}(\mathfrak{K}) \subseteqq Fe_1 + Fe_2 + Fe_3$. Also, since $\mathfrak{C}_{\mathfrak{M}}(e_1) = \mathfrak{M}_{11} + \mathfrak{M}_{22} + \mathfrak{M}_{23} + \mathfrak{M}_{33} \supseteqq \mathfrak{K}$, we see that Fe_1 o-commutes with $\mathfrak{K}$, i.e. $Fe_1 \subseteqq \mathfrak{C}_{\mathfrak{M}}(\mathfrak{K})$. Let now $\alpha e_2 + \beta e_3$ belong to $\mathfrak{C}_{\mathfrak{M}}(\mathfrak{K})$: then this element o-commutes with a, so

$$\begin{aligned}[e_2 \circ (\alpha e_2 + \beta e_3)] \circ a &= \alpha e_2 \circ a = (1/2)\alpha a = (e_2 \circ a) \circ (\alpha e_2 + \beta e_3)\\ &= (1/2)a \circ (\alpha e_2 + \beta e_3) = (1/4)(\alpha + \beta)a\ .\end{aligned}$$

Thus $\alpha = \beta$, i.e. $\mathfrak{C}_{\mathfrak{M}}(\mathfrak{K}) \subseteqq Fe_1 + F(e_2 + e_3)$. On the other hand, $(e_2 + e_3) \in \mathfrak{C}_{\mathfrak{M}}(\mathfrak{K})$ since $\mathfrak{K} \subseteqq \mathfrak{M}_{11} + \mathfrak{M}'_{11}$, so $\mathfrak{C}_{\mathfrak{M}}(\mathfrak{K}) = Fe_1 + F(e_2 + e_3)$, and evidently these matrices commute with the matrices in $\mathfrak{K}$. Conversely, let $x = \sum_{i\leqq j}x_{ij}$, $x_{ij} \in \mathfrak{M}_{ij}$, commute with the matrices of $\mathfrak{K}$. Since e_1, e_2, e_3

are in $\mathfrak{K}$, $x = \alpha e_1 + \beta e_2 + \gamma e_3$. Letting e_{ij}, $i \neq j$ denote the matrix units, the element a of $\mathfrak{K}$ can be written as $a = de_{23} + \bar{d}e_{32}$, d an element of C, and $a \circ a = 4(d\bar{d}e_2 + \bar{d}de_3) = 4(e_2 + e_3)$, thus $d\bar{d} = \bar{d}d = 1$.

$$ax = (de_{23} + \bar{d}e_{32})(\alpha e_1 + \beta e_2 + \gamma e_3) = (1/2)(d\gamma e_{23} + \bar{d}\beta e_{32}) ; \quad \alpha, \beta, \gamma \in F .$$

$$xa = (\alpha e_1 + \beta e_2 + \gamma e_3)(de_{23} + \bar{d}e_{32}) = (1/2)(\beta de_{23} + \bar{\gamma}de_{32}) .$$

Since we are supposing $ax = xa$, we have $\beta d = d\gamma$, but $d\gamma = \gamma d$ as $\gamma \in F$, so $\beta = \gamma$, i.e. $x = \alpha e_1 + \beta(e_2 + e_3)$ and x is in $\mathfrak{C}_{\mathfrak{M}}(\mathfrak{K})$.

2. $\mathfrak{K}$ is simple; then $\mathfrak{K}$ has degree one, two, or three

(a) $\mathfrak{K}$ has degree one: $\mathfrak{K} = F \cdot 1$. Then $\mathfrak{C}_{\mathfrak{M}}(\mathfrak{K}) = \mathfrak{M}$, and the matrices of $\mathfrak{M}$ commuting with those of $\mathfrak{K}$ are just the elements of $\mathfrak{C}_{\mathfrak{M}}(\mathfrak{K})$.

(b) $\mathfrak{K}$ has degree two: we show this is impossible. For let $1 = u_1 + u_2$, u_i primitive idempotents in $\mathfrak{K}$. Then, as in case 1b, we may assume u_1 is primitive in $\mathfrak{M}$ and u_2 not, so that $u_1 = e_1$, $u_2 = e_2 + e_3$. Since $\mathfrak{K}$ is simple of degree two, it contains an element x such that $x \circ x = 2x^2 = u_1 + u_2$ and $u_1 \circ x = (1/2)x = u_2 \circ x$. Let $x = \sum_{i \leq j} x_{ij}$, x_{ij} in $\mathfrak{M}_{ij}$. Since $e_1 \circ x = (1/2)x$, $x = x_{12} + x_{13}$, therefore $x \circ x = x_{12} \circ x_{12} + x_{13} \circ x_{13} + 2x_{12} \circ x_{13}$. But

$$x_{12} \circ x_{12} = \lambda_1(e_1 + e_2),\ x_{13} \circ x_{13} = \lambda_2(e_1 + e_3),\ \lambda_i \in F,\ x_{12} \circ x_{13} \in \mathfrak{M}_{23}$$

so $x \circ x = (\lambda_1 + \lambda_2)e_1 + \lambda_1 e_2 + \lambda_2 e_3 + 2x_{12} \circ x_{13} = u_1 + u_2 = e_1 + e_2 + e_3$. Since $\mathfrak{M}$ is a direct sum of the $\mathfrak{M}_{ij}$, $2x_{12} \circ x_{13} = 0$, $\lambda_1 + \lambda_2 = 1$, $\lambda_1 = 1$, $\lambda_2 = 1$, a contradiction.

(c) $\mathfrak{K}$ has degree three: Then $\mathfrak{K}$ contains idempotents e_1, e_2, e_3 and the subspaces $\mathfrak{K}_{ij} = \mathfrak{K} \cap \mathfrak{M}_{ij}$ are all non-zero. $\mathfrak{K}$ contains the subalgebra $\mathfrak{K}_{11} + \mathfrak{K}_{22} + \mathfrak{K}_{23} + \mathfrak{K}_{33}$ of the type considered in 1b (ii), and the centralizer of this subalgebra, as well as the set of matrices commuting with its elements, is $Fe_1 + F(e_2 + e_3)$. Arguing the same way but replacing the index 1 by 2 and then 3, we see that $\mathfrak{C}_{\mathfrak{M}}(\mathfrak{K}) = F(e_1 + e_2 + e_3) = F \cdot 1$, and that any matrix commuting with the elements of $\mathfrak{K}$ is in $F \cdot 1$. Conversely, the matrices in $F \cdot 1$ obviously commute with those in $\mathfrak{K}$. We have shown that in each case $\mathfrak{C}_{\mathfrak{M}}(\mathfrak{K})$ is semi-simple.

If now F is not necessarily algebraically closed, the centralizer of $\mathfrak{K}$ in $\mathfrak{M}$ remains semi-simple on extending F to its algebraic closure, so that $\mathfrak{C}_{\mathfrak{M}}(\mathfrak{K})$ is separable. Also, in the algebraically closed case, we have shown that for every $\mathfrak{K}$ there is a matrix representation of $\mathfrak{M}$ such that the elements of $\mathfrak{C}_{\mathfrak{M}}(\mathfrak{K})$ are represented by the matrices commuting with the matrices representing the elements of $\mathfrak{K}$. This concludes the proof.

2. Subalgebras Generated by a Single Element. In this section we study the centralizer and double centralizer in a central simple Jordan algebra of the subalgebra generated by one element, and generalize to Jordan algebras some of the known results on simple associative algebras contained in, say, [6], Chapter 3.

If x is an element of $\mathfrak{J}$, we denote by (x) the subalgebra generated by x and 1. The following facts are known about $\mathfrak{C}_{\mathfrak{J}}((x))$ if $\mathfrak{J} = \mathfrak{A}_j$ and $\mathfrak{A}$ is the associative algebra of all $n \times n$ matrices over a field F:

1. Let x have invariant factors $\delta_1, \cdots, \delta_r$ of respective degrees $d_1 \geqq d_2 \geqq \cdots \geqq d_r$. Let $(F[\lambda])_r$ be the algebra of all $r \times r$ matrices with coefficients polynomials in an indeterminate λ, (δ) the matrix diag $(\delta_1(\lambda), \cdots, \delta_r(\lambda))$ and $\mathfrak{B}$ the subalgebra of matrices (α) in $(F[\lambda])_r$ such that $(\alpha)'(\delta) = (\delta)(\beta)$ for some (β) in $(F[\lambda])_r$, which condition we will also write as $(\delta)^{-1}(\alpha)'(\delta) = (\beta) \in (F[\lambda])_r$, $(\alpha)'$ denoting the transpose of (α). Let $\mathfrak{R}$ be the ideal in $\mathfrak{B}$ of matrices of the form $(\beta)(\delta)$. Then $\mathfrak{C}_{\mathfrak{A}}((x))$ is isomorphic to $\mathfrak{B}/\mathfrak{R}$, so that by the results of section 1, $\mathfrak{C}_{\mathfrak{J}}((x))$ is isomorphic to $\mathfrak{B}_j/\mathfrak{R}_j$. From 1. easily follows the following theorem of Frobenius:

2. Let x be as above with invariant factor degrees $d_1 \geqq \cdots \geqq d_r$. Then $\mathfrak{C}_{\mathfrak{A}}((x))$ is of dimension $\sum_{k=0}^{r-1}(2k+1)d_{k+1}$.

3. Finally, $\mathfrak{C}_{\mathfrak{A}}(\mathfrak{C}_{\mathfrak{A}}((x))) = (x)$.

We will give appropriate generalizations of these results for Jordan algebras. The method will be to examine one by one the various types of simple algebras. In this way we will obtain some information on each of the various types of simple Jordan algebras, but it would also be interesting to have a general method which works for all the algebras at once.

Since all the special algebras $\mathfrak{J}$ that occur will have semi-simple enveloping associative algebras, we will be able to apply Proposition 1.2 to obtain that $\mathfrak{C}_{\mathfrak{J}}((x))$ is just the set of elements of $\mathfrak{J}$ commuting with x. Proposition 1.3b will be used in a similar way for the exceptional algebras.

In most proofs the key case will be that of a nilpotent element x, and the reduction to that case will be made by using a special case of the decomposition of an associative algebra over a perfect field into a direct sum of a semi-simple algebra and the radical: if we consider the associative commutative subalgebra (x) of $\mathfrak{J}$ and decompose it, then we can write $x = s + n$, s a linear combination of orthogonal idempotents e_i in (x) and n a sum of nilpotent elements n_i in (x) such that $e_i n_i = n_i$, $e_j n_i = 0$ for $j \neq i$.

The known theorems about the algebra of all linear transformations quoted above are best proved by considering the vector space V on which the given transformation x acts as a module over a principal ideal ring and decomposing it into a direct sum of cyclic modules. If we are interested in vector spaces with inner product we have to try to obtain orthogonal decompositions into cyclic subspaces. This is done in the first two lemmas (which are essentially known, see e.g. [13], last chapter, for algebraically closed fields, but hard to find in the literature in a form useful for us).

Let V be a finite dimensional left vector space over an involutorial division ring D. We will use the letters $x, y, \cdots$ for elements of V, α, $\beta, \cdots$ for scalars in D, $\alpha \to \alpha'$ denoting the involution, and $A, N, \cdots$ for linear transformations. We will assume V is a self-dual space, i.e. it has a non-degenerate scalar product (x, y) which is either hermitian: $(y, x) = (x, y)'$, $(\alpha x, \beta y) = \alpha(x, y)\beta'$, or else alternate: $(x, y) = -(y, x)$ and D is a field with $\alpha' = \alpha$. V will be called hermitian or alternate, respectively. The characteristic is assumed to be different from two.

LEMMA 2.1. *Let V be a finite dimensional self-dual hermitian space over D, and F the subfield of self-adjoint elements of D. Let A be a self-adjoint linear transformation which is algebraic over F,* i.e. *A satisfies a polynomial equation over F, and assume also that if $D \neq F$ then the roots of this equation lie in F. Then V can be decomposed into a direct sum of mutually orthogonal non-isotropic spaces which are indecomposable cyclic for A. If A is nilpotent, then each cyclic subspace has a basis of the form $\{x, xA, \cdots, xA^{n-1}\}$ such that $(xA^i, xA^j) = 0$ if $i + j \neq n - 1$, and $(xA^i, xA^{n-1-i}) = \mu$ for all i, μ a self-adjoint non-zero element of D.*

Proof. If the minimum polynomial of A over F has distinct prime factors, then corresponding to these there exist mutually orthogonal idempotents E_i with sum 1 which are polynomials in A with coefficients in F, and so are self-adjoint. Then $V = \sum \oplus VE_i$, and $(VE_i, VE_j) = (VE_iE_j, VE_j) = 0$ for $i \neq j$. Thus the $V_i = VE_i$ are mutually orthogonal and so are themselves hermitian self-dual spaces, and the linear transformation induced by A on each of these satisfies a primary polynomial $p(\lambda)^n$ over F. Thus it is sufficient to assume the minimum polynomial of A over F is of the form $p(\lambda)^n$, and by assumption $p(\lambda) = \lambda - \alpha$, α in F, if D is not equal to F. Moreover, if $p(\lambda) = \lambda - \alpha$, then it is sufficient to prove the first statement of the lemma for $A - \alpha 1$ instead of A, that is, we may assume A nilpotent if D is not F. In what follows we will often write p or p^r for $p(\lambda)$ or $p(\lambda)^r$, and if $m(\lambda)$ is a polynomial over F in the indeterminate λ, we will write $zm(\lambda)$ or zm for $zm(A)$, if z is in V.

If x is a vector of order p^r, that is, $xp(A)^r = 0$ and $xm(A) = 0$ implies that p^r divides m, then the vectors $xA^0 = x$, xA^s, $0 < s <$ degree of p^r, form a basis for the cyclic subspace generated by x. This is well known if $D = F$, and if $D \neq F$, then $p(\lambda) = \lambda$ and so $xA^r = 0$, $xA^{r-1} \neq 0$, and it is easy to check that the vectors $x, xA, \cdots, xA^{r-1}$ are linearly independent over D. We also note that $p(\lambda)^n$, the minimum polynomial of A over F, is also the minimum polynomial of A with coefficients in D, since if $p(\lambda)^n = \lambda^n$, the only monic factors of $p(\lambda)^n$ are the polynomials λ^r which have coefficients in F. From now on all polynomials we will consider will be assumed to have coefficients in F.

Since $p(\lambda)^n$ is the minimum polynomial of A over D, there is an x in V of order p^n. Let U be the cyclic subspace generated by x, and suppose U is isotropic. We show that this implies $(xp^{n-1}, x) = 0$: first let $D = F$, and let $xq(A)$ be a non-zero vector in the radical of U. Then $(xq(A), xt(A)) = 0 = (xq(A)t(A), x)$ for all $t = t(\lambda)$. Since $xq(A) \neq 0$, there is a t such that $xq(A)t(A) = xp(A)^{n-1}$. If $D \neq F$, let $z = \sum_{i=1}^{n-1}\gamma_i xA^i$ be non-zero and in the radical of U. Suppose $\gamma_k \neq 0$, and $\gamma_j = 0$ for $j < k$. Then $0 = (z, xA^{n-1-k}) = \gamma_k(xA^k, xA^{n-1-k}) = \gamma_k(xA^{n-1}, x)$. We now show that we can find a non-isotropic cyclic subspace U of maximum order: once we have it, its orthogonal complement $U^{\perp}$ will also be invariant for A, and will satisfy the same hypotheses as A but will have lower dimension, and we can use induction. Therefore, suppose $(x, xp^{n-1}) = 0$. There exists y in V such that $(y, xp^{n-1}) = 1 = (yp^{n-1}, x)$. Thus $yp^{n-1} \neq 0$, and we may assume $(y, yp^{n-1}) = 0$, otherwise y generates the desired cyclic space. Consider now the cyclic subspace spanned by $x + y$:

$$\begin{aligned}(x + y, (x + y)p^{n-1}) &= (x, xp^{n-1}) + (y, yp^{n-1}) + (x, yp^{n-1}) + (y, xp^{n-1}) \\ &= 0 + 0 + 1 + 1 = 2 \neq 0 .\end{aligned}$$

Thus, in particular, $x + y$ has order p^n, and it generates a non-isotropic cyclic space of order p^n. This completes the proof of the first statement of the lemma.

Suppose A is nilpotent and V is cyclic of dimension n. Let $x, xA, \cdots, xA^{n-1}$ be a basis for V. Clearly $(xA^{n-1}, xA^i) = 0$ for $i \geqq 1$, so we may assume $(x, xA^{n-1}) = \mu \neq 0$. If $n = 1$, there is nothing to prove, so let $n \geqq 2$. Suppose k is some integer such that $0 \leqq k < n - 1$ and $(x, xA^r) = 0$ for all r satisfying $k < r < n - 1$. We show that x may be replaced by another cyclic generator y such that $(y, yA^r) = 0$ for $k - 1 < r < n - 1$: Let $y = x + \alpha xA^{n-1-k}$, α an element of D to be specified shortly. For $k < r \leqq n - 1$, $yA^r = \alpha xA^{n-1+r-k} + xA^r = xA^r$ since $A^n = 0$; in particular $yA^{n-1} \neq 0$ so y is a cyclic generator also. Further, for $k < r < n - 1$,

$$\begin{aligned}(y, yA^r) = (y, xA^r) &= (x + \alpha xA^{n-1-k}, xA^r) \\ &= (x, xA^r) + \alpha(x, xA^{n-1+r-k}) \\ &= (x, xA^r) \text{ since } r - k > 1 .\end{aligned}$$

Finally,

$$\begin{aligned}(y, yA^k) &= (x + \alpha x A^{n-1-k}, xA^k + \alpha x A^{n-1}) \\ &= (x, xA^k) + \alpha(x, xA^{n-1}) + (x, xA^{n-1})\alpha' + \alpha(xA^{n-1-k}, xA^{n-1})\alpha' .\end{aligned}$$

The last term is zero since $n - 1 - k \geqq 1$, and $(x, xA^{n-1}) = \mu \neq 0$. Let $\alpha = -(1/2)(x, xA^k)\mu^{-1}$, then $\alpha(x, xA^{n-1}) = \alpha\mu = -(1/2)(x, xA^k) = (x, xA^{n-1})\alpha'$ since $\mu' = \mu$, and so $(y, yA^k) = 0$ also. This completes the proof.

This last type of basis can be used to give a relation between the index of the hermitian form and the index of nilpotency of a nilpotent self-adjoint linear transformation. Let V have dimension m and index r, that is, r is the dimension of any maximal totally isotropic subspace, so that $m \geqq 2r$. Let N be nilpotent of index n, i.e. $N^n = 0$, $N^{n-1} \neq 0$, and self-adjoint. Then : $n \leqq \text{Min}\,(m, 2r + 1)$, (i.e. if $m = 2r$ then $n \leqq 2r$ and if $m > 2r$ then $n \leqq 2r + 1$), and equality is achieved for some N). To see this, choose a set of vectors $x, xN, \cdots, xN^{n-1}$ satisfying $(xN^j, xN^k) = 0$ for $j + k \neq n - 1$, $(xn^j, xN^{n-1-j}) \neq 0$. The vectors xN^j for $j < (n-1)/2$ then span a totally isotropic subspace of dimension $[n/2]$ (greatest integer $\leqq n/2$) so $[n/2] \leqq r$ and $n \leqq 2r + 1$. The condition $n \leqq m$ is always satisfied. Conversely, let V have index r. Then we can find vectors $x_1, \cdots, x_r$ and $y_1, \cdots, y_r$ spanning totally isotropic spaces respectively and satisfying $(x_i, y_j) = \gamma\delta_{ij}$, γ any preassigned element of D. First assume $m > 2r$; then we can find a vector z orthogonal to the space spanned by the x_i and y_i such that $(z, z) \neq 0$. We may assume $(x_i, y_i) = \gamma = (z, z)$ for all i. Now define a linear transformation N as follows :

$$\begin{aligned}&x_1N = x_2, x_2N = x_3, \cdots, x_{r-1}N = x_r, x_rN = z, \\ &\qquad\qquad zN = y_r, y_rN = y_{r-1}, \cdots, y_2N = y_1, y_1N = 0.\end{aligned}$$

Let U be the space spanned by the x_i, y_i, z; U is non-isotropic, so $U \oplus U^{\perp} = V$, $U^{\perp}$ denoting the orthogonal complement of U. Define N to be zero on $U^{\perp}$. It is clear that N is nilpotent of index $2r + 1$ with cyclic subspace U, and self-adjoint. If $m = 2r$, we merely omit z and define N by : $x_1N = x_2, \cdots, x_{r-1}N = x_r, x_rN = y_r, y_rN = y_{r-1}, \cdots, y_2N = y_1, y_1N = 0$, where $(x_i, y_j) = \delta_{ij}$.

Lemma 2.1 is a generalization of the result that a hermitian self-dual finite-dimensional space has an orthogonal basis, the one-dimensional orthogonal subspaces being replaced by cyclic subspaces of a self-adjoint linear transformation. We obtain an analogous generalization of a symplectic basis (i.e. $(x_i, y_j) = \delta_{ij} = -(y_j, x_i)$, $(x_i, x_j) = 0 = (y_i, y_j)$) for a symplectic space :

LEMMA 2.2. *Let V be a symplectic space and A a self-adjoint linear transformation in V. Then V can be decomposed into a direct sum of* *indecomposable cyclic subspaces* U_i, U'_i *such that* U_i *and* U'_i *are isomorphic as* (A) *modules and totally isotropic and* $U_i + U'_i$ *is non-isotropic and orthogonal to* $U_j + U'_j$ *if* $i \neq j$. *If* A *is nilpotent and* U, U' *are a pair of isomorphic cyclic subspaces with* $U + U'$ *non-isotropic, then we can find bases of the form* $\{x, xA, \cdots, xA^{r-1}\}$, $\{y, yA, \cdots, yA^{r-1}\}$ *for* U, U' *respectively such that* $(xA^i, yA^{r-1-i}) = 1 = -(yA^i, xA^{r-1-i})$, *and all other scalar products are zero.*

Proof. As in the hermitian case, we can immediately reduce the proof to the case of a primary minimum polynomial $p(\lambda)^n$ for A. We note that every cyclic subspace is totally isotropic, for $(xA^i, xA^j) = -(xA^j, xA^i)$ since V is symplectic, but $(xA^i, xA^j) = (xA^j, xA^i)$ since A is self-adjoint, so $(xA^i, xA^j) = 0$. This implies that if U and U' are cyclic subspaces such that $U + U'$ is non-isotropic, then their intersection contains just the zero vector, since any vector in the intersection is orthogonal to both U and U' and so is orthogonal to $U + U'$. As in the proof of Lemma 2.1, it will suffice to find isomorphic cyclic subspaces U, U' such that $U + U'$ is non-isotropic (and therefore the sum is automatically direct).

Let x be a vector of maximum order p^n in V, and let U be the cyclic subspace generated by x. xp^{n-1} is not zero, so there exists y in V such that $(xp^{n-1}, y) = 1$. Then also $(x, yp^{n-1}) = 1$, so $yp^{n-1} \neq 0$ and y also has order p^n. Let U' be the cyclic subspace generated by y: then U' has order p^n and so is isomorphic to U. Suppose z is in the radical of $U + U'$, and let $z = xf(A) + yg(A)$, f, g polynomials. If $f(A)$ is not zero, then f is not divisible by p^n, and so there is a polynomial h such that $f(A)h(A) = p(A)^{n-1}$. Then

$$\begin{aligned} 0 = (z, yh(A)) &= (zh(A), y) = (xp(A)^{n-1} + yg(A)h(A), y) \\ &= 1 \text{ since } (xp(A)^{n-1}, y) = 1 \text{ and } (yg(A)h(A), y) = 0 , \end{aligned}$$

a contradiction. Thus $f(A) = 0$, and since $(yp(A)^{n-1}, x) = -1$, we obtain in the same way $g(A) = 0$, so $z = 0$. Thus $U + U'$ is non-isotropic.

Now assume A is nilpotent, and U, U' are isomorphic cyclic subspaces such that $U + U'$ is non-isotropic. Let $x, xA, \cdots, xA^{r-1}$, and $y, yA, \cdots, yA^{r-1}$ be bases for U, U' respectively, and $A^r = 0$. $(yA^{r-1}, xA^i) = (yA^{r+i-1}, x) = 0$ if $i > 1$, and also $(yA^{r-1}, z) = 0$ for every z in U', thus we must have $(yA^{r-1}, x) \neq 0$ since otherwise yA^{r-1} is in the radical of $U + U'$. Replacing x by a scalar multiple if necessary, we may assume $(x, yA^{r-1}) = 1$.

Fix k, $0 \leqq k < r - 1$, and assume that $(x, yA^j) = 0$ for $k < j < r - 1$. Let $x' = x + \alpha xA^{r-1-k}$, α a scalar to be specified shortly. Then x' is also a cyclic generator of U, and

$$(x', yA^{r-1}) = (x, yA^{r-1}) + \alpha(xA^{r-1-k}, yA^{r-1}) = (x, yA^{r-1}) = 1$$

since $r-1-k \geqq 1$. Also, for $k < j < r-1$,

$$(x', yA^j) = (x, yA^j) + \alpha(xA^{r-1-k}, yA^j) = (x, yA^j) + \alpha(x, yA^{r-1+j-k}) \\ = (x, yA^j)$$

since $j-k \geqq 1$, so $(x', yA^j) = 0$ for $k < j < r-1$. Finally,

$$(x', yA^k) = (x, yA^k) + \alpha(x, yA^{r-1}) = (x, yA^k) + \alpha\ .$$

Let $\alpha = -(x, yA^k)$, then $(x', yA^k) = 0$ also. Proceeding in this way we can obtain a cyclic generator x'' of U such that

$$(x'', yA^{n-1}) = 1 = -(y, x''A^{n-1}),\ \text{and}\ (x'', yA^j) = 0\ \text{for}\ j \neq n-1\ .$$

This completes the proof.

The above lemma has several immediate consequences:

1. The invariant factors $\delta_1, \cdots, \delta_r$ of a self-adjoint linear transformation in a symplectic space come in pairs: $\delta_1 = \delta_2, \cdots, \delta_{r-1} = \delta_r$ (and r is even).

2. If V has dimension m there is a self-adjoint nilpotent linear transformation of index of nilpotency m.

3. If the base field is algebraically closed, two self-adjoint linear transformations are conjugate by an isometry of V if and only if they have the same invariant factors.

The last statement, 3., is also valid for a space V with symmetric scalar product over an algebraically closed field. This follows easily from Lemma 2.1.

In order to extend the theorem of Frobenius mentioned above to arbitrary central simple Jordan algebras, we have to define now invariant factor degrees, and preferably also invariant factor polynomials, for an element x of such a Jordan algebra. For characteristic zero, there is a general method for doing this, due to Professor Jacobson (unpublished), based on Lie algebra methods: it is proved that if x is a nilpotent element, then there exists another nilpotent element y such that x, y and the identity element of $\mathfrak{J}$ generate a semisimple subalgebra which is a direct sum of simple algebras H_i, where H_i is the Jordan algebra of all self-adjoint linear transformations in a vector space of dimension n_i with non-degenerate symmetric scalar product of maximal index, and if $n_1 \geqq n_2 \geqq \cdots \geqq n_r$, then n_1 is the index of nilpotency of x. It is then natural to define $\delta_i(\lambda) = \lambda^{n_i}$, $d_i = n_i$. We shall use the full statement of the above theorem only for the exceptional Jordan algebra over an algebraically closed field. In this case of the exceptional algebra the theorem has also been proved for characteristic $p \neq 2$ or 3, based on (unpublished) work of Professor Jacobson on the representation theory of a simple 3-dimensional Lie algebra, so that we have to assume characteristic not 2 or 3 for the exceptional Jordan algebra in Theorems 2.2, 2.4.

To avoid the difficulties of Lie algebra arguments in characteristic p, we will define the d_i differently, and will use the above theorem of Professor Jacobson only in some of the proofs for the exceptional algebra.

Definition of the d_i: Let $\mathfrak{J}$ be central simple, and extend the base field to its algebraic closure. If $\mathfrak{J}$ becomes the algebra of all $n \times n$ symmetric matrices (Type B) or all $n \times n$ matrices, Type A, take the usual definition of d_i, δ_i for the $n \times n$ matrix x. For Type C, the "symplectic-symmetric" matrix x has invariant factors equal in pairs, and we take for δ_i one member of each pair, and for d_i the degree of δ_i. For type D, the element x satisfies a minimum polynomial μ of degree one or two; let $\delta_2 = \mu = \delta_1$, $d_1 = d_2 = 1$ if μ is of degree one, and $\delta_1 = \mu$, $\delta_2 = 0$, $d_1 = 2$, $d_2 = 0$ if μ is of degree two. Finally, if $\mathfrak{J}$ is an exceptional algebra, every element satisfies a minimum polynomial of degree at most 3 ([5]): set $d_1 = 3$, $d_2 = d_1 = 0$, or $d_1 = 2$, $d_2 = 1$, $d_3 = 0$, or $d_1 = d_2 = d_3 = 1$ according as the degree is 3, 2, or 1; set $\delta_1 =$ minimum polynomial of x. Finally, if $\mathfrak{J}$ is of degree one, then $\mathfrak{J} = F \cdot 1$, and we can set $d_1 = 1$, $d_i = 0$ for $i > 1$. In the following proofs the case of $\mathfrak{J}$ of degree one will not be mentioned because of its triviality.

The d_i, δ_i as defined above have the following properties:

(a) $\delta_1 =$ minimum polynomial of x in $\mathfrak{J}$. $d_i \geqq d_{i+1}$ and δ_{i+1} divides δ_i whenever defined.

(b) $d_1 \leqq$ degree of $\mathfrak{J}$ (defined as the maximum number of orthogonal idempotents in a decomposition of the identity when the base field is extended to its algebraic closure). $\sum_i d_i =$ degree of $\mathfrak{J}$.

(c) If $\mathfrak{J}$ is special (Types A–D) and the base field is algebraically closed, two elements are taken into one another by an automorphism of $\mathfrak{J}$ leaving the center fixed if and only if they have the same invariant factors. The same thing is true for $\mathfrak{J}$ an exceptional algebra, at least for characteristic zero, as will become apparent from some of the later proofs. However, we will not use this, and so do not give the details of a proof.

The next theorem describes the structure of $\mathfrak{C}_{\mathfrak{J}}((x))$ if $\mathfrak{J}$ is central simple over an algebraically closed field and special. The exceptional algebra is studied during the proof of the succeeding theorem, but we cannot give a simple statement for it.

THEOREM 2.1. (a) *Let $\mathfrak{J}$ be the Jordan algebra of all $n \times n$ symmetric matrices over an algebraically closed field F, N a nilpotent element of $\mathfrak{J}$ with invariant factors $\delta_1, \delta_2, \cdots, \delta_r$ where δ_{i+1} divides δ_i. Let $(F[\lambda])_r$ be the algebra of all $r \times r$ matrices with coefficients polynomials in the indeterminate λ, $\mathfrak{K}$ the subalgebra of matrices $(\beta) = (\beta_{ij}(\lambda))$ satisfying $(\delta)^{-1}(\beta)'(\delta) = (\beta)$ where $(\delta) = \mathrm{diag}\,(\delta_1, \cdots, \delta_r)$ and $(\beta)'$ is the transpose of β, and $\mathfrak{N}$ the ideal in $\mathfrak{K}$ of matrices (β) of the form $(\beta) = (\alpha)(\delta)$ for some*

(α) in $(F[\lambda])_r$. Then $\mathfrak{C}_{\mathfrak{J}}((N))$ is isomorphic to $\mathfrak{R}/\mathfrak{N}$. If A is an arbitrary element of $\mathfrak{J}$, then $\mathfrak{C}_{\mathfrak{J}}((A))$ is a direct sum of algebras $\mathfrak{R}_i/\mathfrak{N}_i$, one for each characteristic root. If $\mathfrak{J}$ consists of all hermitian matrices over an arbitrary field and N is a nilpotent element, then $\mathfrak{C}_{\mathfrak{J}}(N)$ is as above with δ_i replaced by $\alpha_i\delta_i$, α_i a self-adjoint scalar, and $(\beta)'$ the conjugate transpose of (β).

(b) *Let $\mathfrak{J}$ be the Jordan algebra of all $n \times n$ symplectic-symmetric matrices over an arbitrary field, N a nilpotent element with invariant factors $\delta_1, \cdots, \delta_r$ in $\mathfrak{J}$* (i.e. *the δ_i are every other one of the ordinary invariant factors). Then $\mathfrak{C}_{\mathfrak{J}}((N))$ is isomorphic to $\mathfrak{R}/\mathfrak{N}$ where $\mathfrak{R}$ is the subalgebra of $(F[\lambda])_{2r}$ of matrices (β) satisfying $(\delta)^{-1}(\beta)'(\delta) = (\beta)$, $(\delta) = \delta_1(e_{12} - e_{21}) + \delta_2(e_{34} - e_{43}) + \cdots + \delta_r(e_{2r-1,2r} - e_{2r,2r-1})$ (e_{ij} matrix units in $(F[\lambda])_{2r}$), and $\mathfrak{N}$ is the ideal in $\mathfrak{R}$ of matrices of the form $(\beta) = (\alpha)(\delta)$.*

(c) *Let $\mathfrak{J}$ be the algebra $(F_n)_j$ of all $n \times n$ matrices over F, A an element of $\mathfrak{J}$ with invariant factors $\delta_1, \cdots, \delta_r$. Let $\mathfrak{R}$ be the subalgebra of $(F[\lambda])_r$ of all (β) such that $(\delta)^{-1}(\beta)'(\delta)$ is in $(F[\lambda])_r$, (δ) as in* (a), *$\mathfrak{N}$ the ideal of (β) of the form $(\beta) = (\alpha)(\delta)$, (α) in $(F[\lambda])_r$. Then $\mathfrak{C}_{\mathfrak{J}}((A))$ is isomorphic to $\mathfrak{R}/\mathfrak{N}$.*

(d) *Let $\mathfrak{J}$ be central simple of Type D. Then $\mathfrak{C}_{\mathfrak{J}}((x)) = (x)$ if x is not in the center of $\mathfrak{J}$, and $\mathfrak{C}_{\mathfrak{J}}((x)) = \mathfrak{J}$ otherwise.*

Proof. (a) We take $\mathfrak{J}$ to be the set of all self-adjoint linear transformations in a hermitian self-dual space V, and N a nilpotent element of $\mathfrak{J}$. Let $x_1, \cdots, x_r$ generate cyclic non-isotropic mutually orthogonal subspaces of orders $d_i =$ degree of δ_i, $i = 1, \cdots, r$ and satisfy $(x_i, x_iN^{j-1}) = \alpha_i \neq 0$ if $j = d_i - 1$, $(x_i, x_iN^{j-1}) = 0$ otherwise.

Let B be any linear transformation commuting with N and write $x_iB = \sum_{j=1}^r x_j\beta_{ij}(N)$, β_{ij} polynomials. Since $x_i\delta_i(N) = 0$, $\delta_i\beta_{ij} \equiv 0 \pmod{\delta_j}$. Also, since $BN = NB$, $(x_iN^pB, x_jN^q) = (x_iN^{p+q}B, x_j)$, so that B is self-adjoint if and only if $(x_iN^kB, x_j) = (x_iN^k, x_jB)$ for all i, j, k.

Now fix i, j, $1 \leqq i \leqq j \leqq r$ and let

$$\beta_{ij}(\lambda) = \mu_0 + \mu_1\lambda + \cdots + \mu_{p-1}\lambda^{p-1}, \quad p = d_j$$
$$\beta_{ji}(\lambda) = \nu_0 + \nu_1\lambda + \cdots + \nu_{q-1}\lambda^{q-1}, \quad q = d_i .$$

Since $(x_j, x_jN^p) = \alpha_j$, $(x_i, x_iN^q) = \alpha_i$, and $(x_i, x_iN^k) = 0$ for $k \neq q$, we have the condition:

$$(x_iN^kB, x_j) = (x_iBN^k, x_j) = \mu_{p-1-k}(x_jN^{p-1}, x_j) = \alpha_j\mu_{p-1-k}$$
$$(x_iN^k, x_jB) = \nu'_{q-1-k}(x_iN^{q-1}, x_i) = \alpha_i\nu'_{q-1-k}$$

for all k, that is $\alpha_i^{-1}\mu_{p-1-k} = \alpha_j^{-1}\nu'_{q-1-k}$ or $\alpha_i^{-1}\beta_{ij}(\lambda) = \lambda^{(d_j-d_i)}\alpha_j^{-1}\beta_{ji}(\lambda)'$

$$(1) \qquad \beta_{ij}(\lambda) = [\alpha_i^{-1}\delta_i(\lambda)]^{-1}\beta_{ji}(\lambda)'(\alpha_j^{-1}\delta_j(\lambda))$$

$\beta_{ji}(\lambda)'$ denoting the conjugate of $\beta_{ji}(\lambda)$.

Since $\delta_i(\lambda)\beta_{ij}(\lambda) \equiv 0 \pmod{\delta_j(\lambda)}$ is a consequence of the condition (1), every matrix (β) satisfying the above condition (1) defines a B in $\mathfrak{C}_{\mathfrak{J}}((N))$ if we define for all i, k $x_iN^kB = \sum_j x_j\beta_{ij}(N)N^k$, and conversely every B in $\mathfrak{C}_{\mathfrak{J}}(N)$ gives rise to such a matrix. Since δ_j is the order of x_j, it is clear that B is zero if and only if $\beta_{ij} \equiv 0 \pmod{\delta_j}$. This proves the *isomorphism for N nilpotent. If A is an arbitrary element of $\mathfrak{J}$, in the* symmetric case, and all the characteristic values λ_i of A lie in F, then we can write $A = \sum_i \lambda_i E_i + N_i$, the E_i, N_i being the idempotent and nilpotent elements of A. Set $V_i = VE_i$, then $\mathfrak{C}_{\mathfrak{J}}(A)$ is the direct sum of $\mathfrak{C}_{\mathfrak{J}_i}((N_i))$, $\mathfrak{J}_i$ the symmetric linear transformations on V_i. If F is algebraically closed, we may assume $\alpha_i = 1$ for V symmetric.

(b) We denote by A the nilpotent symplectic-symmetric matrix under consideration. Regarding the elements of $\mathfrak{J}$ as self-adjoint linear transformations in V, we can choose a canonical basis $\{x_iA^k, y_iA^k\}$ as in Lemma 2.2: $(x_iA^k, y_i) = 1 = -(y_iA^k, x_i)$ if $k = d_i - 1$, all other scalar products are zero. Let B be a linear transformation in V commuting with A, and let

$$x_iB = \textstyle\sum_j x_j\varphi_{ij}(A) + y_j\psi_{ij}(A)$$
$$y_iB = \textstyle\sum_j x_j\eta_{ij}(A) + y_j\rho_{ij}(A)\ .$$

Then

$$(x_iA^kB, y_j) = (x_j\varphi_{ij}(A)A^k, y_j)$$
$$(x_iA^k, y_jB) = (x_iA^k, y_i\rho_{ji}(A)) = (x_i\rho_{ji}(A)A^k, y_i)$$
$$(x_iA^kB, x_j) = (y_j\psi_{ij}(A)A^k, x_j) = -(x_j\psi_{ij}(A)A^k, y_j)$$
$$(x_iA^k, x_jB) = (x_iA^k, y_i\psi_{ji}(A)) = (x_i\psi_{ji}(A)A^k, y_i)$$
$$(y_iA^kB, y_j) = (x_j\eta_{ij}(A)A^k, y_j)$$
$$(y_iA^k, y_jB) = (y_iA^k, x_i\eta_{ji}(A)) = -(x_i\eta_{ji}(A)A^k, y_i)\ .$$

Thus B is self-adjoint if and only if

$$(x_j\varphi_{ij}(A)A^k, y_j) = (x_i\rho_{ji}(A)A^k, y_i)$$
$$(x_j\psi_{ij}(A)A^k, y_j) = -(x_i\psi_{ji}(A)A^k, y_i)\ .$$
$$(x_j\eta_{ij}(A)A^k, y_j) = -(x_i\eta_{ji}(A)A^k, y_i)\ .$$

Since $(x_iA^k, y_i) = 1$ if $k = d_i - 1$, $= 0$ otherwise, we have, exactly as in the symmetric case,

$$(2) \qquad \varphi_{ij} = \delta_i^{-1}\rho_{ji}\delta_j,\ \ \psi_{ij} = -\delta_i^{-1}\psi_{ji}\delta_j,\ \ \eta_{ij} = -\delta_i^{-1}\eta_{ji}\delta_j\ .$$

To B we now assign the $2r \times 2r$ matrix $(\beta_{u,v}(\lambda))$ where

$$\beta_{2i-1,2j-1} = \varphi_{i,j}$$

$$\begin{aligned} \beta_{2i-1,2j} &= \psi_{i,j} \\ \beta_{2i,2j-1} &= \eta_{i,j} \\ \beta_{2i,2j} &= \rho_{i,j} \end{aligned} \qquad i, j = 1, \cdots, r$$

i.e. we arrange the elements φ, ψ, η, ρ in 2×2 blocks

$$\begin{pmatrix} \varphi_{ij} & \psi_{ij} \\ \eta_{ij} & \rho_{ij} \end{pmatrix}.$$

If now (δ) is the matrix with the blocks

$$\begin{pmatrix} 0 & \delta_i \\ -\delta_i & 0 \end{pmatrix}$$

on the main diagonal and zeros elsewhere, condition (2) in matrix form is $(\delta)^{-1}(\beta)'(\delta) = (\beta)$.

If B corresponds to (β), clearly $B = 0$ if and only if $\varphi_{ij}, \psi_{ij}, \eta_{ij}, \rho_{ij}$ are all $\equiv 0 \pmod{\delta_j}$, i.e. $(\beta) = (\alpha)(\delta)$ for some (α) in $(F[\lambda])_{2r}$. It is now clear that the map B into β gives an isomorphism of $\mathfrak{C}_{\mathfrak{J}}((A))$ and $\mathfrak{R}/\mathfrak{N}$.

(c) This case is treated in [6], Chapt. 3, where however the defining relation for $\mathfrak{R}$ is

$$(3) \qquad \delta_i \beta_{ij} \equiv 0 \pmod{\delta_j}.$$

However, the condition $\delta_i\beta_{ij} \equiv 0 \pmod{\delta_j}$ is clearly equivalent to $(\delta)^{-1}(\beta)'(\delta)$ being in $(F[\lambda])_r$, i.e. to the existence of (γ) in $(F[\lambda])_r$ such that $(\beta)'(\delta) = (\delta)(\gamma)$.

(d) Let $\mathfrak{J}$ of be type D, then $\mathfrak{J} = F\cdot 1 \oplus V$, 1 the identity element and V a vector space of dimension at least two and with non-degenerate symmetric scalar product (x, y). $\mathfrak{J}$ can be considered as a subspace of the Clifford algebra C determined by V, and the product in $\mathfrak{J}$ can be written as $x \circ y = xy + yx$, where xy is the product in C. The vector space C can be identified with the vector space of the exterior algebra E over V, with multiplication $x \wedge y$, in such a way that $xy = x \wedge y + (x, y)1$ for x, y in V (See [3]). Let now $a = \alpha 1 + v$, $\alpha \in F$, $v \in V$ be an element of $\mathfrak{J}$. Clearly $\mathfrak{C}_{\mathfrak{J}}((a)) = \mathfrak{C}_{\mathfrak{J}}((v))$, and $\mathfrak{C}_{\mathfrak{J}}((a)) = \mathfrak{J}$ if $v = 0$. Let now $v \neq 0$: we have to show $\mathfrak{C}_{\mathfrak{J}}((v)) = (v)$. Let $w \in V$ and $w \in \mathfrak{C}_{\mathfrak{J}}((v))$. As shown in § 1, $vw = wv$, but $vw = v \wedge w + (v, w)1$, so $v \wedge w = w \wedge v = 0$, thus w is in (v). This proves $\mathfrak{C}_{\mathfrak{J}}((v)) = (v)$ if v is not zero, and completes the proof of the theorem.

COROLLARY 2.1. *If $\mathfrak{J}$ is a central simple Jordan algebra of type B with enveloping associative algebra $\mathfrak{A}$ and x is any element of $\mathfrak{J}$, then the enveloping associative algebra of $\mathfrak{C}_{\mathfrak{J}}((x))$ is $\mathfrak{C}_{\mathfrak{A}}((x))$. The same statement does not hold if $\mathfrak{J}$ is of type C.*

Proof. As usual, we assume the base field F is algebraically closed, and consider first the case of a nilpotent element x. Then, in the notation of previous theorem, $\mathfrak{C}_{\mathfrak{A}}((x))$ can be identified with the associative algebra $\mathfrak{K}/\mathfrak{N}$, $\mathfrak{K}$ the subalgebra of $(F[\lambda])_r$ of matrices (α) such that $(\delta)^{-1}(\alpha)'(\delta)$ is in $(F[\lambda])_r$ and $\mathfrak{N}$ the ideal of matrices of the form $(\alpha)(\delta)$, while $\mathfrak{C}_{\mathfrak{J}}((x))$ is the Jordan subalgebra $(H+\mathfrak{N})/\mathfrak{N}$ of $\mathfrak{K}/\mathfrak{N}$, H denoting the matrices (α) such that $(\delta)^{-1}(\alpha)'(\delta)=(\alpha)$. It is sufficient to prove that the enveloping algebra of H in $(F[\lambda])_r$ contains every element of $\mathfrak{K}$. If $r>1$, H contains elements φe_{ii} and $\varphi e_{ij}+\delta_i\delta_j^{-1}\varphi e_{ji}$ for $1\leqq i<j\leqq r$ and φ any polynomial. Thus the enveloping algebra contains $e_{ii}(\varphi e_{ij}+\delta_i\delta_j^{-1}\varphi e_{ji})=\varphi e_{ij}$ for $i<j$ and also $\delta_i\delta_j^{-1}\varphi e_{ji}=(\varphi e_{ij}+\delta_i\delta_j^{-1}\varphi e_{ji})e_{ii}$ as well as φe_{ii} for all i; however the elements φe_{ij}, $i\leqq j$, and ψe_{ij} with $\psi\equiv 0 \pmod{\delta_i\delta_j^{-1}}$ for $i>j$, clearly are a basis for $\mathfrak{K}$. Note that if $r=1$, then $\mathfrak{C}_{\mathfrak{J}}((x))=\mathfrak{C}_{\mathfrak{A}}((x))$. If now x is not nilpotent, we write $x=\sum\lambda_i e_i+n_i$, e_i idempotent and n_i nilpotent as before, and set $\mathfrak{A}_i=e_i\mathfrak{A}e_i$, $\mathfrak{J}_i=e_i\mathfrak{J}e_i$. Then $\mathfrak{C}_{\mathfrak{A}}((x))=\sum_i\oplus\,\mathfrak{C}_{\mathfrak{A}_i}((n_i))$, $\mathfrak{C}_{\mathfrak{J}}((x))=\sum\oplus\,\mathfrak{C}_{\mathfrak{J}_i}((n_i))$, therefore the enveloping algebra of $\mathfrak{C}_{\mathfrak{J}}((x))$ is $\mathfrak{C}_{\mathfrak{A}}((x))$. If $\mathfrak{J}$ is of type C, and A is a nilpotent $2n\times 2n$ symplectic-symmetric element of $\mathfrak{J}$ of index of nilpotency n, i.e. $r=1$, $\delta_1=\lambda^n$, then it is easy to see that $\mathfrak{C}_{\mathfrak{J}}((x))=(x)$ and is a commutative associative algebra, whereas $\mathfrak{C}_{\mathfrak{A}}((x))$ is not commutative.

THEOREM 2.2. *Let $\mathfrak{J}$ be a central simple Jordan algebra, n its degree and $n+(1/2)n(n-1)s$ its dimension* [*thus* $s=1,2,4,8$ if $n\geqq 3$, $s=1,2,4$ if $n>3$, $s\geqq 1$ if $n=2$]. *Let x be an element of $\mathfrak{J}$ with invariant factor degrees $d_1\geqq d_2\geqq\cdots\geqq d_r$, as defined before. Then $\mathfrak{C}_{\mathfrak{J}}((x))$ has dimension $\sum_{k=0}^{r-1}(sk+1)d_{k+1}$.* If $s=8$, *we assume characteristic* $\neq 2$ *or* 3.

Proof. We may assume the base field is algebraically closed.

(a) Let $\mathfrak{J}$ be of type A: this is the theorem of Frobenius and follows from Th. 2.1, part C). Here $s=2$.

(b) $\mathfrak{J}$ of type B. First let x be nilpotent. We merely have to calculate the dimension of the space of matrices (β) with $(\delta)^{-1}(\beta)'(\delta)=(\beta)$ and subtract the dimension of the (β) of the form $(\alpha)(\delta)$, i.e. reduce $\beta_{ij} \bmod \delta_j$. Clearly (β) is determined by the elements β_{ij} with $i\leqq j$ and β_{ij} of degree $\leqq d_j$, thus the dimension is $d_1+2d_2+\cdots+rd_r$, and here $s=1$. If x is not nilpotent, $x=\sum\alpha_i e_i+n_i$, and $\mathfrak{C}_{\mathfrak{J}}((x))$ is a direct sum of the algebras $\mathfrak{C}_{\mathfrak{K}}((n))$, $\mathfrak{K}=\mathfrak{J}_1(e_i)$ and $n=n_i$. The $\mathfrak{J}_1(e_i)$ are again algebras of all symmetric matrices (of degree $\geqq 1$), and the invariant factor degree d_k of x is the sum of the $d_k(n_i)$. Thus the formula holds also for x.

(c) Type C. Again we need only consider nilpotent elements, and the calculation of the dimension is like that for symmetric matrices, using Th. 2.1 b and its proof.

(d) $\mathfrak{J}$ of type D. Let $x = \alpha 1 + v$, $v \in V$, using the notation of the proof of Cor. 2.1. If the minimum polynomial δ_1 has degree one, then $v = 0$, $d_1 = d_2 = 1$, and $\mathfrak{C}_{\mathfrak{J}}((x)) = \mathfrak{J}$. Here $2 + s =$ dimension of $\mathfrak{J}$, so $\sum_{k=0}^{r-1}(sk+1)d_{k+1} = d_1 + (s+1)d_2 = s + 2$. Let now the minimum polynomial have degree two, then $d_1 = 2$, $d_2 = 0$, and v is not zero, so $\mathfrak{C}_{\mathfrak{J}}((x)) = (x)$ has dimension $2 = \sum_{k=0}^{r-1}(sk+1)d_{k+1}$.

(e) $\mathfrak{J}$ the exceptional algebra, type E. We use the previously mentioned result of Professor Jacobson that if x is a nilpotent element of $\mathfrak{J}$, with $x^n = 0$, $x^{n-1} \neq 0$, then (x) can be imbedded in a direct sum of algebras $H^{(i)}$ of all $n_i \times n_i$ symmetric matrices with $n = n_1 \geqq n_2 \geqq \cdots \geqq n_r$. The identity element of $\mathfrak{J}$ is in (x) and $\mathfrak{J}$ has degree 3, so r is at most 3, and x satisfies a cubic polynomial over the center F. Thus we have only a small number of cases to consider. Let the cubic polynomial be $(\lambda - \alpha_1)(\lambda - \alpha_2)(\lambda - \alpha_3)$, $\alpha_i \in F$.

(1) The α_i are all equal, say to α_1. Since $\mathfrak{C}_{\mathfrak{J}}((x - \alpha_1 1)) = \mathfrak{C}_{\mathfrak{J}}((x))$, we may replace x by $x - \alpha_1 1$, and thus have x nilpotent.

(i) Let $x^3 = 0$, $x^2 \neq 0$. Since no element of a 2×2 or 1×1 matrix algebra can satisfy this condition, $H^{(1)}$ must be of degree 3, and $H^{(i)} = 0$ for $i > 1$. Then $H^{(1)}$ contains primitive orthogonal idempotents e_i with sum 1, and elements u_{12}, u_{13}, u_{23} with $u_{ij} \circ u_{ij} = 4(e_i + e_j)$, $u_{ij} \circ u_{jk} = u_{ik}$ for i, j, k distinct, and $e_i \circ u_{ij} = (1/2)u_{ij}$. By Theorem 9.1 of [9], $\mathfrak{J}$ can be represented as $H(C_3)$ with these elements as matrix units and involution $\sum c_{ij}e_{ij} \to \sum \bar{c}_{ij}e_{ji}$, $(e_{ii} = 2e_i, e_{ij} + e_{ji} = u_{ij})$. The element x is represented by a matrix N with elements in F. Since $N^3 = 0$, $N^2 \neq 0$, and N is a 3×3 symmetric matrix, it follows from the remarks after Lemma 2.1 that there is an orthogonal matrix T with elements in F such that $TNT^{-1} = M$,

$$M = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & \alpha \\ 1 & \alpha & 0 \end{pmatrix} \quad \alpha \in F \quad \text{and} \quad \alpha^2 = -1$$

since also $M^3 = 0$, $M^2 \neq 0$. The matrix T is in the nucleus of C_3 (i.e. associates with all elements), and so $X \to TXT^{-1} = X^A$ is an automorphism A of $H(C_3)$. Using the matrix units e_i^A, u_{ij}^A, by Th. 9.1 of [9] we may represent $\mathfrak{J}$ as $H(C_3)$ with involution the ordinary conjugate transpose operation with x being represented by M. Since M has coefficients in the center of C, by Proposition 1.3b, $\mathfrak{C}_{\mathfrak{J}}((M))$ is just the set of matrices commuting with M: these turn out to be all matrices of the form

$$X = \gamma_1 e_{11} + \gamma_2 e_{22} + \gamma_3 e_{33} + x_3 e_{12} + \bar{x}_3 e_{21} + y_2 e_{13} + \bar{y}_2 e_{31} + z_1 e_{23} + \bar{z}_1 e_{31}$$

such that the γ_i are in F, $\bar{x}_3 = x_3$, $\bar{y}_2 = y_2$, $\bar{z}_1 = z_1 = \alpha y_2$

and $$\gamma_1 + \alpha x_3 = \gamma_3 = \gamma_2 - \alpha x_3 \, .$$

Thus the x_i, y_i, z_i are in F, and $X = (\gamma_1 + \alpha x_3)1 + \alpha z_1 M + y_2 M^2$. Hence $\mathfrak{C}_{\mathfrak{J}}((x)) = (x)$. Also, $d_1 = 3$, $d_2 = d_1 = 0$ and $\sum(8k+1)d_{k+1} = 3 =$ dimension of (x).

(ii) $x^2 = 0$: then $H^{(1)}$ is of degree two, so $H^{(2)}$ must be of degree one and equal to Fe, e a primitive idempotent. We note for future reference that $H^{(1)} \oplus H^{(2)}$ is contained in $\mathfrak{J}_e(1) + \mathfrak{J}_e(0) = \{a \text{ in } \mathfrak{J} \text{ such that } e \circ a = a \text{ or } e \circ a = 0\}$ and so e is in $\mathfrak{C}_{\mathfrak{H}}((x))$. $H^{(1)}$ contains primitive orthogonal idempotents e_1, e_2 and an element u_{12} with $u_{12} \circ u_{12} = 4(e_1 + e_2)$. Write e_3 for e, then there are elements v_{13}, w_{23} in $\mathfrak{J}$ such that $v_{13} \circ v_{13} = 4(e_1 + e_3)$, $w_{23} \circ w_{23} = 4(e_2 + e_3)$, $u_{12} \circ v_{13} = w_{23}$, $u_{12} \circ w_{23} = v_{13}$, and $v_{13} \circ w_{23} = u_{12}$ (this can be seen, for instance, by writing $\mathfrak{J}$ as $H(C_3)$ with $2e_i = e_{ii}$, so that $u_{12} = ce_{12} + \bar{c}e_{21}$ where $c\bar{c} = 1$ in C, and setting $v_{13} = ce_{13} + \bar{c}e_{31}$, $w_{23} = e_{23} + e_{32}$). As before we can write $\mathfrak{J}$ as $H(C_3)$ using the $e_i, u_{12}, v_{13}, w_{23}$ as matrix units: then x is a linear combination of e_1, e_2, u_{12}, and e_3 and so is represented by a matrix with elements in F. Since $x^2 = 0$, the coefficient of e_3 must be zero, and as before the 2×2 symmetric matrix $\gamma_1 e_1 + \gamma_2 e_2 + \gamma_3(e_{12} + e_{21})$, γ_i in F, may be transformed to the form $\alpha(e_1 - e_2) + e_{12} + e_{21}$ by a 2×2 orthogonal matrix T with coefficients in F. Then the 3×3 orthogonal matrix $T_1 = \begin{pmatrix} \multicolumn{2}{c}{T} & 0 \\ & & 0 \\ 0 & 0 & 1 \end{pmatrix}$ induces an automorphism of $H(C_3)$, taking x into

$$N = \begin{pmatrix} \alpha & 1 & 0 \\ 1 & -\alpha & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \alpha^2 = -1 \, .$$

Again we have only to find the matrices commuting with N, which turn out to have the form

$$\begin{pmatrix} \beta_1 & (1/2)\alpha(\beta_2 - \beta_1) & \alpha z_3 \\ (1/2)\alpha(\beta_2 - \beta_1) & \beta_2 & z_3 \\ \alpha\bar{z}_3 & \bar{z}_3 & \beta_3 \end{pmatrix}, \quad \beta_i \text{ in } F,\ z_3 \text{ in } C \, .$$

Thus the dimension of $\mathfrak{C}_{\mathfrak{J}}((x))$ is $3 \cdot \dim F + \dim C = 11$; also $d_1 = 2$, $d_2 = 1$, $d_3 = 0$ so $\sum(8k+1)d_{k+1} = 2 + 9 = 11$.

(2) The cubic polynomial of x has two distinct roots. Then (x) contains two orthogonal idempotents e, f with sum 1, which we may write as $e = e_3$, $f = e_1 + e_2$, where the e_i are orthogonal primitive

idempotents in $\mathfrak{J}$. Then $x = \gamma_1 f + n + \gamma_3 e$, where n is nilpotent, in fact $n^2 = 0$, $e \circ n = 0$, $f \circ n = n$, and e, f, n are all in (x). Note that $(x) \subseteqq J_e(0) + J_e(1)$ and e is in $\mathfrak{C}_{\mathfrak{J}}((x))$. Suppose first that $n = 0$. Then $\mathfrak{C}_{\mathfrak{J}}((x)) = \mathfrak{C}_{\mathfrak{J}}(e) = \mathfrak{C}_{\mathfrak{J}}(f) = \mathfrak{J}_e(0) + J_e(1)$ and has dimension 11, while $d_1 = 2$, $d_2 = 1$ and $\sum(8k + 1)d_{k+1} = 11$. Now suppose $n \neq 0$, and let y be in $\mathfrak{C}_{\mathfrak{J}}((x))$. Since y o-commutes with e, $y = y_0 + y_1$, y_i in $\mathfrak{J}_e(i)$. Since $\mathfrak{J}_e(0)$ is an algebra, y_0 o-commutes with n in $\mathfrak{J}_e(0)$—a simple algebra of degree two; thus by part (d) of this theorem, y_0 is a linear combination of f and n since n is not in the center of $\mathfrak{J}_e(0)$. Also, y_1 is a scalar multiple of e since $\mathfrak{J}_e(1) = Fe$; thus $y = y_0 + y_1$ belongs to (x) and $\mathfrak{C}_{\mathfrak{J}}((x)) = (x)$ and has dimension 3. $d_1 = 3$, $d_2 = 0$ and $\sum(8k + 1)d_{k+1} = 3$.

(3) The cubic polynomial has 3 distinct roots. Then $(x) = Fe_1 + Fe_2 + Fe_3$, e_i mutually orthogonal, and $\mathfrak{C}_{\mathfrak{J}}((x))$, being the intersection of the $\mathfrak{J}_e(0) + \mathfrak{J}_e(1)$ for $e = e_1, e_2, e_3$ respectively, equals (x). Here $d_3 = 3$, $d_2 = d_1 = 0$, and $\sum(8k + 1)d_{k+1} = 3 = \dim \mathfrak{C}_{\mathfrak{J}}((x))$. Incidentally, we also note that for each x in $\mathfrak{J}$ there is a representation of $\mathfrak{J}$ as $H(C_3)$ such that the matrices of $\mathfrak{C}_{\mathfrak{J}}((x))$ are just those commuting with the matrices of (x).

COROLLARY 2.2. *Let $\mathfrak{J}$ be a central simple Jordan algebra of degree n, $(n \geqq 1)$, x an element of $\mathfrak{J}$. Then the minimum polynomial of x in $\mathfrak{J}$ has degree at most n, and equals n if and only if $\mathfrak{C}_{\mathfrak{J}}((x)) = (x)$. Elements x with minimum polynomial of degree n always exist.*

Note. such elements with minimum polynomial of degree equal to the degree of the algebra are a natural generalization of non-derogatory $n \times n$ matrices in the Jordan algebra $(F_n)_j$.

Proof of the Corollary. Since the dimension of (x) is d_1, the degree of the minimum polynomial of x, and (x) is always contained in $\mathfrak{C}_{\mathfrak{J}}((x))$, we see that $\sum_{k=0}^{r-1}(sk + 1)d_{k+1} = d_1$ if and only if $d_2 = \cdots = d_r = 0$. But since $d_1 + \cdots + d_r = n$, this means $d_1 = n$. If the base field is algebraically closed, hence infinite, and $e_1, \cdots, e_n$ are primitive orthogonal idempotents with sum 1, then $x = \alpha_1 e_1 + \cdots + \alpha_n e_n$ is of degree n if the α_i are distinct. Since the degree of the minimum polynomial is unaffected by field extension, there must exist elements with $d_1 = n$ in a central simple $\mathfrak{J}$ over any base field.

We next consider the double centralizer $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{C}_{\mathfrak{J}}(x))$, and prove that it is always equal to (x): a known theorem for $\mathfrak{J}$ the algebra of all $n \times n$ matrices over a field. We can also prove this for $\mathfrak{J}$ a Jordan algebra analogous to a finite dimensional Jordan algebra of type A, B, or C but obtained from a simple ring with minimum condition instead of a simple algebra.

THEOREM 2.3. *Let* $\mathfrak{A}$ *be an involutorial simple ring with minimum condition,* F *the set of self-adjoint elements of the center of* $\mathfrak{A}$, *and* $\mathfrak{J}$ *the set of self-adjoint elements of* $\mathfrak{A}$, *regarded as a Jordan algebra over* F. *Let* x *be an algebraic element of* $\mathfrak{J}$, i.e. (x) *is finite dimensional over* F. *Then* $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{C}_{\mathfrak{J}}((x))) = (x)$.

Proof. We can also consider $\mathfrak{A}$ as an algebra over F, and if K is a finite extension field of F containing the eigenvalues of x, form $\mathfrak{A} \otimes_F K$, $\mathfrak{J} \otimes_F K$, noting that the involution in $\mathfrak{A}$ can be extended to $\mathfrak{A} \otimes K$ by letting it be the identity on $K = 1 \otimes K$ and that $\mathfrak{J} \otimes K$ is then the set of self-adjoint elements of $\mathfrak{A} \otimes K$. We then need only prove that if x is an algebraic element of $\mathfrak{J} \otimes K$ with eigenvalues in K, then (x) is its own double centralizer.

Case 1. Let E, the center of $\mathfrak{A}$, equal F. Then $\mathfrak{A} \otimes K$ is again an involutorial simple ring with minimum condition so that we may as well assume F already contained the eigenvalues of x. We may represent $\mathfrak{A}$ as the algebra (over F) of all linear transformations on a finite dimensional vector space V, over an involutorial division ring D with center F, with a non-degenerate scalar product which either is hermitian or else is symplectic and $D = F$.

(a) Hermitian scalar product. We use the following lemma:

LEMMA 2.3. *Let* V *be a vector space of dimension* $n \geqq 1$ *with non-degenerate hermitian scalar product over a division ring* D. *Let* F *be the subfield of self-adjoint elements of* D, $\mathfrak{J}$ *the Jordan algebra, over* F, *of self-adjoint linear transformations in* D. *Let* A *be an element of* $\mathfrak{J}$ *such that* $A^n = 0$, $A^{n-1} \neq 0$ (set $A^0 = 1$). *Then* $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{C}_{\mathfrak{J}}((A))) = (A)$.

Proof. First consider the case $n = 1$, that is $V = D$ (as left D space). Then $(A) = F$, $\mathfrak{C}_{\mathfrak{J}}((A)) = \mathfrak{J}$. Now it is known that either $\mathfrak{J}$ generates D or else $\mathfrak{J} = F$ ([7], p. 187). In either case, $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{J}) = \mathfrak{C}_{\mathfrak{J}}(\mathfrak{C}_{\mathfrak{J}}(A))) = (A) =$ center of $\mathfrak{J} = F$. Let now $n > 1$. By Lemma 2.1 we can find a basis for V of the form $(v, vA, \cdots, vA^{n-1})$ such that $(vA^i, vA^j) = \lambda$, a self-adjoint non-zero element of D, if $i + j = n - 1$, and $(vA^i, vA^j) = 0$ otherwise. Let B be a self-adjoint linear transformation commuting with $\mathfrak{A}$, and let $vB = \beta_0 v + \beta_1 vA + \cdots + \beta_{n-1} vA^{n-1}$, β_i in D. Then

$$(vB,\ vA^{n-1-i}) = \beta_i (vA^i,\ vA^{n-1-i}) = \beta_i \lambda$$
$$(v,\ vA^{n-1-i}B) = (v,\ \beta_i vA^{n-1}) = \lambda \beta_i'$$

for all i, and conversely, any linear transformation B commuting with A such that the β_i satisfy $\beta_i \lambda = \lambda \beta_i'$ is self-adjoint, $(i = 0, 1, \cdots, n-1)$. For γ in D, let $\gamma^* = \lambda \gamma' \lambda^{-1}$. Then $\gamma \to \gamma^*$ is an involution in D, and β_i^*

$= \beta_i$. The self-adjoint elements of this new involution either generate D or else lie in the center of D. First suppose they lie in the center of D: then $\beta_i^* = \beta_i' = \beta_i$ so the β_i are in F, and B is in (A), ($B = \beta_0 1 + \beta_1 A + \cdots + \beta_{n-1}A^{n-1}$) i.e. $\mathfrak{C}_{\mathfrak{J}}((A)) = (A)$ so $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{C}_{\mathfrak{J}}((A))) = \mathfrak{C}_{\mathfrak{J}}((A)) = (A)$. Next suppose that the self-adjoint elements of * generate D, and let M be in $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{C}_{\mathfrak{J}}((A)))$. Since A is in $\mathfrak{C}_{\mathfrak{J}}((A))$, M commutes with A and so $vM = \mu_0 v + \cdots + \mu_{n-1}vA^{n-1}$. Also M commutes with the linear transformation $B_0(\beta): \rho v A^i \to \rho\beta v A^i$ ($i = 1, n-1$, in D) for every β such that $\beta = \beta^*$. Thus $vB_0M = \beta\mu_0 v + \cdots + \beta\mu_{n-1}vA^{n-1} = vMB_0 = \mu_0\beta v + \cdots + \mu_{n-1}\beta v A^{n-1}$ so that the μ_i commute with all such β, and since the latter generate D, every μ_i is in the center of D, so $\mu_i^* = \mu_i' = \mu_i$ and the μ_i are in F. Thus B is in (A). This completes the proof of the lemma.

Let now A be an algebraic element of $\mathfrak{J}$ with eigenvalues in F: thus $A = \sum \lambda_i E_i + N_i$ with λ_i distinct, in F, and N_i nilpotent, E_i, N_i being in (A). First let the degree of the minimum polynomial of A on V equal the dimension of V. Then V is a direct sum of the non-isotropic mutually orthogonal subspaces $V_i = VE_i$ and the index of nilpotency of $N_i = E_i N_i$ equals the dimension of V_i. Since the E_i are in (A) and so in $\mathfrak{C}_{\mathfrak{J}}((A))$, every element of $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{C}_{\mathfrak{J}}((A)))$ maps each V_i into itself, commutes with N_i on V_i, and is self-adjoint on V_i. By the lemma the induced transformation on V_i is a polynomial $\varphi_i(A)$ with coefficients in F. Since the minimum polynomials of A in V_i are relatively prime in pairs, there is a polynomial $\varphi(\lambda)$ such that $\varphi(A)E_i = \varphi_i(A)E_i$, that is $\varphi(A)$ induces $\varphi_i(A)$ on V_i. Thus every element of $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{C}_{\mathfrak{J}}((A)))$ is a polynomial in A. Finally, consider the general case. Then V is a direct sum of cyclic mutually orthogonal subspaces W_i such that W_{i+1} is a homomorphic image of W_i as A-space and the minimum polynomial of A on W_i has degree equal to the dimension of W_i. Since the orthogonal projections of V on W_i are self-adjoint and in $\mathfrak{C}_{\mathfrak{J}}((A))$, every element C of $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{C}_{\mathfrak{J}}((A)))$ maps each W_i on itself. Let S_i be a self-adjoint linear transformation of W_i into itself which commutes with A on W_i. S_i can be extended to V by letting it act as zero on the W_j with $j \neq i$, and will then belong to $\mathfrak{C}_{\mathfrak{J}}((A))$. Thus C commutes with S_i, and so on each W_i, C is equal to a polynomial $\psi_i(A)$. Finally, we must show that all the $\psi_i(A)$ may be assumed equal. Let x_1 be a cyclic generator of W_1, and x_2 its image under an A-homomorphism of W_1 onto W_2. Then x_2 is a cyclic generator of W_2. Denote by T the mapping of W_1 on W_2 with $x_1T = x_2$, and define $T = 0$ on W_i for $i \geqq 2$. Let T' be the adjoint of T. Since $(W_jT, W_1) = 0$, $(W_j, W_1T') = 0$ for all j, so $W_1T' = 0$. Let $S = T + T'$, then $S = T$ on W_1, S is self-adjoint and since $TA = AT$, also $T'A = AT'$ and $SA = AS$. Thus S is in $\mathfrak{C}_{\mathfrak{J}}((A))$, $SC = CS$, and $x_1S = x_2$. For all i, $x_iC = x_i\psi_i(A)$. Since $x_2 = x_1S$, $x_2\psi_2(A) = x_2C = x_1SC = x_1CS = x_1\psi_1(A)S = x_1S\psi_1(A) = x_2\psi_1(A)$. Thus $\psi_1(A) = \psi_2(A)$ on W_2, and in the same way

$\psi_1(A) = \psi_i(A)$ on W_i for all i. Thus $C = \psi_1(A)$ on V.

(b) Skew-symmetric scalar product: We now write A for the element x of $\mathfrak{J}$. By Lemma 2.2, we can write V as a direct sum of cyclic totally isotropic spaces W_i, W_i' such that $W_i + W_i'$ is non-isotropic and orthogonal to $W_j + W_j'$ for $i \neq j$, W_i is isomorphic to W_i', and W_{i+1} is a homomorphic image of W_i as (A) modules. Let $V_i = W_i + W_i'$ and A_i the restriction of A to V_i. Let $\mathfrak{J}_i$ be the algebra of self-adjoint linear transformations on V_i; by Corollary 2.2, $\mathfrak{C}_{\mathfrak{J}_i}((A_i)) = (A_i)$, since the degree of $\mathfrak{J}_i$ equals the dimension of W_i. Let E_i be the self-adjoint projection on V_i, and let C be in $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{C}_{\mathfrak{J}}((A)))$; then C commutes with the E_i and maps V_i on itself, inducing C_i on V_i. Since C_i is in $\mathfrak{C}_{\mathfrak{J}_i}((A_i))$, C_i is a polynomial $\varphi_i(A)$. To show the φ_i are all equal, we choose cyclic generators x_i, x_i' of W_i, W_i' and define a linear transformation T by: $TA = AT$, $x_iT = x_{i+1}$, $x_i'T = x_{i+1}'$, and T is zero on V_j for $j \neq i$. As in the symmetric case, the adjoint T' of T is zero on V_i, so if we set $S = T + T'$, then $x_iS = x_{i+1}$, $x_i'S = x_{i+1}'$ and S is in $\mathfrak{C}_{\mathfrak{J}}((A))$. Then $x_iSC = x_{i+1}C = x_{i+1}\varphi_{i+1}(A) = x_iCS = x_{i+1}\varphi_i(A)$, and similarly, $x_{i+1}'\varphi_{i+1}(A) = x_{i+1}'\varphi_i(A)$. Thus $C = \varphi(A) = \varphi_1(A)$ on all of V. This completes the proof of Case 1.

Case 2. E, the center of $\mathfrak{A}$, is a quadratic extension of F. Just as in the case of finite dimensional algebras, the ring $\mathfrak{A} \otimes_F E$ is a direct sum of two copies $\mathfrak{A}_1$, $\mathfrak{A}_2$ of $\mathfrak{A}$ and if we extend the involution of $\mathfrak{A}$ to $\mathfrak{A} \otimes E$ by letting it be the identity on $1 \otimes E$ (i.e. if $a \to a'$ is the involution in $\mathfrak{A}$, we set $(a \otimes e)' = a' \otimes e$ for e in E) then the involution maps $\mathfrak{A}_1$ on $\mathfrak{A}_2$, $\mathfrak{A}_2$ on $\mathfrak{A}_1$. Thus the Jordan algebra $\mathfrak{J} \otimes_F E$ of self-adjoint elements of $\mathfrak{A} \otimes_F E$ is isomorphic to the set of elements $a_1 \oplus a_1'$ of $\mathfrak{A}_1 \oplus \mathfrak{A}_2$, where a_1 is in $\mathfrak{A}_1$ and a_1' the image of a_1 under the involution; as $\mathfrak{A}_1$ is isomorphic to $\mathfrak{A}$, $\mathfrak{J} \otimes_F E$ is also isomorphic to $\mathfrak{A}_j$ (i.e. $\mathfrak{A}$ with the Jordan product). In $\mathfrak{A}_j$, the set of elements o-commuting with an element x is just the set of elements commuting with it in the ordinary multiplication, and the double centralizer of (x) for an algebraic element x (over the base field E) is (x): the usual proof for matrices over a field goes over for a division ring, since we can quickly reduce it to the case of a cyclic nilpotent matrix x and note that for such an x, every matrix commuting with it is a polynomial in x and 1 with coefficients in the division ring, and every matrix commuting with these is a polynomial in x with coefficients in the center E of the division ring. This completes the proof.

THEOREM 2.4. *Let $\mathfrak{J}$ be a central simple Jordan algebra and x an element of $\mathfrak{J}$. Then $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{C}_{\mathfrak{J}}((x))) = (x)$. If $\mathfrak{J}$ is exceptional we assume the characteristic is not 2 or 3.*

Proof. Theorem 2.3 covers algebras of types A, B, C. Let $\mathfrak{J}$ be of type D: then, by Th. 2.1 d, if x is not in the center of $\mathfrak{J}$ then $\mathfrak{C}_{\mathfrak{J}}((x)) = (x)$ so $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{C}_{\mathfrak{J}}((x))) = \mathfrak{C}_{\mathfrak{J}}((x)) = (x)$, whereas if x is in the center of $\mathfrak{J}$, then $\mathfrak{C}_{\mathfrak{J}}((x)) = \mathfrak{J}$, and $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{C}_{\mathfrak{J}}((x))) =$ Center of $\mathfrak{J} = (x)$.

It remains to consider the exceptional Jordan algebra $\mathfrak{J}$ over an algebraically closed field F. If the minimum polynomial of x is of degree three, then by Corollary 2.2, $\mathfrak{C}_{\mathfrak{J}}((x)) = (x)$ so $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{C}_{\mathfrak{J}}((x))) = (x)$ also. If the minimum polynomial is of degree one, then $(x) =$ Center of $\mathfrak{J}$, so $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{C}_{\mathfrak{J}}((x))) = (x)$. Finally, let the minimum polynomial be of degree two: then there are two cases—minimum polynomial has one root or distinct roots.

Let the minimum polynomial of x be $(x - \gamma 1)^2$. Replacing x by $x - \gamma 1$ does not change (x), and so we may assume $x^2 = 0$, $x \neq 0$. In the proof of Theorem 2.2, we showed that there is a representation of $\mathfrak{J}$ as $H(C_3)$ such that x is represented by the matrix $\alpha(e_{11} - e_{22}) + e_{12} + e_{21}$, where α is in F and $\alpha^2 = -1$, and so the elements of $H(C_3)$ o-commuting with x are exactly the matrices commuting with the matrix of x. We also showed that $e_3 = (1/2)e_{33}$ is in $\mathfrak{C}_{\mathfrak{J}}((x))$. Let y be in $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{C}_{\mathfrak{J}}((x)))$: then y o-commutes with e_3 and y is also in $\mathfrak{C}_{\mathfrak{J}}((x))$ so that its matrix is as in the proof of Theorem 2.2, and the element z_3 is zero since y o-commutes with e_3. Thus y also belongs to the simple subalgebra $\mathfrak{K}$ of $H(C_3)$ of elements with coefficients in F. Since y o-commutes on $\mathfrak{J}$ with every element of $\mathfrak{C}_{\mathfrak{J}}((x))$ and since $\mathfrak{C}_{\mathfrak{K}}((x)) \subseteqq \mathfrak{C}_{\mathfrak{J}}((x))$ (each consists of the matrices, in $\mathfrak{K}$ or $\mathfrak{J}$ respectively, commuting with x), y o-commutes on $\mathfrak{J}$ with every element of $\mathfrak{C}_{\mathfrak{K}}((x))$. and in particular y o-commutes on $\mathfrak{K}$ with every element of $\mathfrak{C}_{\mathfrak{K}}((x))$. As y is in $\mathfrak{K}$, we have shown that y is in $\mathfrak{C}_{\mathfrak{K}}(\mathfrak{C}_{\mathfrak{K}}((x)))$, that is, that $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{C}_{\mathfrak{J}}((x))) \subseteqq \mathfrak{C}_{\mathfrak{K}}(\mathfrak{C}_{\mathfrak{K}}((x)))$. But $\mathfrak{K} = H(F_3)$ is a simple Jordan algebra of degree three and type B, and we have already proved for such algebras that $\mathfrak{C}_{\mathfrak{K}}(\mathfrak{C}_{\mathfrak{K}}((x))) = (x)$. Thus also $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{C}_{\mathfrak{J}}((x))) = (x)$.

The only case left is that of a minimum polynomial of degree two and distinct roots. Here $x = \alpha e + \beta(1 - e)$, e a primitive idempotent and $\alpha \neq \beta$. Clearly $(x) = (e)$, so we may assume $x = e$. $\mathfrak{C}_{\mathfrak{J}}((e)) = \mathfrak{J}_e(0) + \mathfrak{J}_e(1)$. Let y be in $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{C}_{\mathfrak{J}}((e)))$. Then, since y is in $\mathfrak{C}_{\mathfrak{J}}((e))$, y is in $\mathfrak{J}_e(0) + \mathfrak{J}_e(1)$ and y o-commutes on $\mathfrak{J}$, and therefore on $\mathfrak{J}_e(0) + \mathfrak{J}_e(1)$ also, with every element of $\mathfrak{J}_e(0) + \mathfrak{J}_e(1)$. Thus y is in the center of this algebra which is the sum of the centers of $\mathfrak{J}_e(0)$ and $\mathfrak{J}_e(1)$, central simple algebras with respective identities $1 - e$ and e. Thus $y = \gamma(1 - e) + \delta e$ belongs to (e), and $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{C}_{\mathfrak{J}}((x))) = (x)$. This completes the proof.

Section 3. Centralizers of Simple Algebras. In this section we study the centralizer and double centralizer of a simple subalgebra in a central simple Jordan algebra. We are also able to study the centralizer theory in certain infinite dimensional Jordan algeberas, namely the algebras of all self-adjoint linear transformations on a self-dual vector space which may be infinite dimensional, such as a Hilbert space. The method is much simpler than that of the last section: we use the enveloping associative algebra and the known centralizer theory of simple associative algebras (see [7] and [15]).

We shall prove analogues of the following theorems on associative algebras ([15]): Let M, N be left and right vector spaces, respectively, over a division ring D, dually paired by an inner product (x, y) (i.e. if $(x, z) = 0$ for all z in N, then $x = 0$, and similarly $(u, y) = 0$ for all u in M implies $y = 0$). Denote by $\mathfrak{A} = L(M, N)$ the ring of all linear transformations on M having adjoints on N, regarded as an algebra over its center. ($M = N$, that is, M is self-dual, if and only if $\mathfrak{A}$ has an involution ([6]). The involution can be assumed to be the adjoint map.) Let $\mathfrak{B}$ be a simple subalgebra of $\mathfrak{A}$ containing the identity element. Then:

1. $\mathfrak{C}_{\mathfrak{A}}(\mathfrak{B})$ is also isomorphic to a ring $L(V, W)$ for a pair of dual spaces V and W. If $\mathfrak{A}$ is a simple finite dimensional algebra then so is $\mathfrak{C}_{\mathfrak{A}}(\mathfrak{B})$.

2. $\mathfrak{C}_{\mathfrak{A}}(\mathfrak{C}_{\mathfrak{A}}(\mathfrak{B})) = \mathfrak{B}$.

Actually we will also generalize the above associative theorems from rings $L(M, N)$ to the more general primitive rings with minimal ideals, and obtain a corresponding generalization for Jordan algebras. In what follows the term "simple algebra" will be used only for finite-dimensional algebras.

THEOREM 3.1. *Let $\mathfrak{J}$ be a special central simple Jordan algebra, $\mathfrak{K}$ a semi-simple subalgebra containing the identity of $\mathfrak{J}$. Then $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{K})$ is semi-simple. The same result holds for $\mathfrak{J}$ an exceptional algebra provided $\mathfrak{K}$ is separable.*

Proof. If $\mathfrak{J}$ is of degree one then $\mathfrak{K} = \mathfrak{J}$ and $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{K}) = \mathfrak{J} = \mathfrak{K}$. Next let $\mathfrak{J}$ be of degree two. Then $\mathfrak{J} = F \cdot 1 + V$, V *a* vector space with symmetric scalar product. Since $F \cdot 1$ is in $\mathfrak{K}$, $\mathfrak{K} = F \cdot 1 + \mathfrak{K} \cap V$. If $\mathfrak{K} \cap V$ is one-dimensional, then $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{K}) = \mathfrak{K}$; if $\mathfrak{K} \cap V$ has dimension greater than one, then $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{K}) = F \cdot 1$, and if $\mathfrak{K} \cap V = (0)$ then $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{K}) = \mathfrak{J}$: these statements follow immediately from Theorem 2.1d. Thus $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{K})$ is semi-simple.

Finally, let $\mathfrak{J}$ have degree three or more, and be special. If $\mathfrak{J}$ is of type A_2, i.e. isomorphic to $\mathfrak{A}_j$ for a central simple associative algebra $\mathfrak{A}$, and $\mathfrak{K}$ has enveloping associative algebra $\mathfrak{B}$ in $\mathfrak{A}$, then $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{K}) = \mathfrak{C}_{\mathfrak{A}}(\mathfrak{B})_j$. But $\mathfrak{B}$ is a semi-simple associative algebra, therefore $\mathfrak{C}_{\mathfrak{A}}(\mathfrak{B})$ is also semi-simple, and so $\mathfrak{C}_{\mathfrak{A}}(\mathfrak{B})_j$ is a semi-simple Jordan algebra. The only remaining possibility is that $\mathfrak{J} = H(\mathfrak{A})$, $\mathfrak{A}$ a simple involutorial algebra. Let $\mathfrak{B}$ be the enveloping algebra of $\mathfrak{R}$. Then $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{R}) = \mathfrak{C}_{\mathfrak{A}}(\mathfrak{B}) \cap \mathfrak{J} =$ set of self-adjoint elements of $\mathfrak{C}_{\mathfrak{A}}(\mathfrak{B})$. But $\mathfrak{B}$ and $\mathfrak{C}_{\mathfrak{A}}(\mathfrak{B})$ are semi-simple, and $\mathfrak{C}_{\mathfrak{A}}(\mathfrak{B})$ is a self-adjoint subalgebra of $\mathfrak{A}$ since $\mathfrak{B}$ is self-adjoint. Therefore the set of self-adjoint elements of $\mathfrak{C}_{\mathfrak{A}}(\mathfrak{B})$ is a semi-simple Jordan algebra: $\mathfrak{C}_{\mathfrak{A}}(\mathfrak{B})$ is a direct sum of simple ideals which are either self-adjoint or interchanged in pairs by the involution, their self-adjoint elements are simple Jordan algebras of types A_1, B, C or type A_2 respectively. This completes the proof for special algebras $\mathfrak{J}$. The exceptional algebra case was proved at the end of Section 1.

As we saw in the above proof, if $\mathfrak{J}$ is of type A_2, i.e. $\mathfrak{J} = \mathfrak{A}_j$, $\mathfrak{A}$ a central simple associative algebra, and $\mathfrak{R}$ is a simple subalgebra of $\mathfrak{J}$ with enveloping algebra $\mathfrak{B}$, then $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{R}) = \mathfrak{C}_{\mathfrak{A}}(\mathfrak{B})_j$; also, if $\mathfrak{J}$ is central simple and $\mathfrak{R}$ contains the identity, then $\mathfrak{C}_{\mathfrak{A}}(\mathfrak{C}_{\mathfrak{A}}(\mathfrak{B})) = \mathfrak{B}$, so that $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{C}_{\mathfrak{J}}(\mathfrak{R})) = \mathfrak{B}_j$. Thus $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{C}_{\mathfrak{J}}(\mathfrak{R})) = \mathfrak{R}$ if and only if $\mathfrak{R} = \mathfrak{B}_j$ is also an algebra of type A_2, and the centralizer theory here is identical with the associative theory. The theory for algebras of degree two is not very interesting, since we will always have $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{R}) = \mathfrak{J}$, $\mathfrak{R}$, or center of $\mathfrak{J}$, as we have seen in the proof of the above theorem.

The remaining type, $\mathfrak{J} = H(\mathfrak{A})$, $\mathfrak{A}$ an involutorial ring of linear transformations, is the only interesting one. We note first of all that the double centralizer of a simple subalgebra $\mathfrak{R}$ may be actually larger than $\mathfrak{R}$, as is shown by the following examples:

1. Let $\mathfrak{R} = F \cdot 1 + V$, V even-dimensional, be an algebra of type D and let $\mathfrak{A}$ be the Clifford algebra determined by V. We may then assume that $\mathfrak{R}$ is contained in $\mathfrak{A}$. $\mathfrak{A}$ has an involution such that the elements of $\mathfrak{R}$ are self-adjoint; let $\mathfrak{J} = H(\mathfrak{A})$, then $\mathfrak{J}$ is a simple Jordan algebra of degree greater than two if V has dimension greater than two, and $\mathfrak{J}$ properly contains $\mathfrak{R}$. $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{R}) = F \cdot 1$ since $\mathfrak{R}$ generates $\mathfrak{A}$, and $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{C}_{\mathfrak{J}}(\mathfrak{R})) = \mathfrak{J}$. We may also embed V in a larger space W with symmetric non-degenerate scalar product, and take $\mathfrak{A}$ to be the Clifford algebra of W, $\mathfrak{J} = H(\mathfrak{A})$. If $\mathfrak{B}$ is the enveloping algebra of $\mathfrak{R}$, then $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{C}_{\mathfrak{J}}(\mathfrak{R}))$ will contain $H(\mathfrak{B})$, so will be larger than $\mathfrak{R}$. To exclude this possibility we will at least have to assume that if $\mathfrak{B}$ is the enveloping algebra of $\mathfrak{R}$, then $H(\mathfrak{B}) = \mathfrak{R}$, that is, that $\mathfrak{R}$ is not of type D.

2. Let $\mathfrak{J} = H(Q_n)$, the algebra of $n \times n$ hermitian matrices with quaternion coefficients, $n \geqq 3$, and $\mathfrak{R}$ the subalgebra of $n \times n$ symmetric matrices, $H(F_n)$. It is easy to see that $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{R}) =$ center of $\mathfrak{J} = F \cdot 1$, and $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{C}_{\mathfrak{J}}(\mathfrak{R})) = \mathfrak{J}$. Here $\mathfrak{J}$ is the set of self-adjoint linear transformations on an n-dimensional space and $\mathfrak{R}$ generates an irreducible algebra of linear transformations on the same space, i.e. $\mathfrak{R}$ is almost all of $\mathfrak{J}$.

To rule out such a case we will have to assume the dimension of the vector space on which $\mathfrak{J}$ operates is not too small relative to $\mathfrak{K}$.

3. Let $\mathfrak{J} = H(C_n)$, ordinary $n \times n$ complex-hermitian matrices for a suitably large n—an algebra over the field of real numbers, and $\mathfrak{K}$ the subalgebra of matrices with r equal $s \times s$ symmetric (i.e. with real coefficients) matrix blocks on the main diagonal and zeros elsewhere ($rs = n$). Then $\mathfrak{K}$ is isomorphic to $H(R_s)$, and if $r, s \geqq 2$, $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{K})$ is isomorphic to $H(C_r)$, and $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{C}_{\mathfrak{J}}(\mathfrak{K}))$ to $H(C_s)$ which properly contains $\mathfrak{K}$. To exclude this, we will assume that the center of the enveloping associative algebra of K contains the center of the enveloping associative algebra of $\mathfrak{J}$.

We start with a discussion of centralizers of simple subalgebras of a primitive ring with minimal ideals and identity element. Let M, N be dual left and right spaces over the division ring D with center F. Then $F \cdot 1$ is the center of the ring $L(M, N)$. Let $F(M, N)$ be the socle of $L(M, N)$, i.e. the linear transformations whose range is finite-dimensional. We will consider primitive rings $\mathfrak{R}$ containing $F \cdot 1$ and such that $L(M, N) \supseteqq \mathfrak{R} \supset F(M, N)$.

LEMMA 3.1. *Let $\mathfrak{A}$ be a ring of endomorphisms of a module M, and $\mathfrak{S}$ a right ideal in $\mathfrak{A}$ such that $M\mathfrak{S} = M$. If e is an endomorphism of M commuting with every element of $\mathfrak{S}$, then e commutes with every element of $\mathfrak{A}$.*

Proof. Let s, a belong to $\mathfrak{S}, \mathfrak{A}$ respectively. Then $s(ea - ae) = esa - sae$ since $es = se$, but since sa is in $\mathfrak{S}$, $e(sa) = (sa)e$, so $s(ea - ae) = 0$ for all s in $\mathfrak{S}$. Since $M\mathfrak{S} = M$, $M(ea - ae) = M\mathfrak{S}(ea - ae) = 0$. Thus $ea = ae$.

THEOREM 3.2. *Let $\mathfrak{A} = L(M, N) \supseteqq \mathfrak{R} \supseteqq F(M, N) + F \cdot 1$, and let $\mathfrak{B}$ be a simple finite-dimensional subalgebra, over F, of $\mathfrak{R}$ which contains $F \cdot 1$. Then $\mathfrak{C} = \mathfrak{C}_{\mathfrak{R}}(\mathfrak{B})$ is also a primitive ring with minimal ideals and identity element, and $\mathfrak{C}_{\mathfrak{R}}(\mathfrak{C}_{\mathfrak{R}}(\mathfrak{B})) = \mathfrak{B}$.*

Proof. The above theorem is a generalization of a result of Rosenberg, [15], but follows immediately from his result and the above lemma. $\mathfrak{C}_{\mathfrak{R}}(\mathfrak{B}) = \mathfrak{C}_{\mathfrak{A}}(\mathfrak{B}) \cap \mathfrak{R}$, since $\mathfrak{R} \subseteqq \mathfrak{A}$. As shown in [15], $\mathfrak{C}_{\mathfrak{A}}(\mathfrak{B})$ is a ring of the form $L(V, W)$ for dual spaces V and W, and the socle $\mathfrak{S}$ of $\mathfrak{C}_{\mathfrak{A}}(\mathfrak{B})$ is contained in $F(M, N)$. Since $F(M, N) \subset \mathfrak{R}$, $\mathfrak{S} \subset \mathfrak{R}$, and so $\mathfrak{S} \subset \mathfrak{C}_{\mathfrak{R}}(\mathfrak{B}) \subset \mathfrak{C}_{\mathfrak{A}}(\mathfrak{B}) = L(V, W)$. Thus $\mathfrak{C}_{\mathfrak{R}}(\mathfrak{B})$ is a primitive ring with minimal ideals and identity, since $\mathfrak{B}$ and $\mathfrak{R}$ contain the identity of $\mathfrak{A}$. This proves the first statement.

It is known that $M\mathfrak{S} = M$ ([15]). Since $\mathfrak{S}$ is a two-sided ideal in $\mathfrak{C}_{\mathfrak{A}}(\mathfrak{B})$, it follows from the lemma that $\mathfrak{C}_{\mathfrak{A}}(\mathfrak{S}) = \mathfrak{C}_{\mathfrak{A}}(\mathfrak{C}_{\mathfrak{A}}(\mathfrak{B}))$, and $\mathfrak{C}_{\mathfrak{A}}(\mathfrak{C}_{\mathfrak{A}}(\mathfrak{B})) = \mathfrak{B}$

by [15]. Thus $\mathfrak{C}_{\mathfrak{R}}(\mathfrak{S}) \subseteqq \mathfrak{C}_{\mathfrak{A}}(\mathfrak{S}) = \mathfrak{B}$. Since $\mathfrak{S} \subset \mathfrak{C}_{\mathfrak{R}}(\mathfrak{B})$, $\mathfrak{C}_{\mathfrak{R}}(\mathfrak{C}_{\mathfrak{R}}(\mathfrak{B})) \subseteqq \mathfrak{C}_{\mathfrak{R}}(\mathfrak{S}) \subseteqq \mathfrak{B}$, but $\mathfrak{C}_{\mathfrak{R}}(\mathfrak{C}_{\mathfrak{R}}(\mathfrak{B})) \supseteqq \mathfrak{B}$ always, so $\mathfrak{C}_{\mathfrak{R}}(\mathfrak{C}_{\mathfrak{R}}(\mathfrak{B})) = \mathfrak{B}$, which we had to prove.

The method of the above proof can be used to extend the Galois theory of rings $L(M, N)$, contained in [16], to primitive rings with minimal ideals.

We state now the main result of this section:

THEOREM 3.3. *Let M be a self-dual space* (see § 2 for the definition) *over a division ring D, E the center of D, and F the subfield of self-adjoint elements of E. Let $\mathfrak{A} = L(M, N)$, $\mathfrak{J} = H(\mathfrak{A})$, the Jordan algebra* (*over* $F \cdot 1$) *of self-adjoint linear transformations, and $\mathfrak{K}$ a simple subalgebra of $\mathfrak{J}$ of degree greater than two containing $F \cdot 1$ and finite dimensional over it. Assume that $\mathfrak{B}$, the enveloping associative algebra of $\mathfrak{K}$ in $\mathfrak{A}$, contains $E \cdot 1$. Then:*

1. *If $\mathfrak{K}$ is of type A_2, thus is isomorphic to $L(V, W)_j$ for a pair of dual spaces V and W, then $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{K})$ is isomorphic to $L(P, Q)_j$ for a pair of dual spaces P, Q.*

2. *If $\mathfrak{K}$ is of type A_1, B or C,* i.e. *is isomorphic to $H(L(V, V))$ for a self-dual space V, then $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{K})$ isomorphic to $H(L(P, P))$, P self-dual.*

3. *$\mathfrak{C}_{\mathfrak{J}}(\mathfrak{C}_{\mathfrak{J}}(\mathfrak{K})) = \mathfrak{K}$, provided the dimension of M is greater than twice the dimension over D of any minimal right ideal of $\mathfrak{B} \cdot D$, the ring of endomorphisms of M generated by $\mathfrak{B}$ and the scalar multiplications by elements of D.*

Proof. Let $\mathfrak{K}$ be of type A_2. Then its enveloping associative algebra $\mathfrak{B}$ is a homomorphic image of the "universal" enveloping algebra $\mathfrak{U}$ of $\mathfrak{K}$, and $\mathfrak{U} = \mathfrak{U}_1 \oplus \mathfrak{U}_2$ where $\mathfrak{U}_2$ is anti-isomorphic to $\mathfrak{U}_1$ and $\mathfrak{U}_1$ and $\mathfrak{U}_2$ are simple algebras. Thus either $\mathfrak{B}$ is isomorphic to $\mathfrak{U}_1$ or else $\mathfrak{B}$ is isomorphic to $\mathfrak{U}$. $\mathfrak{B}$ is a self-adjoint subalgebra of $\mathfrak{A}$ since $\mathfrak{K}$ consists of self-adjoint elements. Moreover the dimensions over F of $\mathfrak{U}_1$ and $\mathfrak{K}$ are the same, so if $\mathfrak{B}$ is isomorphic to $\mathfrak{U}_1$ then $\mathfrak{B} = \mathfrak{K}$. Since $\mathfrak{K}$ has degree at least three, $\mathfrak{B}$ is not commutative and so cannot consist only of self-adjoint elements. Thus $\mathfrak{B} = \mathfrak{B}_1 \oplus \mathfrak{B}_2$, where $\mathfrak{B}_i$ is isomorphic to $\mathfrak{U}_i$. Since $\mathfrak{B}$ is self-adjoint, either $\mathfrak{B}_1$ and $\mathfrak{B}_2$ are also self-adjoint, or else $\mathfrak{B}_2 = \mathfrak{B}_1'$, the image of $\mathfrak{B}_1$ under the involution in $\mathfrak{A}$. But if the $\mathfrak{B}_i$ are self-adjoint we again get a contradiction: for each k in $\mathfrak{K}$, write $k = k_1 + k_2$, k_i in $\mathfrak{B}_i$ and self-adjoint. Then the map k into k_1 is a homomorphism of Jordan algebras and so is either zero or an isomorphism since $\mathfrak{K}$ is simple. If it is an isomorphism it is onto $\mathfrak{B}_1$ since $\mathfrak{B}_1$ and $\mathfrak{K}$ have the same dimension over F. But then $\mathfrak{B}_1$ consists of self-adjoint elements and this is impossible, as before. If the map k into k_1 is zero then the map k into k_2 must be onto, which is equally impossible. Hence $\mathfrak{B} = \mathfrak{B}_1 \oplus \mathfrak{B}_1'$.

Let e_i be the identity of $\mathfrak{B}_i$, $i = 1, 2$. Then $e_2 = e_1'$ and $e_1 + e_2 = 1$, $e_1 e_2 = 0$. Let $M_i = Me_i$: then $M = M_1 \oplus M_2$ and the M_i are each totally isotropic, for $(Me_i, Me_i) = (Me_i e_i', Me_i) = 0$ since $e_i e_i' = 0$. As M is self-dual, M_1 and M_2 are dually paired by the scalar product in M.

Let $\mathfrak{C} = \mathfrak{C}_{\mathfrak{A}}(\mathfrak{B})$. Since $\mathfrak{B}$ contains e_1 and e_2, clearly

$$\mathfrak{C} = \mathfrak{C} \cap e_1\mathfrak{A}e_1 \oplus \mathfrak{C} \cap e_2\mathfrak{A}e_2 = \mathfrak{C}_1 \oplus \mathfrak{C}_2$$
$$= \text{centralizer of } \mathfrak{B}_1 \text{ in } e_1\mathfrak{A}e_1 \oplus \text{centralizer of } \mathfrak{B}_2 \text{ in } e_2\mathfrak{A}e_2.$$

Since $\mathfrak{B}_2 = \mathfrak{B}_1'$, $\mathfrak{C}_2 = \mathfrak{C}_1'$. Also, by § 3.20 of [16], $e_1\mathfrak{A}e_1$ is isomorphic to $L(Me_1, Me_1') = L(M_1, M_2)$ and in the same way $e_2\mathfrak{A}e_2$ is isomorphic to $L(M_2, M_1)$. $\mathfrak{C}_1$, being the centralizer in $L(M_1, M_2)$ of the simple subalgebra $\mathfrak{B}_1$ which contains the center $E \cdot 1$, is a ring $L(N_1, N_2)$, by [15], and $\mathfrak{C}_2 = \mathfrak{C}_1'$. Thus $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{K}) = H(\mathfrak{C}_1 \oplus \mathfrak{C}_1')$ is isomorphic to $(\mathfrak{C}_1)_J$, i.e. to $L(N_1, N_2)_J$. This proves statement 1.

Let $\mathfrak{K}$ be of type A_2 as above, and let M satisfy the dimensionality condition of 3. Since E is the center of D and $\mathfrak{B}$ contains $E \cdot 1$, $\mathfrak{B} \cdot D$ is isomorphic to $\mathfrak{B} \otimes_E D$ and so $\mathfrak{B} \cdot D = \mathfrak{B}_1 \cdot D \oplus \mathfrak{B}_2 \cdot D = \mathfrak{B}_1 \otimes_E D \oplus \mathfrak{B}_2 \otimes_E D$. As the $\mathfrak{B}_i$ are simple and contain Ee_i in their centers, $\mathfrak{B}_i \cdot D$ is a simple ring with minimum condition operating on M_i and therefore M_i is completely reducible and homogeneous as $\mathfrak{B}_i \cdot D$ (right) module. By the dimensionality assumption, M_i is a direct sum of at least two irreducible submodules. Let M be finite dimensional, then $\mathfrak{C}_i$ is the ring of all endomorphisms of M commuting with $\mathfrak{B} \cdot D$ and so is a matrix ring G_n over a division ring G with $n \geqq 2$, i.e. $\mathfrak{C}_1 = L(N_1, N_2)$ and N_1 is a vector space of dimension greater than one. If M is infinite dimensional, so is N_1. From this it follows that the enveloping ring in $\mathfrak{A}$ of $\mathfrak{C} \cap \mathfrak{J}$ contains the socles of $\mathfrak{C}_1$ and $\mathfrak{C}_2$: for $\mathfrak{C}_1$ is not commutative and so contains elements a and b with $ab \neq ba$. Then $a + a'$, $b + b'$ and $ba + a'b'$ all belong to $\mathfrak{C} \cap \mathfrak{J}$, so $(a + a')(b + b') - (ba + a'b') = ab - ba$ is in the enveloping ring of $\mathfrak{C} \cap \mathfrak{J}$. Also $(ab - ba)(g + g') = (ab - ba)g$ and $(g + g')(ab - ba) = g(ab - ba)$ are in this enveloping ring for all g in $\mathfrak{C}$ (since $\mathfrak{C}\mathfrak{C}' = (0) = \mathfrak{C}'\mathfrak{C}$), and so this ring contains a non-zero two-sided ideal in $\mathfrak{C}_1$. Since every two-sided ideal contains the socle, the enveloping ring does also. If V, W are dual spaces and $\mathfrak{R}$ is a subalgebra of $L(V, W)$ which is the centralizer of a simple finite dimensional subalgebra, then the socle F of $\mathfrak{R}$ satisfies $VF = V$ (see [15]). Therefore if F_1 is the socle of $\mathfrak{C}_1$, $M_1F_1 = M_1$. By Lemma 3.1, the centralizer of F_1 in $L(M_1, M_2)$ is also the centralizer of $\mathfrak{C}_1$. Thus $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{C}_{\mathfrak{J}}(\mathfrak{K})) = \mathfrak{J} \cap \mathfrak{C}_{\mathfrak{A}}(\mathfrak{C})$. But $\mathfrak{C}_{\mathfrak{A}}(\mathfrak{C})$ is just $\mathfrak{C}_{\mathfrak{A}}(\mathfrak{C}_1 \oplus \mathfrak{C}_2)$ and the latter is $\mathfrak{B}_1 \oplus \mathfrak{B}_2$ since $\mathfrak{B}_1$ is the centralizer of $\mathfrak{C}_1$ in $e_1\mathfrak{A}e_1 = L(M_1, M_2)$ and similarly for $\mathfrak{B}_2$. Therefore $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{C}_{\mathfrak{J}}(\mathfrak{K})) = \mathfrak{B} \cap \mathfrak{J}$ and it is known that $\mathfrak{B} \cap \mathfrak{J} = H(\mathfrak{B}_1 \oplus \mathfrak{B}_1') = \mathfrak{K}$ since $\mathfrak{K}$ is of type A_2.

Next let $\mathfrak{K}$ be of type A_1, B, or C. Then $\mathfrak{B}$ is simple and $\mathfrak{K} = H(\mathfrak{B})$. $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{K}) = \mathfrak{C}_{\mathfrak{A}}(\mathfrak{B}) \cap \mathfrak{J}$. But $\mathfrak{C}_{\mathfrak{A}}(\mathfrak{B})$ is self-adjoint and so is of the form

$L(P, P)$ for a self-dual space P, since $\mathfrak{B}$ is simple and contains the center $E \cdot 1$ of $\mathfrak{A}$. This proves statement 2.

Let now M satisfy the dimension condition of 3. $\mathfrak{B} \cdot D$ is a simple ring with minimum condition since $\mathfrak{B}$ is simple and contains $E \cdot 1$, and so is homogeneous completely reducible on M, and M is a direct sum of at least three irreducible submodules, i.e. P has dimension at least three. Write $\mathfrak{C}$ for $\mathfrak{C}_{\mathfrak{A}}(\mathfrak{B})$ and F for the socle of $\mathfrak{C}$. Then F is locally canonically matrix of degree $\geqq 3$, [12], and so is generated by its self-adjoint elements—i.e. the enveloping ring of $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{K}) = H(\mathfrak{C})$ contains F. Therefore $\mathfrak{C}_{\mathfrak{A}}(F) = \mathfrak{C}_{\mathfrak{A}}(\mathfrak{C}) = \mathfrak{C}_{\mathfrak{A}}(\mathfrak{C}_{\mathfrak{J}}(\mathfrak{K}))$ and $\mathfrak{C}_{\mathfrak{A}}(\mathfrak{C}) = \mathfrak{B}$. Thus $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{C}_{\mathfrak{J}}(\mathfrak{K})) = \mathfrak{C}_{\mathfrak{A}}(\mathfrak{C}) \cap \mathfrak{J} = \mathfrak{B} \cap \mathfrak{J}$, and finally $\mathfrak{B} \cap \mathfrak{J} = \mathfrak{K}$ since $\mathfrak{J}$ is of type A_1, B or C. Thus $\mathfrak{C}_{\mathfrak{J}}(\mathfrak{C}_{\mathfrak{J}}(\mathfrak{K})) = \mathfrak{K}$. This concludes the proof.

REFERENCES

1. A. A. Albert, *A structure theory for Jordan algebras*, Ann. of Math. **48** (1947), 446–467.
2. ———, *A theory of power-associative commutative algebras*, Trans. Amer. Math. Soc. **69** (1950), 503–527.
3. C. Chevalley, *The algebraic theory of spinors*, New York, 1954.
4. J. Dieudonné, *Sur la reduction canonique des couples de matrices*, Bull. Soc. Math. France, **74** (1946), 130–146.
5. H. Freudenthal, *Oktaven, Ausnahmengruppen, und Oktavengeometrie*, Utrecht, 1951.
6. N. Jacobson, *Lectures in abstract algebra*, Vol. 2, New York 1953.
7. ———, *Structure of rings*, Amer. Math. Soc., 1956.
8. ———, *Operator commutativity in Jordan algebras*, Proc. Amer. Math. Soc., **3** (1952), 973-976.
9. ———, *Structure of alternative and Jordan bimodules*, Osaka Math. Journal, **6** (1954), 1–71.
10. ———, *On the structure of Jordan algebras*, Proc. Nat. Acad. Sci. U.S.A., **42** (1956), 140-147.
11. N. Jacobson and C. E. Rickart, *Jordan homomorphisms of rings*, Trans. Amer. Math. Soc., **69** (1950), 479–502.
12. ———, *Homomorphisms of Jordan rings of self-adjoint elements*, Trans. Amer. Math. Soc., **72** (1952), 310–322.
13. A. I. Malcev, *Foundations of linear algebra*, Moscow-Leningrad, 1947.
14. N. H. McCoy, *On quasi-commutative matrices*, Trans. Amer. Math. Soc., **35** (1934), 332.
15. A. Rosenberg, *Finite dimensional simple subalgebras of the ring of all continuous linear transformations*, Math. Z. **61** (1954), 150–159.
16. A. Rosenberg and D. Zelinsky, *Galois theory of continuous transformation rings*, Trans. Amer. Math. Soc., **79** (1955), 429–452.
17. P. Jordan, J. von Neumann, and E. Wigner, *On an algebraic generalization of the quantum mechanical formalism*, Ann. of Math., **35** (1934), 29–64.

YALE UNIVERSITY

MODULUS OF A BOUNDARY COMPONENT

MARTIN JURCHESCU

§1. PRELIMINARIES AND SUMMARY

1.1 Preliminary definitions. Let R be an open Riemann surface, and let $\{G_n\}$ $(n = 1, 2, \cdots)$ be an infinite sequence of subregions of R such that:

(a) the relative boundary of each G_n is compact,

(b) $G_n \supset G_{n+1}$, and

(c) $\bigcap_{n=1}^{\infty} \overline{G}_n = 0$.

$\{G_n\}$ is said to define a *boundary component* γ of R in the sense of Kerékjártó [6] and Stoilow [16]. Here two sequences of subregions $\{G_n\}$ and $\{G'_n\}$ are considered to be equivalent and to define the same γ if each region G_n includes a region G'_m. That this is a proper equivalence relation follows immediately.

Let γ be a boundary component of R, and let S be a subregion of R. If there exists a defining sequence $\{G_n\}$ of γ with $G_{n_0} = S$, for some n_0, we call S a *neighborhood of* γ. Throughout this paper we shall consider only neighborhoods S of γ such that the relative boundary of S is a closed analytic Jordan curve γ_0.

By an *exhaustion* of R, we mean an infinite sequence $\{R_n\}$ $(n = 1, 2, \cdots)$ of subregions of R as follows (see [16]):

(1) each R_n is compact relative to R and the relative boundary β_n of R_n consists of a finite number of closed analytic Jordan curves β_{ni},

(2) $R_n \subset R_{n+1}$,

(3) $\bigcup^{\infty} R_n = R$, and

(4) each connected component S_{ni} of $R - \overline{R}_n$ is non-compact (relative to R) and its boundary consists of a single curve β_{ni}.

Each set $R - \overline{R}_n$ is said to be a *boundary neighborhood* of R. It is easy to see that, for any boundary component γ of R, there exists a single connected component S_{ni} which is a neighborhood of γ.

A property is said to be a *boundary property* (respectively a *γ-property*) if the following is true. If a Riemann surface R has the property then every Riemann surface R' which admits a conformal mapping from a boundary neighborhood of R' (a neighborhood of γ', where γ' is a boundary

Received April, 1, 1958, and in received form May 1, 1958. This paper is a part of the author's dissertation realized under the guidance of Professor Stoilow at Mathematical Institute of the R. P. P. Academy, Bucarest. The author wishes to express his gratitude to Professor Stoilow for his encouragement and permanent support. The author is also indebted to Dr. Kotaro Oikawa of Tokyo Institute of Technology who read the manuscript and made valuable remarks.

component of R') onto a boundary neighborhood of R (a neighborhood of γ) has the property.

Let u be a harmonic function on a subregion S of R. We shall denote by $\bar{u}$ the conjugate harmonic function of u and by $D(u; S)$ the Dirichlet integral of u over S.

1.2. Capacity of a boundary component. Let γ be a boundary component of an open Riemann surface R, P_0 a point of R, and $K_z: |z| \leqq 1$ a fixed parametric disc on R with $z = 0$ corresponding to P_0. Let $\{R_n\}$ be an exhaustion of R with $P_0 \in R_1$, and let γ_n denote the curve β_{ni} which separates γ from P_0. This means that γ_n separates a neighborhood of γ from P_0.

We consider the class $\{t\}_\gamma$ of single-valued functions on R which satisfy the following conditions:

(1.1) each t is harmonic on $R - P_0$ and has the form

$$t = \log |z| + h(z)$$

in K_z, where h is harmonic and $h(0) = 0$.

$$\int_{\gamma n} d\bar{t} = 2\pi \text{ and } \int_{\beta ni \neq \gamma n} d\bar{t} = 0 , \qquad \text{for all } n , \tag{1.2}$$

where γ_n and β_{ni} are described in the positive sense with respect to R_n.

We further consider the corresponding class $\{t\}_{\gamma n}$ on R_n, and we denote by t_n the function of this class with $t_n = k_n$ on γ_n and $t_n = k_{ni}$ on $\beta_{ni} \neq \gamma_n$, where k_n and k_{ni} are real numbers.

The following theorem due to Sario is proved in [14] (see also Savage [15]). Let $t \in \{t\}_\gamma$, and let

$$I(t) = \lim \frac{1}{2\pi} \int_{\beta n} t d\bar{t} \ .$$

THEOREM 1. *The sequence of functions* $\{t_n\}$ *is compact. Let* t_γ *denote a limit function of* $\{t_n\}$. *Then we have the following conclusions*:

$$t_\gamma \in \{t\}_\gamma \textit{ and, for any } t, \ \min I(t) = I(t_\gamma) \ . \tag{1.3}$$

$$I(t) = I(t_\gamma) + D(t - t_\gamma; R) \ . \tag{1.4}$$

$$k_n \leqq k_{n+1} \textit{ and } I(t_\gamma) = \lim k_n \equiv k_\gamma \ . \tag{1.5}$$

By (1.4), for $k_\gamma < \infty$, the minimizing function t_γ is unique. t_γ is called the *capacity function* of R for γ, and the quantity $c_\gamma = e^{-k\gamma}$ is called the *capacity* of γ (with respect to K_z). Let $z' = az + \cdots, a \neq 0$, be a new local parameter in the neighborhood of P_0, and let c'_γ denote the capacity of γ with respect to this local parameter. It follows, from the definition of the capacity, that

$$c_\gamma = |a|\, c'_\gamma \,. \tag{1.6}$$

Hence, the condition $c_\gamma = 0$ is independent of the local parameter which is used in the neighborhood of P_0. Using Green's formula, it is easy to see that this condition is also independent of P_0. A boundary component γ is called *weak* if it has a capacity $c_\gamma = 0$. The class of Riemann surfaces for which all γ are weak is denoted by C_γ. The boundary of a Riemann surface R belonging to C_γ is called *absolutely disconnected* [14, 15].

1.3. Summary. Let R be an open Riemann surface, γ a boundary component of R, S a neighborhood of γ, and γ_0 the relative boundary of S. The present paper deals with a conformal invariant of S which is denoted by $\mu(S;\gamma_0,\gamma)$ (or, simply, for fixed S, by μ_γ) and is called the *modulus of* S *for* γ_0 *and* γ (the *modulus of* γ).

In §2 harmonic functions u on S with $u = 0$ on γ_0 and satisfying conditions (2.3) are considered, and a theorem is proved which establishes the existence of a minimizing function $u_\gamma = u(z;S;\gamma_0,\gamma)$ for the Dirichlet integral $D(u;S)$. The modulus is defined by setting $\mu_\gamma = D(u_\gamma;S)$. The notion of a parabolic boundary component is defined by the condition $\mu_\gamma = \infty$, and a theorem is proved which shows the equivalence of parabolicity and weakness.

In §3 measurable conformal metrics are considered. An important minimal property of the conformal metric $\rho_\gamma = |\operatorname{grad} u_\gamma|$ corresponding to a result of Wolontis [17] and Strebel [18] is proved, which connects μ_γ with the extremal length of a certain family of curves on S. As an application, a characterization of a parabolic boundary component is obtained in terms of conformal metrics. Another characterization of a parabolic boundary component is given by means of the divergence of a modular series $\sum \mu(E_n;\gamma_{n-1},\gamma_n)$. The sufficient part of this theorem implies the modular criterion of Savage [15]. A theorem shows the equivalence of perimeter in Ahlfors and Beurling's sense and capacity in Sario's sense.

Section 4 deals with the class M_γ of Riemann surfaces for which all γ are parabolic in the case of a finite genus. The conformal mapping properties of u_γ and t_γ are discussed, and, for planar Riemann surfaces, the equalities $O_{SB} = M_\gamma = O_{SD}$ [1, 14] are proved. Finally a theorem is proved which shows the connection between M_γ and the class of Riemann surfaces for which the continuation is topologically unique, or which do not possess essential continuations.

§2. Harmonic Functions and Modulus

2.1. Moduli of a compact subregion. Let S_0 denote a relatively compact subregion of a Riemann surface R. We assume that the boundary of S_0 is a set $\gamma_0 \cup \alpha_0$, where γ_0 is a closed analytic Jordan curve and α_0 consists of a finite number of closed analytic Jordan curves $\alpha_{01}, \cdots, \alpha_{0k}$ $(k \geqq 1)$. We assign to each $\alpha_{0i} (i = 1, \cdots, k)$ as positive orientation the positive sense with respect to S_0 and to γ_0 the sense for which γ_0 and α_0 are homologous.

If u is a harmonic function on S_0 then we denote the conjugate period of u around α_{0i} by $p_i(u)$. This is defined by the integral $\int_{\alpha'_{0i}} d\bar{u}$, where α'_{0i} is any closed Jordan curve on S_0 such that α_{0i} and α'_{0i} are homologous. If u is harmonic on $S_0 \cup \alpha_{0i}$ then clearly $p_i(u) = \int_{\alpha_{0i}} d\bar{u}$. The period vector $(p_1(u), \cdots, p_k(u))$ will be denoted by $p(u)$.

LEMMA 1. *There is a harmonic function* $u_0 = u(z; S_0; \gamma_0, k_{01})$ *on* S_0 *satisfying the following conditions*:

(a) $u_0 = 0$ *on* γ_0 *and* $u_0 = \mu_{0i} = const.$ *on* $\alpha_{0i} (i = 1, \cdots, k)$,

(b) $p(u_0) = (1, 0, \cdots, 0)$.

(c) $0 < u_0(z) < \mu_{01}$ *on* S_0 *and on the boundary curves* $\alpha_{02}, \cdots, \alpha_{0k}$.

Proof. Denote the harmonic measure of α_{0i} with respect to S_0 by ω_i, and consider the function

$$u(z) = \sum_{i=1}^{k} \mu_i \omega_i(z) , \tag{2.1}$$

where μ_i are arbitrary real numbers. Clearly, this function is harmonic on $\bar{S}_0 = S_0 \cup \gamma_0 \cup \alpha_0$. Setting $a_{ij} = p_i(\omega_i)$, we obtain

$$p_i(u) = \int_{\alpha_{0i}} d\bar{u} = \sum_{j=1}^{k} a_{ij} \mu_j .$$

We assert that this linear mapping of the k-dimensional cartesian space into itself is one-to-one. In fact, from Green's formula

$$D(u) \equiv D(u; S_0 = \sum_{i=1}^{k} \int_{\alpha_{0i}} u d\bar{u} = \sum_{i=1}^{k} \mu_i p_i(u) ,$$

we see that the condition $p_i(u) = 0$, for all i, implies $D(u) = 0$, that is $u \equiv 0$ (since $u = 0$ on γ_0) and consequently $\mu_i = 0$, for all i, which proves our assertion. Hence we deduce in particular that the above linear mapping is onto, i.e., for any p, there is a function $u = \sum \mu_i \omega_i(z)$ such that $p(u) = p$. Let u_0 denote the function (1.1) corresponding to $p_0 = (1, 0, \cdots, 0)$. This is clearly the unique bounded harmonic function on S_0 satisfying (a) and (b).

Now denote the maximum and the minimum of u_0 on the boundary of S_0 by M_0 and m_0 respectively. From the maximum principle, we have

$m_0 < u_0(z) < M_0$ on S_0. It follows that $\partial u_0/\partial n \leqq 0$ on each boundary curve $\gamma(M_0)$ on which $u_0(z) = M_0$. Here $\partial/\partial n$ denotes the derivative in the direction of the interior normal. Since u_0 is not constant and $\partial u_0/\partial n$ is continuous, there exists a subarc of $\gamma(M_0)$ on which $\partial u_0/\partial n < 0$ and therefore

$$\int_{\gamma(M_0)} d\bar{u}_0 = -\int_{\gamma(M_0)} \frac{\partial u_0}{\partial n}|dz| > 0 ,$$

where $\gamma(M_0)$ is described in the positive sense with respect to S_0. This and condition (b) implies that $\gamma(M_0)$ coincides necessarily with α_{01}, whence $M_0 = \mu_{01}$ and this maximum is attained only on α_{01}. Similarly, it can be proved that $m_0 = 0$ and that this minimum is attained only on γ_0 This completes the proof of Lemma 1.

LEMMA 2. *The function u_0 gives the minimum of $D(u)$,*

$$\min D(u) = D(u_0) ,$$

in the class of all harmonic functions u on S_0 with $u = 0$ on γ_0 and $p(u) =$ $(1, 0, \cdots, 0)$.

Proof. Clearly, the function u_0 belongs to the class of admissible functions and, by Green's formula,

$$D(u_0) = \sum_{i=1}^{k} \mu_{0i} p_i(u_0) = \mu_{01} < \infty .$$

Let u be any admissible function with $D(u) < \infty$. Setting $u - u_0 = h$, we have

$$D(u) = D(u_0) + D(h) + 2D(u_0, h) ,$$

where $D(u_0, h) = D(u_0, h; S_0)$ is the mixed Dirichlet integral of u_0 and h over S_0. We shall show that $D(u_0, h) = 0$. If u is harmonic on $\overline{S}_0$ then Green's formula gives immediately

$$D(u_0, h) = \int_{\alpha_0} u_0 d\bar{h} = \sum_{i=1}^{k} \mu_{0i}\, p_i(h) = 0$$

since, for all i, $p_i(h) = p_i(u) - p_i(u_0) = 0$. If the above assumption is not true, we consider the open set $S_0(\varepsilon) = S_0 - \cup_{i=1}^{k} E_{0i}(\varepsilon)$, where ε is a positive number, sufficiently small, and $E_{0i}(\varepsilon)$ is the set (of points of S_0 for which) $\mu_{0i} - \varepsilon \leqq u_0(z) \leqq \mu_{0i} + \varepsilon$. The boundary of $S_0(\varepsilon)$ consists only of level lines of u_0. On the other hand each level line $c(\mu): u_0(z) = \mu$ $(0 < \mu < \mu_{01}$, $\mu \neq \mu_{0i}, i = 1, \cdots, k)$ is a dividing cycle on S_0 (that is, $c(\mu)$ is homologous with a sum of α_{0i}) and therefore $\int_{c(\mu)} d\bar{h} = 0$. Hence, Green's formula gives again $D(u_0, h; S_0(\varepsilon)) = 0$ and, as $\varepsilon \to 0$, $D(u_0, h) = 0$. We conclude finally that

$$(2.2) \qquad D(u) = D(u_0) + D(u - u_0) ,$$

which proves our lemma.

The uniqueness of the minimizing function u_0 is an immediate consequence of (2.2). For, if $D(u) = D(u_0)$, we conclude from (2.2) that $D(u - u_0) = 0$, that is $u \equiv u_0$, since $u - u_0 = 0$ on γ_0.

The function $u_0 = u(z; S; \gamma_0, \alpha_{01})$ will be called the *extremal function* of S_0 for γ_0 and α_{01}. The quantity $\mu_{01} = D(u_0)$ will be called the *modulus* of S_0 for γ_0 and α_{01} and denoted generally by $\mu(S_0; \gamma_0, \alpha_{01})$.

2.2. Modulus of a boundary component. Let us consider a boundary component γ of an open Riemann surface R, and let S be a given neighborhood of γ. Let γ_0 be the relative boundary of S (see 1.1). An exhaustion of S is a sequence $\{S_n\}$ $(n = 1, 2, \cdots)$ of subregions of R such that: (1) S_n is a relatively compact subregion of R and the relative boundary of S_n is a set $\gamma_0 \cup \alpha_n$, where $\gamma_0 \cap \alpha_n = 0$ and α_n consists of a finite number of closed analytic Jordan curves α_{ni}, (2) $S_n \subset S_{n+1}$, (3) $\bigcup_{n+1}^{\infty} S_n = S$, and (4) each connected component of $S - S_n$ is non-compact and its relative boundary consists of a single α_{ni}. We assign to each α_{ni} as positive orientation the positive sense with respect to S_n and to γ_0 the sense for which γ_0 and α_n are homologous.

Let γ_n be the curve α_{ni} which separates γ from γ_0, and let $\{n\}_\gamma$ be the class of all harmonic functions u on S with $u = 0$ on γ_0 and

$$(2.3) \qquad \int_{\gamma_n} d\bar{u} = 1 \text{ and } \int_{\alpha_{ni} \neq \gamma_n} d\bar{u} = 0 ,$$

for all n. It is easy to see, using Green's formula, that conditions (2.3) are independent of the particular exhaustion which is used.

THEOREM 2. *In $\{u\}_\gamma$ there exists a function u_γ with the property*

$$\min D(u; S) = D(u_\gamma; S) .$$

Moreover, for any u,

$$(2.4) \qquad D(u; S) = D(u_\gamma; S) + D(u - u_\gamma; S) .$$

Proof. Denote by u_n the extremal function of S_n for γ_0 and γ_n, and put $\mu_n = D(\mu_n; S_n) =$ value of u_n on γ_n; μ_n is the modulus of S_n for γ_0 and γ_n.

Since the restriction of u_{n+1} to S_n satisfies the condition of Lemma 2 (where S_0 is replaced by S_n and α_{01} by γ_n), we have

$$\mu_n = D(u_n; S_n) \leqq D(u_{n+1}; S_n) \leqq D(u_{n+1}; S_{n+1}) = \mu_{n+1} .$$

Similarly, we see that $\mu_n \leqq \mu_\gamma$, where μ_γ is the greatest lower bound of

$D(u; S)$ for u *in* $\{u\}_\gamma$. Thus, $\lim_{n\to\infty} \mu_n$ exists and we have

$$\lim_{n\to\infty} \mu_n \leqq \mu_\gamma .$$

For a fixed N, let s be the bounded harmonic function on S_N with $s = 0$ on γ_0 and $s = d$ on α_N, where d is a constant value determined by $\int_{\alpha_N} d\bar{s} = 1$. From Green's formula $\int_{\alpha_N} u_n d\bar{s} - s d\bar{u}_n) = 0$ and the boundary behavior of u_n and s, we obtain

$$\int_{\alpha_N} u_n d\bar{s} = d ,$$

for all $n \geqq N$, whence $\min_{\alpha_N} u_n \leqq d$. It follows from Harnack's principle that the sequence $\{u_n\}$ is compact. A subsequence, say again $\{u_n\}$, converges, uniformly on each S_N, to a function u_γ. Obviously this function belongs to $\{u\}_\gamma$, so that

$$\mu_\gamma \leqq D(u_\gamma; S) .$$

On the other hand, the lower semicontinuity of the Dirichlet integral gives

$$D(u_\gamma; S) \leqq \lim D(u_n; S_n) = \lim \mu_n .$$

From the three preceding inequalites we conclude that

$$D(u_\gamma; S) = \lim \mu_n = \mu_\gamma ,$$

which proves the first assertion of Theorem 2.

Let us now prove equality (2.4), for any u in $\{u\}_\gamma$. This is evident if $D(u; S) = \infty$. Suppose $D(u; S) < \infty$, and put $u - u_\gamma = h$. For any real number ε, $u_\gamma + \varepsilon h \in \{u\}_\gamma$, and therefore

$$D(u_\gamma + \varepsilon h) = D(u_\gamma) + 2\varepsilon D(u_\gamma, h) + \varepsilon^2 D(h) \geqq D(u_\gamma) .$$

Since $D(u_\gamma + \varepsilon h) < \infty$, this is possible only if $D(u_\gamma, h) = 0$, so that, as $\varepsilon = 1$, we obtain (2.4).

As in Lemma 2, the uniqueness of the minimizing function u_γ in the case $\mu_\gamma < \infty$ is an immediate consequence of (2.4).

The function u_γ will be called the *extremal function* of S for γ_0 and γ and denoted generally by $u(z; S; \gamma_0, \gamma)$. The conformal invariant $\mu_\gamma = D(u_\gamma, S)$ will be called the *modulus* of S for γ_0 and γ or, simply, for fixed S, the modulus of γ. It will be denoted generally by $\mu(S; \gamma_0, \gamma)$.

2.3. Parabolic boundary components. Let γ be a boundary component of an open Riemann surface R. Consider any two neighborhoods S and S' of γ, and denote by γ_0 and γ'_0 the relative boundaries of S and

S' respectively. Set $u(z;S;\gamma_0,\gamma) = u_\gamma$, $u(z;S';\gamma'_0,\gamma) = u'_\gamma$, $\mu(S;\gamma_0,\gamma) = \mu_\gamma$, $\mu(S';\gamma'_0,\gamma) = \mu'_\gamma$.

LEMMA 3. *The moduli μ_γ and μ'_γ are simultaneously finite or infinite.*

Proof. Suppose first $S \subset S'$, and let $\{S'_n\}$ be an exhaustion of S'. The regions $S_n = S \cap S'_n$ give, for n sufficiently large, an exhaustion of S. Set $u(z;\gamma_0,\gamma_n) = u_n$, $u(z;S'_n;\gamma'_0,\gamma_n) = u'_n$, $\mu(S_n;\gamma_0,\gamma_n) = \mu_n$, $\mu(S'_n;\gamma'_0,\gamma_n) = \mu'_n$.
From Green's formula

$$\int_{\alpha_n \cup \gamma_0^{-1}} (u'_n d\bar{u}_n - u_n d\bar{u}'_n) = 0 ,$$

it follows

$$\mu'_n - \mu_n = \int_{\gamma_0} u'_n d\bar{u}_n .$$

Hence, as $n \to \infty$, we obtain

$$\mu'_\gamma - \mu_\gamma = \int_{\gamma_0} u'_\gamma d\bar{u}_\gamma .$$

This proves our lemma in the particular case $S \subset S'$.

Let us now consider the general case, and construct a third neighborhood S'' of γ such that $S'' \subset S \cap S'$. Let γ''_0 denote the relative boundary of S'', and put $\mu(S'';\gamma''_0,\gamma) = \mu''_\gamma$. As before, μ_γ and μ''_γ are simultaneously finite or infinite. The same is valid for μ'_γ and μ''_γ and consequently for μ_γ and μ'_γ, which completes the proof of Lemma 3.

A boundary component γ of R is called *parabolic* if $\mu_\gamma = \infty$ and *hyperbolic* if $\mu_\gamma < \infty$. From Lemma 3, this condition is independent of the neighborhood S which is used, i.e. the parabolicity of a γ is a γ-property of R. The class of all Riemann surfaces for which all boundary components are parabolic will be denoted by M_γ. The property $R \in M_\gamma$ (or $R \notin M_\gamma$) is a boundary property of R.

Consider now the capacity function t_γ of R for γ with respect to a fixed parametric disc $|z| \leqq 1$. Let λ denote a positive number which is sufficiently small such that the level line $c(\lambda)$: $t_\gamma(z) = \log\lambda$ is a closed Jordan curve and the set $t_\gamma(z) \leqq \log\lambda$ is compact. The set $S(\lambda)$: $t_\gamma(z) > \log\lambda$ is then a neighborhood of γ. Put $u(z;S(\lambda);c(\lambda),\gamma) = u_{\gamma,\lambda}$, $\mu(S(\lambda);c(\lambda),\gamma) = \mu_{\gamma,\lambda}$.

LEMMA 4. *If λ satisfies the above conditions, then*

$$t_\gamma(z) - \log\lambda = 2\pi u_{\gamma,\lambda}(z) , \tag{2.5}$$

and

$$k_\gamma - \log\lambda = 2\pi\mu_{\gamma,\lambda}\ . \tag{2.6}$$

Proof. Consider an exhaustion $\{R_n\}$ of R as in 2.1. The regions $S_n(\lambda) = R_n \wedge S(\lambda)$ give, for n sufficiently large, an exhaustion of $S(\lambda)$. Set $u(z; S_n(\lambda); c(\lambda), \gamma_n) = u_{n,\lambda}$, $\mu(S_n(\lambda); c(\lambda), \gamma_n) = \mu_{n,\lambda}$, $t_\gamma - 2\pi u_{\gamma,\lambda} = h$, $t_n - 2\pi u_{n\pi} = h_n$, where t_n is the function on R_n defined in 1.2. From Green's formula, we have

$$D(h_n ; S_n(\lambda)) = \int_{\beta n} h_n d\bar{h}_n - \int_{c(\lambda)} h_n d\bar{h} = - \int_{c(\lambda)} h_n d\bar{h}_n\ ,$$

since h_n = const. on β_{ni} and $\int_{\beta_{ni}} d\bar{h}_n = 0$, for all β_{ni}. Hence, by the lower semicontinuity of the Dirichlet integral,

$$D(h; S(\lambda)) \leqq -\int_{c(\lambda)} hd\bar{h} = 0\ ,$$

since h = const. = $\log\lambda$ on $c(\lambda)$ and $\int_{c(\lambda)} d\bar{h} = 0$. We conclude that $h \equiv \log\lambda$, which proves (2.5).

Now apply Green's formula on $S_n(\lambda)$ to $u_{n,\lambda}$ and t_n. We obtain

$$k_n - 2\pi\mu_{n,\lambda} = \int_{c(\lambda)} t_n d\bar{u}_{n,\lambda}\ ,$$

whence, as $n \to \infty$,

$$k_\gamma - 2\pi\mu_{\gamma,\lambda} = \int_{c(\lambda)} t_\gamma d\bar{u}_{\gamma,\lambda} = \log\lambda\ ,$$

which completes the proof of Lemma 4.

THEOREM 3. *A boundary component γ of R is parabolic if and only if it has a vanishing capacity.*

Proof. This is evident from Lemmas 3 and 4.

COROLLARY. $M_\gamma = C_\gamma$.

§3 MODULUS AND CONFORMAL METRICS

3.1. Definitions. Consider a non-negative function $\rho(z)$ which is defined on each parametric disc K_z: $|z| \leqq 1$ of a subregion S of R and satisfies

$$\rho(z) = \rho(z')\left|\frac{dz'}{dz}\right|$$

at corresponding points z, z' of any two overlapping K_z and $K_{z'}$. We say that ρ is a conformal metric on S. We define the ρ-length of any cycle c (finite set of closed Jordan curves) on S by the lower Darboux integral (see [4])

$$l(\rho; c) = \int_c \rho(z)|dz| .$$

A conformal metric ρ is said to be measurable on S if its restriction to any parametric disc is measurable in Lebesgue's sense. If ρ is a measurable conformal metric on S, we define the ρ-area of S by the Lebesgue integral

$$\mathrm{A}(\rho; S) = \int_S \rho^2(z) d\sigma_z ,$$

where σ_z is the Lebesgue measure on K_z. A measurable conformal metric ρ defined on S is said to be A-bounded on S if $A(\rho; S) < \infty$.

3.2. Extremal conformal metrics. Consider first the relatively compact subregion S_0 of 2.1. We prove the following

LEMMA 5. *The conformal metric* $\rho_0 = |\mathrm{grad} u_0|$ *gives the minimum of* $A(\rho; S_0)$,

$$\min A(\rho; S_0) = A(\rho_0; S_0) , \tag{3.1}$$

in the class of all conformal metrics satisfying $l(\rho; c) \geqq 1$, *for all dividing cycles* c *on* S_0 *which separate* α_{01} *from* γ_0.

Moreover, for any admissible ρ,

$$A(\rho; S_0) \geqq A(\rho_0; S_0) + A(\rho - \rho_0; S_0) . \tag{3.2}$$

Proof. Clearly the conformal metric ρ_0 satisfies the condition of the lemma, and $A(\rho_0; S_0) = D(u_0; S_0) = \mu_{01} < \infty$. Let ρ be any admissible conformal metric on S_0 with $A(\rho; S_0) < \infty$.

We evaluate the integral

$$\int_{S_0} \rho(z)\rho_0(z) d\sigma_z .$$

Take $w_0 = u_0 + i\bar{u}_0$ for the local parameter on S_0, so that $\rho_0(w_0) \equiv 1$. Denote the level line $u_0(z) = \mu$ $(0 \leqq \mu \leqq \mu_{01}$; see Lemma 1) by $c(\mu)$. From Fubini's theorem,

$$\int_{S_0} \rho(z)\rho_0(z) d\sigma_z = \int_0^{\mu_{01}} d\mu \int_{c(\mu)} \rho(w_0) d\bar{u}_0 .$$

Here the integral $\int_{c(\mu)} \rho(w_0)d\bar{u}_0$ exists almost everywhere, for μ on the closed interval $[0, \mu_{01}]$. But $c(\mu)$ is, for any $\mu \neq \mu_{0i}$, a dividing cycle on S_0 which separates α_{01} from γ_0 and therefore, almost everywhere,

$$\int_{c(\mu)} \rho(w_0)d\bar{u}_0 = \int_{c(\mu)} \rho(z)\,|dz| \geqq \int_{c(\mu)} \rho(z)\,|dz| \geqq 1$$

From the two preceding relations it follows that

$$\int_{S_0} \rho(z)\rho_0(z)d\sigma_z \geqq \mu_{01} .$$

Now put $\rho = \rho_0 + (\rho - \rho_0)$ in $A(\rho; S_0)$; we obtain

$$A(\rho; S_0) = \mu_{01} + A(\rho - \rho_0; S_0) + 2\int_{S_0} \rho\rho_0 d\sigma - 2\mu_{01}$$

and, from the preceding inequality, we conclude finally that

$$A(\rho; S_0) \geqq \mu_{01} + A(\rho - \rho_0; S_0) ,$$

which proves our lemma.

Clearly the admissible conformal metric which minimizes $A(\rho; S_0)$ is unique. For, if $A(\rho; S_0) = A(\rho_0; S_0) = \mu_{01} < \infty$, we deduce from (3.2) that $A(\rho - \rho_0; S_0) = 0$, i.e. $\rho = \rho_0$ almost everywhere on S_0.

Now let γ be a boundary of R, and let S be a given neighborhood of γ. Let $\{\rho\}_\gamma$ denote that class of all measurable conformal metrics defined on S which satisfy the condition

$$l(\rho; c) \geqq 1 , \tag{3.3}$$

for all dividing cycles c which separate γ from γ_0. If $u \in \{u\}_\gamma$, then obviously $|\mathrm{grad} u| \in \{\rho\}_\gamma$. This is valid, in particular, for the conformal metric $\rho_\gamma = |\mathrm{grad} u_\gamma|$. The ρ_γ-area of S is $A(\rho_\gamma; S) = D(u_\gamma; S) = \mu_\gamma$.

THEOREM 4. *In $\{\rho\}_\gamma$ the conformal metric $\rho_\gamma = |\mathrm{grad} u_\gamma|$ gives the minimum of $A(\rho; S)$:*

$$\min A(\rho; S) = A(\rho_\gamma; S) . \tag{3.4}$$

Moreover, for any ρ,

$$A(\rho; S) \geqq A(\rho_\gamma; S) + A(\rho - \rho_\gamma; S) . \tag{3.5}$$

Proof. If $A(\rho; S) = \infty$, (3.5) is evident. Assume now that there exists in $\{\rho\}_\gamma$ a conformal metric ρ which is A-bounded.

Set $|\mathrm{grad} u_n| = \rho_n$ (see 2.2). Since $A(\rho; S) \geqq A(\rho; S_n)$, we conclude from Lemma 5 that

$$A(\rho; S) \geqq \mu_n + A(\rho - \rho_n; S_n)$$

As $n \to \infty$, Fatou's Lemma gives immediately

$$A(\rho; S) \geqq \mu_\gamma + \liminf A(\rho - \rho_n; S_n) \geqq \mu_\gamma + A(\rho - \rho_\gamma; S) ,$$

which proves (3.5) and the theorem.

As in Lemma 5, the uniqueness of the minimizing conformal metric ρ_γ in the case $\mu_\gamma < \infty$ is an immediate consequence of (3.5).

By Theorem 4, the quantity $\lambda_\gamma = \mu_\gamma^{-1}$ is equal to the extremal length of the family of all dividing cycles c on S separating γ from γ_0 ([1], [5]).

3.3. Parabolic boundary components. We return to the condition $\mu_\gamma = \infty$ studied in 2.2.

THEOREM 5. *A boundary component γ of R is parabolic if and only if, for any neighborhood S of γ and for any A-bounded conformal metric ρ on S, there exists a dividing cycle separating γ from γ_0 with an arbitrarily small ρ-length.*

Proof. If $\mu_\gamma < \infty$, the conformal metric ρ_γ is A-bounded and satisfies $l(\rho; c) \geqq 1$, for all dividing cycles separating γ from γ_0. Conversely, if there is an A-bounded conformal metric ρ on S satisfying $l(\rho; c) \geqq \varepsilon > 0$, for all dividing cycles c separating γ from γ_0, the conformal metric $\rho^* = (1/\varepsilon)\rho$ is A-bounded and belongs to $\{\rho\}_\gamma$. Therefore, by Theorem 4, $\mu_\gamma < \infty$.

THEOREM 6. *Suppose R is imbedded in a larger Riemann surface R^*. If a boundary component γ of R or a part of γ realized on R^* contains a continuum γ^*, then γ is hyperbolic.*

Proof. Let $K^* : |z^*| \leqq 1$ denote a parametric disc on R^* for which $K^* \cap \gamma^*$ contains a continuum, say again γ^*. Since γ^* is a boundary continuum of R, there exists a disc $\overline{R}_0 \subset K^* \cap R$. In K^* let $Q = aba'b'$ be a rectangle such that its side a is completely interior to R_0 and its opposite sides b, b' have common points with γ^*.

Set $R - \overline{R}_0 = S$. We define a conformal metric ρ_0 on S by setting $\rho_0(z^*) = 1$ on $Q \cap S$ and $\rho_0 = 0$ otherwise. Clearly ρ_0 is A-bounded and satisfies $l(\rho_0; c) \geqq l_0 > 0$, where l_0 is the length of a in K^* and c is any dividing cycle separating γ from γ_0. Hence, by Theorem 5, γ is not parabolic.

Let S be a given neighborhood of a boundary component γ of R, and let $\{S_n\}$ be an exhaustion of S as in 2.2. Let E_n denote the connected component of $S_n - S_{n-1}$ whose boundary includes γ_{n-1} and γ_n. We assert that

$$(3.6) \qquad \mu(S;\gamma_0,\gamma) \geqq \sum_{n=1}^{\infty} \mu(E_n;\gamma_{n-1},\gamma_n) .$$

In fact, since the restriction of ρ_γ to E_n is admissible in Lemma 5 (where S_0 is replaced by E_n, γ_0 and α_{01} by γ_{n-1} and γ_n respectively), we conclude that $A(\rho_\gamma; E_n) \geqq \mu(E_n;\gamma_{n-1},\gamma_n)$. Therefore, $\mu(S;\gamma_0,\gamma) \geqq \sum_{n=1}^{\infty} A(\rho_\gamma; E_n) \geqq \sum_{n=1}^{\infty} \mu(E_n;\gamma_{n-1},\gamma_n)$, which proves (3.6).

Similarly, it may be proved that

$$(3.7) \qquad \mu(S;\gamma_0,\gamma) \geqq \mu(E_1;\gamma_0,\gamma_1) + \mu(S^*_1;\gamma_1,\gamma) ,$$

where S^*_1 is the connected component of $S - \overline{S}_1$ whose relative boundary is γ_1.

THEOREM 7. *A boundary component γ of R is parabolic if and only if there exists an exhaustion of S for which*

$$(3.8) \qquad \sum_{n=1}^{\infty} \mu(E_n;\gamma_{n-1},\gamma_n) = \infty .$$

Proof. By (3.6), the condition (3.8) is sufficient for the parabolicity of γ.

Conversely, assume that γ is parabolic, and let $\{S_n\}$ be a given exhaustion of S. Since

$$\lim_{n\to\infty} \mu(S_n;\gamma_0,\gamma_n) = \mu(S;\gamma_0,\gamma) = \infty ,$$

we can choose $n_1 \geqq 1$ such that $\mu(S_n;\gamma_0,\gamma_{n_1}) \geqq 1$. Let $S^*_{n_1}$ denote the connected component of $S - \overline{S}_{n_1}$ whose relative boundary is γ_{n_1}. $S^*_{n_1}$ is a neighborhood of γ. Since γ is parabolic, we have

$$\lim_{n\to\infty} \mu(S^*_{n_1,n};\gamma_{n_1},\gamma_n) = \mu(S^*_{n_1};\gamma_{n_1},\gamma) = \infty ,$$

where $S^*_{n_1,n} = S^*_{n_1} \cap S_n$. Therefore, we can choose $n_2 > n_1$ such that $\mu(S^*_{n_1,n_2};\gamma_{n_1},\gamma_{n_2}) \geqq 1$. Continuing this procedure, we obtain an exhaustion $\{S_{n_k}\}$ $(k = 1, 2, \cdots)$ of S, which satisfies condition (3.8). Thus Theorem 7 is established.

3.4. Perimeter and capacity. Let $|z| \leqq r_0$ be a fixed parametric disc on R, and let $S(r)$ denote the complement of the disc $|z| \leqq r$ $(0 < r \leqq r_0)$ with respect to R. Set $\mu(S(r); |z| = r, \gamma) = \mu_{\gamma,r}$. By (3.7), for $r' < r$,

$$\mu_{\gamma,r'} \leqq \frac{1}{2\pi} \log \frac{r}{r'} + \mu_{\gamma,r}$$

or

$$-2\pi\mu_{\gamma,r'} - \log r' \leqq -2\pi\mu_{\gamma,r} - \log r .$$

Therefore,

$$\pi_\gamma = \lim_{r\to 0} \frac{1}{r} e^{-2\pi\mu_{\gamma,r}}$$

exists. According to Ahlfors and Beurling [1], we call π_γ perimeter of γ with respect to the fixed parametric dics $|z| \leqq r_0$. Let $z' = \lambda(z) = az + \cdots, a \neq 0$, be a new local parameter in the neighborhood of the point $P_0 \in R$ corresponding to $z = 0$, and let π'_γ denote the perimeter of γ with respect to the parametric disc $|z'| \leqq r'_0$. Set $|z| = r$ and $|z'| = r'$. For corresponding r and r' by $z' = \lambda(z)$, we have

$$|a|r(1 - \varepsilon_r) \leqq r' \leqq |a|r(1 + \varepsilon_r) ,$$

where ε_r is a positive function of r and $\varepsilon_r \to 0$, as $r \to 0$. It follows, from the conformal invariance and the monotony of modulus, that

$$\pi_\gamma = |a| \pi'_\gamma . \tag{3.9}$$

We now prove the following.

THEOREM 8. *If the perimeter π_γ and the capacity c_r are defined with respect to the same parametric disc $|z| \leqq r_0$, then $\pi_\gamma = c_\gamma$.*

Proof. From (1.6) and (3.9), it is sufficient to prove the required equality for a particular parametric disc of the point P_0. We choose this parametric disc, say again $|z| \geqq r_0$, such that $t_\gamma \equiv \log|z|$ on $|z| \leqq r_0$. Then, by (2.6), we conclude immediately that

$$\pi_\gamma = \lim_{\lambda\to 0} \frac{1}{\lambda} e^{-2\pi\mu_{\gamma,\lambda}} = e^{-k_\gamma} = c_\gamma ,$$

which proves our theorem.

COROLLARY. *If P_γ denote the class of Riemann surfaces defined by $\pi_\gamma = 0$, for all γ, then $P_\gamma = c_\gamma = M_\gamma$.*

§ 4. RIEMANN SURFACES OF FINITE GENUS

4.1. Planar subregions. Let γ be a boundary component of an open Riemann surface R, and suppose that γ is hyperbolic and possesses a neighborhood S which is planar.

Set, as usually, $u(z; S; \gamma_0, \gamma) = u_\gamma$, $\mu(S; \gamma_0, \gamma) = \mu_\gamma$, and consider the function $w = F_\gamma(z)$ defined by

$$F_\gamma(z) = \exp 2\pi(u_\gamma(z) + i\bar{u}_\gamma(z)) \tag{4.1}$$

Consider an exhaustion $\{S_n\}$ of S as in 2.2. Since S is planar, the homology group $H^1(S)$ is generated from the boundary curves α_{ni} of $S_n (n = 1, 2, \cdots)$, and we conclude by (2.3) that F_γ is single-valued. We now prove the following [7]:

THEOREM 9. *The function $w = F_\gamma(z)$ maps the region S univalently onto the annulus*

$$A_{0,\mu_\gamma} : 1 < |w| < e^{2\pi\mu_\gamma}$$

slit along a set of circular arcs around the origin. Here the boundary circumferences $|w| = 1$ and $|w| e^{2\pi\mu_\gamma}$ correspond to γ_0 and γ respectively. The total area of the slits vanishes.

Proof. We define the function $w = F_n(z)$ on S_n by

$$F_n(z) = \exp 2\pi(u_n(z) + i\bar{u}(z)) ,$$

where $u_n = u(z; S_n; \gamma_0, \gamma_n)$. As before, we see that F_n is single-valued, for all n.

The function $w = F_n(z)$ gives a one-to-one conformal mapping of S_n onto the covering surface $S_{n,w} = (S_n, w = F_n(z))$. By the definition of u_n, $|F_n(z)|$ assumes constant values on the boundary curves of S_n and satisfies on S_n:

$$1 < |F_n(z)| < e^{2\pi\mu_n} .$$

It follows that $S_{n,w}$ is an unlimited covering surface of the annulus A_{0,μ_n} slit along a finite number of circular arcs. On the other hand, evaluate the ρ_0-area of $S_{n,w}$, where

$$\rho_0(w) = \frac{1}{2\pi |w|} = \frac{1}{2\pi} \left| \frac{d}{dw} \log w \right| .$$

Since, for $w = F_n(z)$,

$$\rho_n(z) = |\operatorname{grad} u_n(z)| = \frac{1}{2\pi} \left| \frac{d}{dz} \log w \right| = \rho_0(w) \left| \frac{dw}{dz} \right| ,$$

we obtain

$$A(\rho_0; S_{n,w}) = A(\rho_n; S_n) = \mu_n .$$

This is equal to the ρ_0- area of the annulus A_{0,μ_n}. It follows that the covering surface $S_{n,w}$ consists necessarily of a single sheet, that is the function F_n is univalent. Since $F_n \to F_\gamma$ uniformly on each S_N, F_γ is also univalent.

Let us now consider the image $S_w = F_\gamma(z)$. Denote the connected components of the boundary of S_w which correspond to γ_0 and γ by γ_w^0 and γ_w respectively. Clearly γ_w^0 is the circumference $|w| = 1$. Further, since $\mu_n \leqq \mu_\gamma$, for all n, S_w is included in the annulus A_{0,μ_γ}. As before, the ρ_0-area of S_w is

$$A(\rho_0; S_w) = A(\rho_\gamma; S) = \mu_\gamma ,$$

since

$$\rho_\gamma(z) = \rho_0(w)\left|\frac{dw}{dz}\right| \qquad (w = F_\gamma(z)) .$$

This is equal to the ρ_0-area of the annulus A_{0,μ_γ}. Accordingly, the complements of S_w with respect to A_{0,μ_γ} has a (logarithmic and Euclidian) vanishing area.

Assume finally that the set $A_{0,\mu_\gamma} - S_w$ possesses a connected component γ^*_w which is not a point or a circular arc around the origin. Construct two circumferences $|w| = r_i$ $(i = 1, 2; r < r_1 < r_2 < e^{2\pi\mu_\gamma})$ having common points with γ^*_w, and consider a point w_0 in the annulus $r_1 < |w| < r_2$. Let K_ε be the disc $|w - w_0| \leqq \varepsilon$. Obviously, for ε sufficiently small, the conformal metric ρ_ε, defined by $\rho_\varepsilon = 0$ on K_ε and $\rho_\varepsilon(w) = \rho_0(w)$ on $S_w - K_\varepsilon$, satisfies the condition (3.3), for all dividing cycles c on S_w separating γ_w from γ_w^0. This contradicts Theorem 4, since $A(\rho_\varepsilon; S_w) < A(\rho_0; S_w) = \mu_\cdot$. Therefore, the continuum γ^*_w does not exist. In particular, γ_w coincides with $|w| = e^{2\pi\mu_\gamma}$. Theorem 9 is completely proved.

4.2. Planar Riemann surfaces. Suppose now that R itself is planar. Let $|z| \leqq 1$ be a fixed parametric disc on R, γ a hyperbolic boundary component of R, and $c_\gamma > 0$ the capacity of γ with respect to $|z| \leqq 1$. Consider the function $w = T_\gamma(z)$ defined by

$$T_\gamma(z) = c_\gamma \exp(t_\gamma(z) + i\bar{t}_\gamma(z)) .$$

By Lemma 4 and Theorem 9, we have the following [14]:

THEOREM 10. *The function $w = T_\gamma(z)$ is univalent and single-valued on R and maps R onto the unit circle slit along a set of circular arcs of vanishing total area. The boundary component γ is mapped into the unit circumference.*

Let SB (SD) be the class of univalent single-valued analytic functions having a bounded modulus (a finite Dirichlet integral), and let $O_{SB}(O_{SD})$ be the class of Riemann surfaces with no functions belonging to $SB(SD)$.

THEOREM 11. [1, 14] *For planar Riemann surfaces,* (4.2) $$O_{SB} = M_\gamma = O_{SB} .$$

Proof. Assume first that the planar surface R possesses a hyperbolic boundary component γ. Then, the function T_γ of Theorem 10 obviously belongs to the class SB and SD.

Conversely, suppose that there exists on R a function $w = T(z)$ which belongs to the class SB or SD. In both cases, the image $R_w = T(R)$ has a finite Euclidian area. Let $K_\varepsilon : |w - w_0| \leqq \varepsilon$ be a disc which is completely included in R_w. Denote by γ_w the connected component of the boundary of R_w which separates $w = 0$ from $w = \infty$ or contains $w = \infty$, The conformal metric $\rho(w) = 1/2\pi\varepsilon$ is clearly A-boundary on $R_w - K_\varepsilon$ and satisfies condition (3.3), for all dividing cycles on $R_w - K_\varepsilon$ which separate γ_w from $|w - w_0| = \varepsilon$. We conclude that the boundary component γ of R which corresponds to γ_w is hyperbolic.

4.3. Riemann surfaces of finite genus. A *continuation* of a Riemann surface R is defined by (1) another Riemann surface R' and (2) a one-to-one conformal mapping $T: R \to R'$, $T(R) \subset R'$, [**2, 4, 8, 9, 11, 12**]. If R' is a compact Riemann surface, the continuation is called *compact*. If $R' - T(R)$ contains interior points, the continuation is called *essential* [**9, 12**].

Let R be a Riemann surface of finite genus. We say that the continuation of R is *topologically unique* if, for any two compact continuations $T_\nu : R \to R'_\nu (\nu = 1, 2)$ of R, there exists a topological mapping $h^*_{12} = R'_1 \to R'_2$, $h^*_{12}(R'_1) = R'_2$, with $h^*_{12}\ T_1(R) = h_{12}$, where $h_{12} = T_2 T_1^{-1}$. If, in addition, h^*_{12} is always a conformal mapping, the continuation of R is said to be *conformally unique*.

Let O_{AD} denote the class of Riemann surfaces with no non-constant single-valued analytic functions having a finite Dirichlet integral. It is well known that the continuation of a Riemann surface R of finite genus is conformally unique if and only if $R \in O_{AD}$ [**1, 8, 12**]. We now prove the following

THEOREM 12. *For Riemann surfaces of finite genus, the following conditions are equivalent*:

(1) $R \in M_\gamma$

(2) *The continuation of R is topologically unique.*

(3) *R does not possess an essential continuation.*

Proof. (1) $\to$ (2). If $R \in M_\gamma$ and $T_\nu : R \to R'_\nu$ $(\nu = 1, 2)$ are compact continuations of R, then, by Theorem 6, the sets $\beta_\nu = R'_\nu - T_\nu(R)$ are totally disconnected. Set $T_2 T_1^{-1} = h_{12}$. We define a topological mapping h^*_{12} of R'_1 onto R'_2 as follows. First, set $h^*_{12}(P_1) = h_{12}(P_1)$, for any $P_1 \in T_1(R)$. Now let $P_1 \in \beta_1$. Since β_1 is totally disconnected, there is a fundamental sequence $\{U_n\}$ of neighborhoods of P_1 such that the open sets $V_n = U_n \cap T_1(R)$ are connected. Set $E(P_1) = \cap_n \overline{h_{12}(V_n)}$. Clearly this is a closed and connected set. On the other hand, $E(P_1) \subset \beta_2$ and, since β_2 is totally disconnected $E(P_1)$ contains a single point P_2. Set $h^*_{12}(P_1) = P_2$. It is easy to see that h^*_{12} is a topological mapping between R'_1 and R'_2.

(2) $\rightarrow$ (3). If R possesses an essential continuation $T_1: R \rightarrow R'_1$, we may construct in an evident manner another compact continuation $T_2: R \rightarrow R'_2$ of R such that R'_1 and R'_2 have different genera.

(3) $\rightarrow$ (1). Assume that $R \notin M_\gamma$, i.e. R possesses some boundary component γ which is hyperbolic. Let S be a neighborhood of γ. We have $\mu_\gamma < \infty$. By Theorem 9, there is a one-to-one conformal mapping of S into the finite annulus $1 < |w| < e^{2\pi\mu_\gamma}$. Let K_w denote the set $|w| > 1$. Clearly the Riemann surface $R' = (R - S) \cup K_w$ defines an essential continuation of R, and therefore (3) $\rightarrow$ (1). Thus, Theorem 12 is established

COROLLARY 1. *For Riemann surfaces of finite genus, we have* $O_{AD} \subset M_\gamma$.

Note that by a theorem of Ahlfors and Beurling [1] this inclusion is strict.

COROLLARY 2. *Let* $R \in M_\gamma - O_{AD}$ *and of finite genus. Then there exist two compact continuations* $T_\nu: R \rightarrow R'_\nu$ $(\nu = 1, 2)$ *of* R *such that the corresponding topogical mapping* h^*_{12} *is not a conformal mapping.*

In particular, we conclude from Corollary 2 that there exist Pompeiu functions which are univalent (see [3], [10], and [16]).

REFERENCES

1. L. Ahlfors and A. Beurling, *Conformal invariants and function-theoretic null sets*, Acta Math. **83** (1950), 101-129

2. S. Bochner, *Fortsezung Riemannscher Flächen*, Math. Ann., **98** (1928), 406-421.

3. A. Denjoy, *Sur les singularités discontinues des fonctions analytiques uniformes*, C. R. Acad. Sci. Paris, **149** (1909), 386-388.

4. M. Heins, *On the continuation of a Riemann surface*, Ann. of Math., **43** (1942), 208-297.

5. J. Hersch, *Longueurs extrémales et théorie des fonctions*, Comm. Math. Helv., **29** (1954). 301-337.

6. B. Kerékjàrtó, *Verlesungen über Topologie*, Berlin, 1923.

7. H. Grötzsch, *Das Kreisbogenschlitztheorem der konformen Abbildung schlichter Bereiche*, Leipziger Ber., **83** (1931), 238-253,

8. A. Mori, *A remark on the prolongation of Riemann surfaces of finite genus*, J. Math. Soc. Japan, **4** (1952), 27-30.

9. R. Nevanlinna, *Uniformisierung*, Berlin, 1953.

10. D. Pompeiu, *Sur la continuité des fonctions de varlables complexes*, Ann. Fac. Sci. Toulouse, **7** (1905) 265-315.

11. T. Rado, *Über eine nichtfortsetzbare Riemannsche Mannigfaltigkeit,* Math. Z., **21** (1924), 1-6.

12. L. Sario, *Uber Riemannsche Flächen mit hebbarem Rand,* Acad. Sci. Fenn., Ser. AI, **50** (1948).

13. ———, *A linear operator method on arbitrary Riemann surfaces,* Trans. Amer. Math. Soc., **72** (1952), 218-295.

14. ———, *Capacity of the boundary and of a boundary component,* Ann. of Math., **59** (1954), 135-144.

15. N. Savage, *Weak boundary components,* Duke Math. J., **24** (1957), 79-95.

16. S. Stoilow, *Lecons sur les principes topologiques de la théorie des fonctions analytiques,* Paris, 2e éd. 1956.

18. V. Wolontis, *Properties of conformal invariants,* Amer. J. Math., **74** (1952), 587-606.

19. K. Strebel, *Die extremale Distanz zweier Enden einer Riemannschen Fläche,* Ann. Acad. Sci. Fenn, Ser. AI, **179** (1955).

MATHEMATICAL INSTITUTE OF THE R.P.R. ACADEMY, BUCAREST

RUNS

H. KENYON AND A. P. MORSE

1. Introduction. A *relation* is a set of ordered pairs. If R is a relation then it helps our intuition to sometimes think that y comes after x if and only if $(x, y) \in R$. With this in mind we search among relations for directing mechanisms among which are to be not only those familiar ones considered by Moore-Smith, but enough more to handle[1] topological convergence.

We agree that

$$\begin{aligned}\text{dmn } R &= \text{domain } R = \text{E}x\,[(x, y) \in R \text{ for some } y] \\ &= \text{the set of points } x \text{ such that } (x, y) \in R \text{ for some } y,\end{aligned}$$

and that

$$\text{rng } R = \text{range } R = \text{E}y\,[(x, y) \in R \text{ for some } x]\,.$$

Now suppose

$$\Gamma = \text{E}x\,(0 \leqq x < \infty)$$

and

$$\omega = \text{the set of non-negative integers.}$$

Also suppose

$$R_1 = \text{E}x, y(0 \leqq x \leqq y < \infty)$$

and

$$R_2 = \text{E}x, y(x \in \omega \text{ and } 0 \leqq x \leqq y < \infty),$$

so that $(x, y) \in R_1$ if and only if $0 \leqq x \leqq y < \infty$ and $(x, y) \in R_2$ if and only if $x \in \omega$ and $0 \leqq x \leqq y < \infty$.

Clearly

$$\text{rng } R_2 = \text{rng } R_1 = \Gamma$$

but on the other hand

$$\text{dmn } R_2 = \omega \neq \text{dmn } R_1 = \Gamma\,.$$

Nevertheless, R_2 and R_1 are intuitively equivalent directing mechanisms.

Now suppose :

Received February 17, 1958.

[1] See Remark 5.2.

$I = \mathrm{E}t(0 \leqq t \leqq 1)$;
$\Gamma' =$ The set of functions on I to Γ;
$\omega' =$ The set of functions on I to ω;
$R_1' = \mathrm{E}x, y\,[x \in \Gamma'$ and $y \in \Gamma'$ and $x(t) \leqq y(t)$ whenever $t \in I]$;

$R_2' = \mathrm{E}x, y\,[x \in \omega'$ and $y \in \Gamma'$ and $x(t) \leqq y(t)$ whenever $t \in I]$.

Very much as before

$$\operatorname{rng} R_2' = \operatorname{rng} R_1' = \Gamma',$$
$$\operatorname{dmn} R_2' = \omega' \neq \operatorname{dmn} R_1' = \Gamma',$$

but nevertheless R_2' and R_1' are intuitively equivalent directing mechanisms.

Let us look more closely at R_2'. R_2' is clearly transitive. That is, $(x, z) \in R_2'$ whenever (x, y) and (y, z) both belong to R_2'. In other words, if y comes after x and z comes after y, then z comes after x. Moreover if $x \in \operatorname{dmn} R_2'$ and $y \in \operatorname{dmn} R_2'$ then there is a $z \in \operatorname{dmn} R_2'$ which comes after both x and y. That is, corresponding to each $x \in \operatorname{dmn} R_2^1$ and each $y \in \operatorname{dmn} R_2'$ there is a $z \in \operatorname{dmn} R_2'$ for which

$$(x, z) \in R_2' \text{ and } (y, z) \in R_2' .$$

We are thus led to

1.1 DEFINITION. R is a *direction* if and only if R is such a non-vacuous transitive relation that corresponding to each $x \in \operatorname{dmn} R$ and each $y \in \operatorname{dmn} R$ there is a $z \in \operatorname{dmn} R$ for which

$$(x, z) \in R \text{ and } (y, z) \in R .$$

Evidently the directing mechanisms of Moore-Smith are directions, but it turns out that even directions are not topologically adequate.[1]

If R is a direction then clearly for each $x \in \operatorname{dmn} R$ and each $y \in \operatorname{dmn} R$ there is a $z \in \operatorname{dmn} R$ such that anything which comes after z also comes after x and after y. We are now on the right track.

1.2 DEFINITION. R is a *run* if and only if R is such a non-vacuous relation that corresponding to each $x \in \operatorname{dmn} R$ and each $y \in \operatorname{dmn} R$ there is a $z \in \operatorname{dmn} R$ for which

$$(x, t) \in R \text{ and } (y, t) \in R$$

whenever t is such that $(z, t) \in R$.

From 1.1 and 1.2 follows

1.3 THEOREM. *Every direction is a run.*

As we shall indicate, runs are topologically adequate. For that matter, so are the filter-bases of Cartan, the nets of Kelley, and the syntaxes of McShane. But among these we do not find such an old friend as the Moore-Smith direction R_1.

It is a curious fact that one can come across situations in which the effect of a direction cannot be duplicated by a filter-base.[2] Suppose

$$R_3 = \mathrm{E}a, b\,[a \subset b \text{ and } b \text{ is a finite set}] .$$

Clearly, R_3 is a direction. Moreover, it is a direction which has been put to use in defining unordered summation. However, no filter-base can do the work of R_3, since in many set theories the family of all finite supersets of a given finite set is a class incapable of belonging to anything.[3]

The runs which first come to mind are directions. However, some runs are very unlike the directions they generalize. The domain of a run is merely an indexing set of sign-posts which seem to say, "Beyond here is far enough." It may be that many things follow such a sign-post yet no sign-post at all is among them. To savor some possibilities along this line let us examine briefly two more runs.

Assume T topologizes S and that $p \in S$ and check intuitively that

$$\mathrm{E}\beta, x[p \in \beta \in T \text{ and } x \in \beta]$$

is a run which converges to p in the topology T.

Next assume ρ metrizes S and $p \in S$ and check intuitively that

$$\mathrm{E}r, x\,[0 < r < \infty \text{ and } \rho(x, p) \leqq r]$$

is a run which converges to p in the metric ρ.

It must be admitted that filter-bases are less intricate than runs. Moreover, filter-bases handle theoretical limits with less emphasis on inessentials than any other method known to us. What disturbs us and others about filter-bases is that in many specific situations, such as limit by refinement, the filter-bases do not correspond vividly enough to the limiting concept one pictures. Perhaps it is for this reason that directions, though inadequate, are still very much with us. We feel that runs retain the virtues of directions and at the same time remove their inadequacies.

---

[2] A filter-base is a non-empty family of non-empty sets such that the intersection of any two of them includes a third.

[3] In this present paper we have in mind a set theory similar to that employed by J.L. Kelley, General Topology, pp. 250 ff. New York, 1953.

2. Some definitions.

2.1 DEFINITIONS.

1. $\sim B =$ The complement of B
2. $\sigma F = \mathrm{E}x\,(x \in \beta$ for some $\beta \in F)$
3. $\pi F = \mathrm{E}x\,(x \in \beta$ for every $\beta \in F)$

2.2 REMARK. Thus σF is the union and πF the intersection of all members of F. If A is the set whose sole member is x, then $\sigma A = x = \pi A$. We assume the integer 0 and the empty set are the same and notice that $\sigma 0 = 0$ and $\pi 0 =$ the universe.

With vertical and horizontal sections in mind we make the following definitions.

2.3 DEFINITIONS.

1. $\mathrm{vs}\, Rx = \mathrm{E}y\,[(x, y) \in R]$.
2. $\mathrm{hs}\, Ry = \mathrm{E}x[(x, y) \in R]$.

When R is a run then we sometimes think:

$y \in \mathrm{vs}\, Rx$ if and only if y comes after x;
$x \in \mathrm{hs}\, Ry$ if and only if x comes before y.

There is no magical significance, as often in analytic geometry, attached to the letters used. Thus we sometimes think:

$x \in \mathrm{vs}\, R\delta$ if and only if x comes after δ;

or even

$x \in \mathrm{vs}\ Ry$ if and only if x comes after y.

2.4 DEFINITIONS.

1. ${}_*RA = \mathrm{E}y\,[(x, y) \in R$ for some $x \in A]$
2. ${}^*RA = \mathrm{E}x[(x, y) \in R$ for some $y \in A]$

2.5 DEFINITION. $\mathrm{inv}\, R =$ inverse $R = \mathrm{E}x, y\,[(y, x) \in R]$.

2.6 DEFINITION. $R : S = \mathrm{E}x, z\,[$There is a y such that $(x, y) \in S$ and $(y, z) \in R.]$

A function is the same as its graph and is hence a special kind of relation. If f and g are functions, then $f : g$ is that function h such that $h(x) = f(g(x))$ for each x.

2.7 DEFINITION. rct $AB =$ rectangle $AB = \mathrm{E}x, y(x \in A$ and $y \in B)$.

3. A few properties of relations.

3.1 THEOREM. *If R is a relation and f is a function, then*:

1. ${}_*R(A \cup B) = {}_*RA \cup {}_*RB$
2. $A \subset B$ *implies* ${}_*RA \subset {}_*RB$
3. ${}_*R(A \cap B) \subset {}_*RA \cap {}_*RB$
4. ${}^*RA = {}_*\text{inv}\, RA$
5. ${}^*f(A \cap B) = {}^*fA \cap {}^*fB$
6. ${}_*f{}^*fA = A \cap \text{rng}\, f$
7. ${}^*R_*RA \supset A \cap \text{dmn}\, R$.

3.2 THEOREM. *If R and S are relations and f is a function, then*:

1. $\text{vs}\,(R\colon S)x = {}_*R\,\text{vs}\, Sx$ *for each* x
2. $B \cap {}_*SA \subset {}_*S(({}^*SB) \cap A)$
3. $B \cap {}_*fA = {}_*f(({}^*fB) \cap A)$
4. $B \cap {}_*fA \neq 0$ *implies* $({}^*fB) \cap A \neq 0$.

4. Properties of runs.

4.1 THEOREM. *R is a run if and only if R is such a non-vacuous relation that for each x and y in the domain of R there exists a z in the domain of R for which* $\text{vs}\, Rz \subset \text{vs}\, Rx \cap \text{vs}\, Ry$.

Accordingly if some vertical section of R is a set belonging to the universe,[3] then the vertical sections form a filter-base *theoretically* as useful as R itself. Only in the peripheral situation that every vertical section of R is a class incapable of belonging to anything are runs more effective than filter-bases. However, runs do operate on an essentially different and, we feel, more convenient level.

The passage from a filter-base W to a run R can always be successfully accomplished by putting $R = \text{E}\beta, x(x \in \beta \in W)$.

4.2 DEFINITIONS.

1. R *runs* in A if and only if R is a run and $\text{rng}\, R \subset A$.
2. R is *eventually* in A if and only if R is a run and vs $Rx \subset A$ for some $x \in \text{dmn}\, R$.
3. R is *frequently* in A if and only if R is a run and vs $Rx \cap A \neq 0$ for each $x \in \text{dmn}\, R$.

4.3 THEOREMS.

1. *If R runs in A then R is eventually in A.*
2. *If R is eventually in A, then R is frequently in A.*

4.4 DEFINITIONS.

1. S is a *subrun* of R if and only if S is a run, R is a run,

and for each $x \in \operatorname{dmn} R$ there exists such a $y \in \operatorname{dmn} S$ that $\operatorname{vs} Sy \subset \operatorname{vs} Rx$.

2. R *runs the same* as S if and only if R is a subrun of S and S is a subrun of R.

We agree that S is a *corun* of R if and only if there exists such an A that R is frequently in A and $S = R \cap \mathrm{E}x, y(y \in A)$.

If S is a corun of R then S is a subrun of R. However coruns are often inadequate in that S may be a subrun of R and yet no corun of R will run the same as S.

4.5 THEOREMS.

1. *If R is a run, then R is a subrun of R.*
2. *If R'' is a subrun of R' and R' is a subrun of R, then R'' is a subrun of R.*
3. *If R is frequently in A and $S = R \cap \mathrm{E}x, y(y \in A)$, then S is a subrun of R,* $\operatorname{dmn} S = \operatorname{dmn} R$, *and* $\operatorname{vs} Sx = A \cap \operatorname{vs} Rx$ *for each x.*
4. *If R' is a subrun of R and R is eventually in A, then R' is eventually in A.*

4.6 THEOREMS.

1. *If S is a relation and R is frequently in* $\operatorname{dmn} S$, *then $S : R$ is a run,* $\operatorname{dmn}(S : R) = \operatorname{dmn} R$, *and* $\operatorname{vs}(S : R)x = {}_*S \operatorname{vs} Rx$ *for each x.*
2. *If S is a relation and R is a run, then R is frequently in* $\operatorname{dmn} S$ *if and only if* $\operatorname{dmn}(S : R) = \operatorname{dmn} R$.
3. *If S is a relation, R' is a subrun of R, and R' is frequently in* $\operatorname{dmn} S$, *then $S : R'$ is a subrun of $S : R$.*
4. *If S is a relation and R is eventually in* $\operatorname{dmn} S$, *then R is a subrun of* $(\operatorname{inv} S) : (S : R)$.
5. *If f is a function and R is eventually in* $\operatorname{rng} f$, *then* $f : (\operatorname{inv} f) : R$ *runs the same as R.*

4.7 DEFINITIONS.

1. merger $RS =$ the set of points of the form $((x, y), z)$, where x, y and z are such that R and S are runs, $(x, z) \in R$, and $(y, z) \in S$.
2. R *merges* with S if and only if R and S are such runs that $\operatorname{vs} Rx \cap \operatorname{vs} Sy \neq 0$ whenever $x \in \operatorname{dmn} R$ and $y \in \operatorname{dmn} S$.

4.8 THEOREMS.

1. *If R and S are runs and $V =$* merger RS, *then V is a relation,* $\operatorname{dmn} V \subset \operatorname{rct} \operatorname{dmn} R \operatorname{dmn} S$, *and* $\operatorname{vs} V(x, y) = \operatorname{vs} Rx \cap \operatorname{vs} Sy$ *whenever* $(x, y) \in \operatorname{dmn} V$.
2. *If R merges with S and $V =$* merger RS, *then dmn* $V = \operatorname{rct} \operatorname{dmn} R \operatorname{dmn} S$, *and V is a subrun of both R and S.*

3. *If W is a subrun of both R and S, then R and S merge and W is a subrun of* merger RS.

4.9 THEOREM. *If f is a function, R is frequently in* dmn f, *and W is a subrun of $f:R$, then there exists such a subrun V of R that W runs the same as $f:V$.*

Proof. *Let* $V' = (\operatorname{inv} f) : W$, *let* $V =$ merger $V'R$, *and let* $W' = f : V$.

Use 4.6.1 to see that

$$\operatorname{dmn}(f:R) = \operatorname{dmn} R \quad \text{and} \quad \operatorname{dmn} V' = \operatorname{dmn} W.$$

We complete the proof in three parts by showing that W' and W are subruns of each other.

Part 1. *V' merges with R, V is a subrun of V' and R, W' is a run, and* $\operatorname{dmn} W' = \operatorname{dmn} V = \operatorname{rct} \operatorname{dmn} W \operatorname{dmn} R$.

Proof. Let $x \in \operatorname{dmn} W$ and $y \in \operatorname{dmn} R$. Since W is a subrun of $f:R$ we have

$$0 \neq \operatorname{vs} Wx \cap \operatorname{vs}(f:R)y = \operatorname{vs} Wx \cap {}_*f \operatorname{vs} Ry.$$

Hence, using 3.2.4, 3.2.1, and 3.1.4, we find that

$$0 \neq ({}^*f \operatorname{vs} Wx) \cap \operatorname{vs} Ry = \operatorname{vs} V'x \cap \operatorname{vs} Ry.$$

Use of 4.6.1 and 4.8.2 completes the proof.

Part 2. *W' is a subrun of W.*

Proof. Use Part 1, 4.6.3, and 4.6.5 to see that $W' = f : V$ is a subrun of $f : V' = f : (\operatorname{inv} f) : W$, which runs the same as W.

Part 3. *W is a subrun of W'.*

Proof. Let $x \in \operatorname{dmn} W$ and $y \in \operatorname{dmn} R$. Select $x' \in \operatorname{dmn} W$ so that $\operatorname{vs} Wx' \subset \operatorname{vs}(f:R)y$, and select $x'' \in \operatorname{dmn} W$ so that $\operatorname{vs} Wx'' \subset \operatorname{vs} Wx \cap \operatorname{vs} Wx'$.
Then

$$\operatorname{vs} Wx'' \subset \operatorname{vs} Wx \cap \operatorname{vs}(f:R)y = \operatorname{vs} Wx \cap {}_*f \operatorname{vs} Ry,$$

which in accordance with 3.2.3 equals

$$\begin{aligned} {}_*f(({}^*f \operatorname{vs} Wx) \cap \operatorname{vs} Ry) &= {}_*f [\operatorname{vs}((\operatorname{inv} f) : W)x \cap \operatorname{vs} Ry] \\ &= {}_*f \operatorname{vs} V(x, y) = \operatorname{vs} W'(x, y). \end{aligned}$$

In view of Part 1 the proof is complete.

4.10. REMARK. Theorems 4.6.1, 4.6.3, and 4.9 show us that under any properly chosen function f, a run R is mapped into a run $S = f : R$, subruns of R are mapped into subruns of S, and any subrun of S runs the same as the map of some subrun of R.

4.11 DEFINITION. indexrun $R = \mathrm{E}x, y[R$ is a run, $x \in \operatorname{dmn} R$, $y \in \operatorname{dmn} R$, and $\operatorname{vs} Ry \subset \operatorname{vs} Rx]$.

4.12 THEOREM. *If R is a run and $D =$ indexrun R, then D is a direction,* $\operatorname{dmn} D = \operatorname{rng} D = \operatorname{dmn} R$, $(x, x) \in D$ *whenever* $x \in \operatorname{dmn} D$, *and* $R = R : D$.

4.13 REMARK. According to Theorem 4.12, every run is the composition of a relation with a reflexive direction. In fact, every run runs the same as the composition of a *function* with a reflexive direction. Suppose R is a run and D is the set of pairs of the form $((x, y), (x', y'))$, where x, y, x', and y' are such that $(x, y) \in R$, $(x', y') \in R$, and $\operatorname{vs} Rx' \subset \operatorname{vs} Rx$. Let f be such a function that $f(x, y) = y$ whenever $(x, y) \in \operatorname{dmn} D$. It is easy to check that D is a reflexive direction and that R runs the same as $f : D$.

In this connection it should be remarked that if (f, D) is a net in the sense of Kelley (op. cit.) then $f : D$ is a corresponding run. The above construction gives a method for passing from a run back to a corresponding net.

4.14 DEFINITIONS.

1. R is a *full* run if and only if R is a run which runs the same as all of its subruns.
2. R is *fillable* if and only if there exists a full subrun of R.

4.15 THEOREMS.

1. *If R is a full run and R is frequently in A, then R is eventually in A.*

Proof. Note that R is a subrun of $R \cap Ex, y(y \in A)$.

2. *R is a full run if and only if for every A, R is either eventually in A or eventually in $\sim A$.*
3. *If R is a full run, f is a function, and R is frequently in* $\operatorname{dmn} f$, *then $f : R$ is a full run.*

Proof. Use 2.

4. *If S is a full run which merges with R, then S is a subrun of R.*

4.16 DEFINITIONS.

1. N is R *nested* if and only if R is a relation and either or $y = x(x, y) \in R \cup \operatorname{inv} R$ whenever x and y are in N.
2. N is *nested* if and only if either $\alpha \subset \beta$ or $\beta \subset \alpha$ whenever α and β are in N.

4.17 DEFINITIONS.

1. F is R *capped* if and only if R is a relation and corresponding to each R nested subfamily N of F there is a $z \in F$ such that $(x, z) \in R$ whenever $x \in N$.
2. F is *capped* if and only if corresponding to each nested subfamily N of F there is $\gamma \in F$ such that $\sigma N \subset \gamma$.

We have found quite useful the following inductive variants of Zorn's

4.18 LEMMAS.

1. *If R is transitive and F is R capped, and if corresponding to each $x \in F \sim K$ there is a $y \in F$ for which $(x, y) \in R \sim \operatorname{inv} R$ then $F \cap K \neq 0$.*
2. *If F is capped and if each member of $F \sim K$ is a proper subset of some member of F, then $F \cap K \neq 0$.*

4.19 REMARK. In accordance with the terminology used by Kelley (op. cit.), we agree that a *set* is a class which is small enough to belong to the universe.

4.20 THEOREMS.

1. *If R is a full run, then R is eventually in some set.*

Outline of proof. Otherwise according to 4.15.2 R is eventually in $\sim A$ whenever A is a set. Advantage may be taken of this fact to construct by transfinite induction two classes B and C for which $B \cap C = 0$, R is frequently in B, and R is frequently in C. In view of 4.15.1 this is impossible.

2. *R is fillable if and only if R is frequently in some set.*

Proof. If R is fillable it is easy to check with the help of 1. that R is frequently in some set. We now assume that R is frequently in some set A and show that R is fillable.

We agree that sng x is the family whose sole member is x, and that $G \cap \cap H = \mathrm{E}\gamma\,[\gamma = \alpha \cap \beta$ for some $\alpha \in G$ and $\beta \in H]$.

Let $B = \mathrm{E}\alpha\,[\alpha \subset A$ and R is frequently in $\alpha]$, let $F = \mathrm{E}W[W$ is a filter-base and $W \subset B]$, and let $K = \mathrm{E}W\,[$for each $\alpha \subset A$ there exists a $\beta \in W$ for which either $\beta \subset \alpha$ or $\beta \subset A \sim \alpha$. If N is a nested subfamily of F then: if $N = 0$ then $\sigma N = 0 \subset \operatorname{sng} A \in F$; if $N \neq 0$, then $\sigma N \subset \sigma N \in F$. Accordingly F is capped.

Now suppose $W \in F \sim K$, and select such a set α that $\beta \cap \alpha \ni B$ and $\beta \sim \alpha \neq 0$ whenever $\beta \in W$. Let $W' = W \cup (W \cap \cap \text{sng } \alpha)$ and check that $W' \in F$ and that W is a proper subfamily of W'. According to 4.18.2 we conclude that $F \cap K \neq 0$ and select $V \in F \cap K$, so that V is a filter base, R is frequently in every member of V, and for each $\alpha \subset A$ there exists such a $\beta \in V$ that $\beta \subset \alpha$ or $\beta \subset A \sim \alpha$.

Let $S = \mathrm{E}\beta, x[x \in \beta \in V]$ and notice that S is a full run which merges with R. According to 4.15.4, S is a full subrun of R. This completes the proof.

4.19 REMARK. The run R_3 is not fillable.

5. Topological convergence.

5.1 DEFINITIONS.

1. R *clusters* about p in the topology T if and only if T is a topology, $p \in \sigma T$, and R is frequently in every T neighborhood of p.
2. R *converges* to p in the topology T if and only if T is a topology, $p \in \sigma T$, and R is eventually in every T neighborhood of p.
3. R *converges* in the topology T if and only if there exists such a point p that R converges to p in the topology T.
4. nhbdrun pT = the neighborhood run of p in the topology $T = \mathrm{E}\beta, x[T$ is a topology, $p \in \beta \in T$, and $x \in \beta]$.
5. nhbdrun$'$ $pT = \mathrm{E}\beta, x[T$ is a topology, $p \in \beta \in T, x \in \beta$, and $p \neq x]$.

5.2 REMARK. If T is a topology, $A \subset \sigma T$, and p is a point in the T closure of A, then $\mathrm{E}\beta, x(p \in \beta \in T$ and $x \in \beta \cap A)$ runs in A and converges to p in the topology T. It is possible that no run which runs in A and converges to p in the topology T can also be a direction. This can be seen by making use of the topology defined in Problem E on page 77 of Kelley (op. cit.).

5.3 THEOREMS.

1. *R clusters about p in the topology T if and only if R merges with* nhbdrun pT.
2. *R converges to p in the topology T if and only if R is a subrun of* nhbdrun pT.

As an application of the foregoing we offer the following characterizations of compactness.

5.4 THEOREM. *Each of the following is a necessary and sufficient condition that a topology T be compact.*

1. *Whenever R runs in σT, then for some point p, R clusters about p in the topology T.*
2. *Whenever R runs in σT, then there exists such a subrun R' of R that R' converges in the topology T.*
3. *Whenever R is a full run which runs in σT, then R converges in the topology T.*

5.5 REMARK. The Tychonoff theorem,[4] which assures us that the topological product of compact topologies is compact, we will now prove following a well-known pattern. Suppose T is the product topology[5] in question and that R is a full run which runs in σT. Considering any coordinate, let P be the usual projection which maps σT into the corresponding coordinate space. According to 4.15.3 and 5.4.3, $P:R$ is a full run which converges in the topology of the coordinate space. Consequently R converges coordinatewise and hence converges in the topology[5] T.

6. Limits.

6.1 DEFINITIONS.

1. far RxP if and only if R is eventually in ExP.

In 1. above we allow "P" to be replaced by an arbitrary formula such as, for example,

$$\text{“}[y < x < x^2]\text{”}.$$

2. $f(x)$ tends to p in the topology T as x runs along R if and only if T is a topology, $p \in \sigma T$, and far $Rx(f(x) \in \beta)$ whenever β is a T neighborhood of p.
3. $f(x)$ tends uniquely to p in the topology T as x runs along R if and only if for every q ($p = q$ if and only if $f(x)$ tends to q in the topology T as x runs along R).
4. lmt $TxRf(x)$ = the limit in the topology T as x runs along R of $f(x) = \pi$Ep [$f(x)$ tends uniquely to p in the topology T as x runs along R].[6]

Thus if $f(x)$ tends uniquely to p in the topology T as x runs along R we know that lmt $TxRf(x) = p$.

6.2 THEOREM. *If T is a topology, $p \in \sigma T$, f is a function, and R is eventually in the domain of f, then*

1. *$f(x)$ tends to p in the topology T as x runs along R if and only if $f:R$ converges to p in the topology T; and*
2. lmt $TxRf(x) = p$ *if and only if $f:R$ converges to p in the*

---

[4] See Kelley (op. cit.) p. 143.
[5] See Kelley (op. cit.) pp. 88–92.
[6] See Remark 2.2.

topology T, *and* $q = p$ *for every* q *such that* $f{:}R$ *converges to* q *in the topology* T.

Very elementary but of considerable use is the

6.3 THEOREM. *If* far $Rx(u(x) = v(x))$ *then* lmt $TxRu(x) =$ lmt $T\,x\,Rv(x)$.

6.4 REMARK. As examples of specialized limit notations in which either the run or the topology or both are suppressed, we give the following definitions. We agree that $\mathscr{T}$ is the usual topology for the extended real number system, and that

$$\mathscr{R} = \text{E}m, n[m \in \omega \text{ and } m \leqq n \in \omega].$$

6.5 DEFINITIONS.

1. lnt $T\,n\,u(n) =$ lmt $\mathscr{T}\,n\,\mathscr{R}u(n)$
2. lm $x\,R\,f(x) =$ lmt $\mathscr{T}\,x\,R\,f(x)$
3. $\overline{\text{lm}}\,x\,R\,f(x) =$ lmt $\mathscr{T}\,t$ indexrun R (sup $x \in$ (vs Rt) $f(x)$)
4. $\underline{\text{lm}}\,x\,R\,f(x) =$ lmt $\mathscr{T}\,t$ indexrun R (inf $x \in$ (vs Rt) $f(x)$)
5. lim x a $f(x) =$ lm x(nhbdrun′ $a\mathscr{T}$) $f(x)$
6. lin $n\,u(n) =$ lnt $\mathscr{T}\,nu(n)$

6.6 REMARK. In 6.5.1 we have a limit notation for ordinary sequences. If u is a sequence, T is a topology, and $p \in \sigma T$, then lnt $T\,n\,u(n) = p$ if and only if p is the unique point such that $u : \mathscr{R}$ converges to p in the topology T.

We give a few more simple but useful theorems.

6.7 THEOREMS.

1. *If* $\delta \in$ dmn R *and if far* $Rx\,\{f(x \geqq f(y)\}$ *whenever* $y \in$ vs $R\delta$,

   then $-\infty \leqq \text{lm}\,x\,R\,f(x) = \sup x \in \text{vs}\,R\delta\ f(x) \leqq \infty$.

2. *If* $\delta \in$ dmn R *and if far* $Rx\{f(x) \leqq f(y)\}$ *whenever* $y \in$ vs $R\delta$,

   then $-\infty \leqq \text{lm}\,x\,R\,f(x) = \inf x \in \text{vs}\,R\delta\ f(x) \leqq \infty$.

From 1. and 2. we infer 3. and 4. below. These results are generalizations of the fact that non-decreasing and non-increasing functions have limits.

3. *If* far Ry far $Rx\{f(x) \geqq f(y)\}$ then

$$-\infty \leqq \text{lm}\,x\,Rf(x) \leqq \infty.$$

4. *If* far Ry far $Rx\{f(x) \leqq f(y)\}$ then

$$-\infty \leqq \text{lm}\,x\,Rf(x) \leqq \infty.$$

6.8 THEOREM. *If R is a run and $-\infty \leqq a \leqq \infty$, then*

$$\operatorname{lm} x\, R\, a = a\,.$$

6.9 THEOREM. *If $A = \operatorname{lm} x\, R\, u(x)$ and $B = \operatorname{lm} x\, R\, v(x)$, then:*
1. *if $-\infty \leqq A + B \leqq \infty$, then $\operatorname{lm} x\, R\{u(x) + v(x)\} = A + B$;*
2. *if $-\infty \leqq A \cdot B \leqq \infty$, then $\operatorname{lm} x\, R\{u(x) \cdot v(x)\} = A \cdot B$.*

In connection with 6.9 above and 6.10 below it is understood that $\infty - \infty$, $0 \cdot \infty$, and $1/0$ are not real numbers.

6.10 THEOREM. *If $A = \operatorname{lm} x\, R\, u(x)$ and $-\infty \leqq 1/A \leqq \infty$, then*

$$\operatorname{lm} x\, R\{1/u(x)\} = 1/A\,.$$

From 4.15.3 and 5.4.3 we infer

6.11 THEOREM. *If R' is a full subrun of R and* far

$$R\, x\, (-\infty \leqq u(x) \leqq \infty), \text{ then}:$$

$$-\infty \leqq \underline{\operatorname{lm}}\, x\, R\, u(x) \leqq \operatorname{lm} x\, R'\, u(x) \leqq \overline{\operatorname{lm}}\, x\, R\, u(x) \leqq \infty\,.$$

If R is fillable, then Theorem 6.11 furnishes us with a generalized limit which, since it is expressed as an actual limit, automatically enjoys the properties found in Theorems 6.8, 6.9, and 6.10. In the event u is bounded, it does not at first glance seem too unreasonable to hope that a similar generalized limit could be arrived at by some Hahn-Banach technique. We, however, are inclined to think this impossible.[7]

We close with an application of limits to integration which expresses the Lebesgue integral as a genuine limit of Riemann-like sums.

6.12 REMARK. Suppose that $\mathscr{L}$ is Lebesgue measure and $\mathfrak{P} = \mathrm{E}P\,[P$ is a countable disjointed family of non-empty $\mathscr{L}$ measurable subsets of the unit interval for which $\sigma P =$ the unit interval]. We agree that Q is a *refinement* of P if and only if every member of Q is included in some member of P, and agree that ξ is a selector function if and only if $\xi(\beta) \in \beta$ whenever $\beta \in \operatorname{dmn} \xi$. Let

$R_4 = \mathrm{E}P, \xi[P \in \mathfrak{P}$ and ξ is a selector function whose domain is a member of $\mathfrak{P}$ and a refinement of $P]$.

We now have the

THEOREM. $\displaystyle\int_0^1 f(x)\,dx = \operatorname{lm} \xi\, R_4 \sum \beta \in \operatorname{dmn} \xi \{f(\xi(\beta)) \cdot \mathscr{L}(\beta)\}$

[7] See R. P. Agnew and A. P. Morse, *Extensions of linear functionals with applications to limits, integrals, measures, and densities*, Ann. Math. Stat. **39**, no. 1, January. 1938. Notice especially the first two lines on page 24.

whenever f is a finite-value $\mathscr{L}$ measurable function defined on the unit interval.

We think it noteworthy that R_4 runs in the selector functions.

UNIVERSITY OF ROCHESTER
UNIVERSITY OF CALIFORNIA

TWO NON-SEPARABLE COMPLETE METRIC SPACES DEFINED ON [0, 1]

BURNETT MEYER AND H D SPRINKLE

Let $\mathfrak{M}$ be the set of all Lebesgue measurable subsets of the closed interval [0, 1], and let $A, B \in \mathfrak{M}$. It is well-known that $\mathfrak{M}$ becomes a pseudo-metric space if distance is defined by

$$d(A, B) = m(A - B) + m(B - A) = m[(A - B) \cup (B - A)] ,$$

m denoting the Lebesgue measure. See [1, pp. 31–32]. It is the purpose of this paper to extend $\mathfrak{M}$ to include the non-measurable sets and to examine some of the properties of the resulting space.

If we remove the restriction that A and B be measurable, and let them be any subsets of [0, 1], then if

$$\rho(A, B) = m^*(A - B) + m^*(B - A), \text{ and } \delta(A, B) = m^*[A - B) \cup (B - A)]$$

(where m^* denotes the exterior Lebesgue measure), it is easily seen that pseudo-metric spaces $\mathfrak{S}$ and $\mathfrak{T}$ are obtained, corresponding to ρ and δ respectively. The properties which we discuss of $\mathfrak{S}$ and $\mathfrak{T}$ are the same and are proved analogously, so we shall state and prove our results for the space $\mathfrak{S}$ only, it being understood that similar theorems and proofs hold for $\mathfrak{T}$.

LEMMA 1. *A necessary and sufficient condition that $\rho(A, B) = 0$ is the existence of sets Z_1 and Z_2, both of Lebesgue measure zero, such that $A \cup Z_1 = B \cup Z_2$.*

Necessity. If $\rho(A, B) = 0$, then $m(A - B) = m(B - A) = 0$. Since $A \cup (B - A) = A \cup B = B \cup A = B \cup (A - B)$, Z_1 and Z_2 may be taken as $B - A$ and $A - B$, respectively.

Sufficiency. If $A \cup Z_1 = B \cup Z_2$, then

$$\rho(A, B) \leqq \rho(A, A \cup Z_1) + \rho(A \cup Z_1, B \cup Z_2) + \rho(B \cup Z_2, B) = 0$$

The relation $\rho(A, B) = 0$ is seen to be an equivalence relation defined on the elements of $\mathfrak{S}$; hence, those elements are partitioned into equivalence classes. Let $[A]$ denote the equivalence class which contains A. It is clear that if $C \in [A]$ and $D \in [B]$, then $\rho(A, B) = \rho(C, D)$. If $\mathfrak{S}^*$ is the set of all equivalence classes defined above, and if $\rho([A], [B]) = \rho(A, B)$, then $\mathfrak{S}^*$ becomes a metric space with the metric $\rho([A], [B])$.

Received June 11, 1958.

LEMMA 2. *If* $B_n \in [A_n]$ *for* $n = 1, 2, \cdots$, *then* $[\bigcup_{n=1}^{\infty} A_n] = [\bigcup_{n=1}^{\infty} B_n]$ and $[\bigcap_{n=1}^{\infty} A_n] = \bigcap_{n=1}^{\infty} B_n]$.

LEMMA 3. *If* A *is measurable and* $B \in [A]$, *then* B *is measurable.*

There exist Z_1 and Z_2 such that $A \cup Z_1 = B \cup Z_2$ with $m(Z_1) = m(Z_2) = 0$. Let $\tilde{B}$ denote $[0, 1] - B$. Then $B \cup Z_2$ is measurable and since $B = (B \cup Z_2) - (\tilde{B} \cap Z_2)$, B is measurable.

It follows from Lemma 3 that the sets in each equivalence class are either all measurable or all non-measurable. Thus the space $\mathfrak{S}^* = \mathfrak{M}^* \cup \mathfrak{N}^*$, where $\mathfrak{M}^*$ is the space of all equivalence classes of measurable sets, and $\mathfrak{N}^*$ is the space of all equivalence classes of non-measurable sets. It should be noted that $\mathfrak{M}^*$ is the metric space corresponding to the well-known pseudo-metric space $\mathfrak{M}$ defined at the beginning of the paper.

In the following we will omit the asterisks and square brackets, and will write $\mathfrak{S}$ for $\mathfrak{S}^*$, etc., and $\rho(A, B)$ for $\rho([A], [B])$. When we write $A \in \mathfrak{S}$, A may be considered either as an equivalence class or as a representative element of that class.

THEOREM 1. *The space* $\mathfrak{S}$ *is complete.*

The proof is similar to that given in [1, p. 32].

THEOREM 2. *For every* $A \in \mathfrak{S}$ *and every positive number* $\varepsilon < 1$, *there exists* $B \in \mathfrak{S}$ such that $0 < \rho(A, B) < \varepsilon$.

Proof Case I. $m(A) = 0$.

If $m(A) = 0$, then $A \in [\phi]$, ϕ denoting the empty set. Let $B \in \mathfrak{S}$ be an interval of length $< \varepsilon$. Then $\rho(A, B) = \rho(\phi, B) = m(B) < \varepsilon$.

Case II. $m^*(A) > 0$.

Let $I \in \mathfrak{S}$ be an interval of length $< \varepsilon$, such that $m^*(I \cap A) > 0$. If $B = A - I$, then

$$\rho(A, B) = \rho(A, A - I) = m^*[A - (A - I)] = m^*(I \cap A) \leqq m^*(I) < \varepsilon .$$

COROLLARY 1. *If in Theorem* 2, $A \in \mathfrak{M}$, *then* B (*as constructed*) $\in \mathfrak{M}$.

THEOREM 3. *If* $A \in \mathfrak{M}$ *and* $\varepsilon > 0$, *then there exists* $C \in \mathfrak{N}$ *such that* $0 < \rho(A, C) < \varepsilon$.

Proof Case I. $m(A) = 0$.

Let M be a set of real numbers such that for every measurable set E, $m^*(M \cap E) = m(E)$ and $m_*(M \cap E) = 0$, m_* denoting the interior Lebesgue measure. (See [2], Theorem E, p. 70.) In Case I of Theorem 2, let $C = B \cap M$. Then

$$\rho(A, C) = \rho(\phi, C) = m^*(C) = m(B) < \varepsilon \text{ and } m_*(C) = 0 .$$

Case II. $m(A) > 0$.

In Case II of Theorem 2, let $C = A - (I \cap M)$, M described above. Then $\rho(A, C) = m^*(A - C) = m^*(A \cap I \cap M) \leqq m(I) < \varepsilon$, and $m^*(A \cap I \cap M) = m(A \cap I) > 0$, $m_*(A \cap I \cap M) = 0$. Since $(A \cap I \cap M) \in \mathfrak{N}$, $C \in \mathfrak{N}$.

THEOREM 4. *$\mathfrak{N}$ is open in $\mathfrak{S}$.*

Proof. Assume Theorem 4 is false. Then there exists $N \in \mathfrak{N}$ and sets $M_m \in \mathfrak{M}$, $m = 1, 2, \cdots$, such that $\lim_{m\to\infty} \rho(N, M_m) = 0$. The sequence M_m, $m = 1, 2, \cdots$, is, therefore, a Cauchy sequence in $\mathfrak{S}$ and so by Theorem 1 has a subsequence M_{m_n}, $n = 1, 2, \cdots$, such that $\lim_{m\to\infty} \rho(\lim \sup_n M_{m_n}, M_m) = 0$. Since $\lim \sup_n M_{m_n}$ is measurable, this means that N is measurable by Lemma 3, a contradiction.

The last few results can be summarized as follows.

THEOREM 5. *$\mathfrak{M}$ is perfect and nowhere dense in $\mathfrak{S}$; $\mathfrak{N}$ is open and dense in $\mathfrak{S}$.*

The remainder of the work is valid for both spaces, as only the equivalence classes are dealt with (these being the same for $\mathfrak{S}$ and $\mathfrak{T}$).

After having proven completeness for $\mathfrak{S}$ in Theorem 1, a natural question to ask is "Is the space separable ?". The theorem proved here which demonstrates the existence of $2^{\mathfrak{c}}(= \mathfrak{f})$, where $2^{\aleph_0} = \mathfrak{c}$, equivalence classes in $\mathfrak{S}$ answers this question (and a similar one about a countable basis) in the negative. It is also interesting to note that the space $\mathfrak{M}$ has exactly $\mathfrak{c}$ equivalence classes. (In the following work Ω is the first ordinal belonging to $\mathfrak{c}$.)

THEOREM 6. *There exist $\mathfrak{f}$ equivalence classes in the space $\mathfrak{S}$.*

Proof. It will be sufficient to construct a well-ordered family $\{A_\alpha \mid 0 \leqq \alpha < \Omega\}$ of mutually disjoint subsets of $[0, 1]$, each of which has $m^*(A_\alpha) = 1$.

Consider $\{B_\beta \mid 0 \leqq \beta < \Omega\}$ as a well-ordering of all closed subsets B_β of $[0, 1]$ which have a positive Lebesgue measure. For each β, $0 \leqq \beta < \Omega$, let $\{x_\alpha^\beta \mid 0 \leqq \alpha \leqq \beta\}$ be a well-ordered subset of B_β such that $x_\alpha^\beta \neq x_{\alpha'}^{\beta'}$, if $\beta \neq \beta'$ or $\alpha \neq \alpha'$. This selection is possible since, for each β, the set of all $x_{\alpha'}^{\beta'}$ with $0 \leqq \alpha' \leqq \beta' < \beta$ has a cardinal number $< \mathfrak{c}$. Set $A_\alpha = \{x_\alpha^\beta \mid \alpha \leqq \beta < \Omega\}$, for each α, $0 \leqq \alpha < \Omega$. By a simple argument $A_\alpha \cap A_{\alpha'} = \phi$, for $\alpha \neq \alpha'$. Now consider any A_α; if $m^*(A_\alpha) \neq 1$, then A_α is contained in some open set Y such that $m(Y) < 1$. The complement of Y is closed and has $m([0, 1] - Y) > 0$. But $m\{[0, x] \cap ([0, 1] - Y)\}$

is a continuous function of x for $0 \leqq x \leqq 1$; therefore, this function takes on all values between 0 and $m([0, 1] - Y)$, inclusive. This means that there are non-denumerably many closed sets whose measures are greater than 0 and which do not intersect A_α. This is, of course, impossible by the construction of A_α. Therefore, $m^*(A_\alpha) = 1$.

Form the set of all subsets of the set of A_x's, and take the sum of each element of this power set. Any two such sums belong to two different equivalence classes since they disagree in a set of exterior measure 1. This set of sums has cardinal $\mathfrak{f}$. There are, therefore, at least $\mathfrak{f}$ equivalence classes, at most $\mathfrak{f}$ such classes; hence, exactly $\mathfrak{f}$.

References

1. H. Hahn and A. Rosenthal, *Set functions*, University of New Mexico Press, 1948.
2. P. R. Halmos, *Measure theory*, van Nostrand, 1950.

University of Colorado
University of Arizona

CESÀRO PARTIAL SUMS OF HARMONIC SERIES EXPANSIONS

M. S. ROBERTSON

1. Introduction. Let the harmonic function $v(r, \theta)$ have the sine series expansion

$$v(r, \theta) = \sum_{1}^{\infty} a_\nu r^\nu \sin \nu\theta \,, \tag{1.1}$$

convergent for $0 \leqq r < 1$ and suppose that $v(r, \theta)$ is non-negative for $0 < \theta < \pi$. Denote the nth partial sum of (1.1) by

$$S_n^{(0)}(r, \theta) = \sum_{1}^{n} a_\nu r^\nu \sin \nu\theta \,, \tag{1.2}$$

and the nth Cesàro partial sum of order k, by

$$S_n^{(k)}(r, \theta) = \sum_{1}^{n} C_k^{n+k-\nu} a_\nu r^\nu \sin \nu\theta \,, \qquad k = 1, 2, \cdots . \tag{1.3}$$

It was shown by Fejér [2, p. 61] and Szász [8] that when $v(r, \theta) \geqq 0$ for $0 < \theta < \pi$, $0 < r < 1$, then $S_n^{(0)}(r, \theta)$ is also non-negative for all n when $0 < \theta < \pi$, $0 < r \leqq 1/4$, and the constant $1/4$ is sharp. Fejér [2] showed that the functions $S_n^{(3)}(1, \theta)$ are also non-negative for all n, $0 < \theta < \pi$. In addition, Szász [8] showed that there exists an $R_n^{(0)}$, depending upon n only, so that $S_n^{(0)}(r, \theta) \geqq 0$ for $0 < r \leqq R_n^{(0)}$, $0 \leqq \theta \leqq \pi$, but not always for $r > R_n^{(0)}$, and that

$$R_n^{(0)} = 1 - 3\frac{\log n}{n} + \frac{\log\log n}{n} + O(1/n) \,. \tag{1.4}$$

In this paper we shall extend the results of Szász to Cesàro partial sums of integral order k, $k = 1, 2, 3$. For $k = 3$ the theorem obtained is a sharpened form of the theorem of Fejér [2]. We prove the following:

THEOREM 1. *Let the harmonic series expansion*

$$v(r, \theta) = \sum_{1}^{\infty} a_\nu r^\nu \sin \nu\theta$$

be convergent for $0 \leqq r < 1$ *and let* $v(r, \theta) \geqq 0$ *for* $0 < \theta < \pi$, $0 < r < 1$. *Then for* $k = 0, 1, 2, 3$ *there exists a positive number* $R_n^{(k)}$ *depending upon* n *only, so that*

Received by the editors February 27, 1958, and in revised form June 3, 1958. The author wishes to express his appreciation to the referee for helpful suggestions.

(1.5) $$S_n^{(k)}(r,\theta) \geqq 0 \quad \text{for} \quad 0 \leqq r \leqq R_n^{(k)}, \quad 0 \leqq \theta \leqq \pi ,$$

but not always for $r > R_n^{(k)}$, *and that*

(1.6) $$R_n^{(k)} = 1 - (3-k)\frac{\log n}{n} + \frac{\log\log n}{n} - \frac{g_k + o(1)}{n}, \qquad k = 0, 1, 2,$$

where

$$g_k = \log\left[\left\{1 + \frac{k+1}{6}(1+(-1)^k)\right\}\mu\right],$$

and where

$$\mu = \begin{cases} 1 \text{ for } n \text{ even} \\ \max\limits_{\pi \leqq h \leqq 3\pi/2} |\sin h|h| = 0.217\cdots \text{ for } n \text{ odd;} \end{cases}$$

and

(1.7) $$R_{2n}^{(3)} = 1 , \quad R_{2n-1}^{(3)} > 1 , \qquad n = 1, 2, \cdots ,$$

(1.8) $$\limsup_{n\to\infty} (2n-1)(R_{2n-1}^{(3)} - 1) \leqq \alpha_0 = 1.07\cdots$$

where α_0 *is the positive root of the equation*

(1.9) $$3 - \alpha - 3\mu e^{\alpha} = 0 .$$

Moreover, $R_n^{(k)}$ *is the largest* r *for which* $\psi_n^{(k)}(r,\theta)$ *is non-negative for all* θ, *where* $\psi_n^{(k)}(r,\theta)$ *is defined for* $k = 0, 1, 2, 3$ *by the equations* (2.18), (2.19), (2.20), *and* (2.21).

Since $v(r,\theta)$ in (1.1) may be regarded as the imaginary part of the analytic function

$$f(z) = \sum_1^\infty a_\nu z^\nu , \quad z = re^{i\theta} , \qquad r < 1, \ a_\nu \text{ real},$$

the property $v(r,\theta) \geqq 0$ for $0 < \theta < \pi$ may be interpreted by saying that $f(z)$ is typically-real in the unit circle, that is $\Im f(z) > 0$ for $\Im z > 0$, and $\Im f(z) < 0$ for $\Im z < 0$, $|z| < 1$. In this case

(1.10) $$F(z) = \int_0^z \frac{f(z)}{z}dz$$

is schlicht and convex in the direction of the imaginary axis for $|z| < 1$. For from (1.10) we have

(1.11) $$\frac{\partial}{\partial\theta}\Re F(re^{i\theta}) = -\Im zF'(z) = -\Im f(z) < 0$$

for $|z| < 1$, $0 < \theta < \pi$.

DEFINITION. Let $\mathscr{F}_r$ and $\mathscr{F}_r^*$ denote the families of functions

$$F(z) = z + b_2 z^2 + \cdots + b_n z^n + \cdots \tag{1.12}$$

which are regular, real on the real axis, schlicht and convex in the direction of the imaginary axis in $|z| < r$, and in $|z| \leqq r$ respectively.

With the help of Theorem 1 we then obtain the following.

THEOREM 2. *Let*

$$F(z) = z + b_2 z^2 + \cdots + b_n z^n + \cdots \tag{1.13}$$

be a member of the family $\mathscr{F}_1$. *Then for* $n = 1, 2, 3, \cdots$ *the* nth *Cesàro partial sum of order one of* (1.13) *is a member of* $\mathscr{F}_{1/2}^*$. *The radius* $1/2$ *cannot be replaced by a larger number. Also the* nth *Cesàro partial sum of order* k, $k = 0, 1, 2, 3$, *is a member of* $\mathscr{F}_{\rho_k}^*$ *where*

$$\rho_0 = 1 - 3n^{-1}\log n + n^{-1}\log\{(3/4 - \varepsilon)\log n\}\ , \qquad n > n_0(\varepsilon),$$

$$\rho_1 = 1 - 2n^{-1}\log n + n^{-1}\log\{(1 - \varepsilon)\log n\}\ , \qquad n > n_1(\varepsilon),$$

$$\rho_2 = 1 - n^{-1}\log n + n^{-1}\log\{1/2 - \varepsilon)\log n\}\ , \qquad n > n_2(\varepsilon),$$

$$\rho_3 \begin{cases} = 1, & n \text{ even}, \\ > 1, & n \text{ odd}, \end{cases}$$

and where $\varepsilon > 0$ *is arbitrarily small and* $n_k(\varepsilon)$, $k = 0, 1, 2$, *are positive integers depending only upon* ε. *The radii* ρ_k *are sharp to within* $O(1/n)$.

2. Preliminary formulas. Before we proceed to the proof of Theorem 1 we shall mention several formulas which will be needed. The following sums are easily calculated:

$$S(z) = z + 2z^2 + \cdots + nz^n + \cdots = z(1-z)^{-2}\ , \tag{2.1}$$

$$S_n^{(0)}(z) = z + 2z^2 + \cdots + nz^n = \{z - (n+1)z^{n+1} + nz^{n+2}\}(1-z)^{-2}\ , \tag{2.2}$$

$$S_n^{(k)}(z) = S_1^{(k-1)}(z) + S_2^{(k-1)}(z) + \cdots + S_n^{(k-1)}(z)\ , \qquad k = 1, 2, \cdots, \tag{2.3}$$

$$S_n^{(k)}(z) = C_k^{n+k-1}z + 2C_k^{n+k-2}z^2 + \cdots + nC_k^k z^n\ , \tag{2.4}$$

$$S_n^{(1)}(z) = \{nz - (n+2)z^2 + (n+2)z^{n+2} - nz^{n+3}\}(1-z)^{-3}\ , \tag{2.5}$$

$$S_n^{(2)}(z) = \frac{1}{2!}\{n(n+1)z - 2n(n+3)z^2 + (n+2)(n+3)z^3 - 2(n+3)z^{n+3} + 2nz^{n+4}\}(1-z)^{-4}\ , \tag{2.6}$$

$$S_n^{(3)}(z) = \frac{1}{3!}\{n(n+1)(n+2)z - 3n(n+1)(n+4)z^2 \tag{2.7}$$

$$+ 3n(n+3)(n+4)z^3 - (n+2)(n+3)(n+4)z^4 + 6(n+4)z^{n+4} - 6nz^{n+5}\}(1-z)^{-5} ,$$

$$(2.8) \quad S_n^{(k)}(z) = \frac{1}{k!}\Big\{\sum_{m=1}^{k+1}(-1)^{m-1}\prod_{p=0}^{k-m}(n+p)\prod_{q=k+3-m}^{k+1}(n+q)z^m + (-1)^{k-1}k!(n+k+1)z^{n+k+1} + (-1)^k k!\, nz^{n+k+2}\Big\}(1-z)^{-k-2} , \quad k = 0, 1, \cdots,$$

where $\prod_{p=i}^{j}(n+p)$ is defined to be 1 if $i > j$.

Let

$$(2.9) \qquad f(z) = \sum_1^\infty a_\nu z^\nu , \qquad a_1 > 0,\ a_\nu \text{ real},$$

be regular and typically-real in $|z| < 1$, which is to say that $v(r, \theta) = \Im f(re^{i\theta})$ is non-negative for $0 \leqq \theta \leqq \pi$, $0 < r < 1$, and $f(z)$ is real on the real axis. As I have shown elsewhere [3] the function $f(z)$ may be represented by the Stieltjes integral

$$(2.10) \qquad f(z) = \frac{a_1}{\pi}\int_0^\pi P(z, \phi)d\alpha(\phi) , \qquad |z| < 1,$$

where $\alpha(\phi)$ is a non-decreasing real function of ϕ in the interval $[0, \pi]$, and where $P(z, \phi)$ is the typically-real, schlicht and star-like function

$$(2.11) \qquad P(z, \phi) \equiv z(1 - 2z\cos\phi + z^2)^{-1} = \sum_1^\infty \frac{\sin \nu\phi}{\sin\phi} z^\nu .$$

For $\phi = 0$ we have $P(z, 0) = S(z)$ where $S(z)$ is defined as in (2.1). From (2.10) we have immediately that

$$(2.12) \qquad S_n^{(k)}(r, \theta) = \frac{a_1}{\pi}\int_0^\pi \Im P_n^{(k)}(re^{i\theta}, \phi)d\alpha(\phi)$$

where $P_n^{(k)}(re^{i\theta}, \phi)$ is the nth Cesàro partial sum of order k of the power series for $P(z, \phi)$ given in (2.11),

$$(2.13) \qquad \Im P_n^{(k)}(re^{i\theta}, \phi) = \sum_{\nu=1}^n C_k^{n+k-\nu} r^\nu \frac{\sin\nu\phi}{\sin\phi}\sin\nu\theta .$$

By a lemma of L. Fejér [8], [9], [4], it follows that

$$(2.14) \qquad \Im P_n^{(k)}(re^{i\theta}, \phi) \geqq 0 , \qquad 0 \leqq \phi \leqq \pi,\ 0 \leqq \theta \leqq \pi,$$

if, and only if,

$$(2.15) \qquad \Im S_n^{(k)}(re^{i\theta}) = \Im P_n^{(k)}(re^{i\theta}, 0) = \sum_{\nu=1}^n \nu C_k^{n+k-\nu} r^\nu \sin\nu\theta \geqq 0, \quad 0 \leqq \theta \leqq \pi.$$

Thus, the behavior of the Cesàro partial sums of the Koebe function

$z(1-z)^{-2}$ determines the extremes to which the Cesàro partial sums of the series expansion of an arbitrary typically-real function $f(z)$ exhibit their properties. Therefore, in order to prove Theorem 1 for the imaginary part of an arbitrary typically-real function $f(z)$ we may confine ourselves to proving these results merely for the Koebe function $S(z) = z(1-z)^{-2}$. For this function the partial sums

$$S_n^{(0)}(z) = \{z - (n+1)z^{n+1} + nz^{n+2}\}(1-z)^{-2} \tag{2.16}$$

are known to be schlicht and star-like with respect to the origin in [1]

$$|z| \leqq 1 - 3n^{-1}\log n\ , \quad n > n_0\ ,$$

and *à fortiori* typically-real in the same radius.

From formulas (2.2), (2.5), (2.6), (2.7), on letting $z = re^{i\theta}$ we obtain by simple, straightforward but long computations the following additional formulas which we shall need.

$$k!\,|1-z|^{2k+4}\Im S_n^{(k)}(re^{i\theta}) = r\sin\theta\,\psi_n^{(k)}(r,\theta)\ , \quad k = 0, 1, 2, 3, \tag{2.17}$$

where

$$\begin{aligned}\psi_n^{(0)}(r,\theta) &= 1 - r^2 - (n+1)r^{n+2}\frac{\sin(n-1)\theta}{\sin\theta} \\ &+ r^{n+1}(2n+2+nr^2)\frac{\sin n\theta}{\sin\theta} - r^n(n+1+2nr^2)\frac{\sin(n+1)\theta}{\sin\theta} \\ &+ nr^{n+1}\frac{\sin(n+2)\theta}{\sin\theta}\ ,\end{aligned} \tag{2.18}$$

$$\begin{aligned}\psi_n^{(1)}(r,\theta) &= \{n + 6r^2 - (n+2)r^4\} - \{(n+2)r - nr^3\}\frac{\sin 2\theta}{\sin\theta} \\ &- (n+2)r^{n+4}\frac{\sin(n-1)\theta}{\sin\theta} + \{(3n+6)r^{n+3} + nr^{n+5}\}\frac{\sin n\theta}{\sin\theta} \\ &- \{(3n+6)r^{n+2} + 3nr^{n+4}\}\frac{\sin(n+1)\theta}{\sin\theta} \\ &+ \{(n+2)r^{n+1} + 3nr^{n+3}\}\frac{\sin(n+2)\theta}{\sin\theta} - nr^{n+2}\frac{\sin(n+3)\theta}{\sin\theta}\ ,\end{aligned} \tag{2.19}$$

$$\begin{aligned}\psi_n^{(2)}(r,\theta) &= \{n(n+1) + (2n^2+18n)r^2 \\ &\qquad - (2n^2 - 6n - 36)r^4 - (n^2+5n+6)r^6\} \\ &- r\{2n^2 + 6n + (16n+24)r^2 - (2n^2+6n)r^4\}\frac{\sin 2\theta}{\sin\theta} \\ &+ r^2\{n^2 + 5n + 6 - (n^2+n)r^2\}\frac{\sin 3\theta}{\sin\theta} \\ &- (2n+6)r^{n+6}\frac{\sin(n-1)\theta}{\sin\theta} + (8n+24+2nr^2)r^{n+5}\frac{\sin n\theta}{\sin\theta}\end{aligned} \tag{2.20}$$

$$- (12n + 36 + 8nr^2)r^{n+4}\frac{\sin (n+1)\theta}{\sin \theta}$$

$$+ (8n + 24 + 12nr^2)r^{n+3}\frac{\sin (n+2)\theta}{\sin \theta}$$

$$- (2n + 6 + 8nr^2)r^{n+2}\frac{\sin (n+3)\theta}{\sin \theta} + 2nr^{n+3}\frac{\sin (n+4)\theta}{\sin \theta},$$

(2.21)
$$\begin{aligned}\varphi_n^{(3)}(r, \theta) = {} & \{n(n+1)(n+2) + 5n(n+1)(n+8)r^2 + 60n(n+4)r^4 \\ & - 5(n+3)(n^2-16)r^6 - (n+2)(n+3)(n+4)r^8\} \\ & - \{3n(n+1)(n+4)r + 5(n^3 + 15n^2 + 32n)r^3 \\ & - 5(n+4)(n^2 - 7n - 12)r^5 - 3n(n+3)(n+4)r^7\}\frac{\sin 2\theta}{\sin \theta} \\ & + \{3n(n+3)(n+4)r^2 + 30(n+2)^2r^4 \\ & - 3n(n+1)(n+4)r^6\}\frac{\sin 3\theta}{\sin \theta} \\ & - \{(n+2)(n+3)(n+4)r^3 - n(n+1)(n+2)r^5\}\frac{\sin 4\theta}{\sin \theta} \\ & - (6n+24)r^{n+8}\frac{\sin (n-1)\theta}{\sin \theta} \\ & + \{(30n+120)r^{n+7} + 6nr^{n+9}\}\frac{\sin n\theta}{\sin \theta} \\ & - \{(60n+240)r^{n+6} + 30nr^{n+8}\}\frac{\sin (n+1)\theta}{\sin \theta} \\ & + \{(60n+240)r^{n+5} + 60nr^{n+7}\}\frac{\sin (n+2)\theta}{\sin \theta} \\ & - \{(30n+120)r^{n+4} + 60nr^{n+6}\}\frac{\sin (n+3)\theta}{\sin \theta} \\ & + \{(6n+24)r^{n+3} + 30nr^{n+5}\}\frac{\sin (n+4)\theta}{\sin \theta} \\ & - 6nr^{n+4}\frac{\sin (n+5)\theta}{\sin \theta}.\end{aligned}$$

3. Proof of Theorem 1 for $k = 1$. We proceed now to the proof of Theorem 1. For $k = 0$ Theorem 1 follows from the theorem of Szász [8]. For $k = 1$ we have $\Im S_n^{(1)}(re^{i\theta}) \geqq 0$ for $0 < \theta < \pi$ provided $\varphi_n^{(1)}(r, \theta) \geqq 0$ for all θ and $r \leqq R_n^{(1)}$. From (2.19) we must determine the largest r for which

(3.1) $$\varphi_n^{(1)}(r, \theta) = \{n + 6r^2 - (n+2)r^4\} - \{(n+2)r - nr^3\}\frac{\sin 2\theta}{\sin \theta}$$

$$- (n+2)r^{n+4}\frac{\sin(n-1)\theta}{\sin\theta} + \{(3n+6)r^{n+3} + nr^{n+5}\}\frac{\sin n\theta}{\sin\theta}$$
$$- \{(3n+6)r^{n+2} + 3nr^{n+4}\}\frac{\sin(n+1)\theta}{\sin\theta}$$
$$+ \{(n+2)r^{n+1} + 3nr^{n+3}\}\frac{\sin(n+2)\theta}{\sin\theta} - nr^{n+2}\frac{\sin(n+3)\theta}{\sin\theta}$$

is non-negative for all θ. We rewrite $\psi_n^{(1)}(r,\theta)$ in the form

$$\psi_n^{(1)}(r,\theta) = A + B(1-\cos\theta) - r^n \sum_{j=0}^{4}(-1)^j C_j \frac{\sin(n-1+j)\theta}{\sin\theta} \tag{3.2}$$

where

$$\begin{aligned} A &= n + 6r^2 - (n+2)r^4 - (2n+4)r + 2nr^3 \\ &= n(1-r)^3(1+r) - 2r(2+r)(1-r)^2, \\ B &= (2n+4)r - 2nr^3, \\ C_0 &= (n+2)r^4, \qquad C_1 = (3n+6)r^3 + nr^5, \\ C_2 &= (3n+6)r^2 + 3nr^4, \qquad C_3 = (n+2)r + 3nr^3, \qquad C_4 = nr^2. \end{aligned} \tag{3.3}$$

Let

$$r = e^{-\varepsilon}, \quad \varepsilon = \frac{2\log n}{n} - \frac{\log\log n}{n} + \frac{p}{n}, \quad r^n = \frac{\log n}{n^2}e^{-p}, \tag{3.4}$$
$$1 - r = 1 - 2\frac{\log n}{n^2} + \frac{\log\log n}{n} - \frac{p}{n} + O\left(\left(\frac{\log n}{n}\right)^2\right).$$

Then

$$A \cong 16\frac{(\log n)^3}{n^2}, \qquad B \cong 8\log n.$$

For fixed k we have

$$r^k = e^{-k\varepsilon} = 1 - k\varepsilon + \frac{k^2}{2}\varepsilon^2 + O(\varepsilon^3),$$

so that

$$\begin{aligned} C_0 &= (n+2) - (4n+8)\varepsilon + 8n\varepsilon^2 + O(\varepsilon^3 n) \\ C_1 &= (4n+6) - (14n+18)\varepsilon + 26n\varepsilon^2 + O(\varepsilon^3 n) \\ C_2 &= (6n+6) - (18n+12)\varepsilon + 30n\varepsilon^2 + O(\varepsilon^3 n) \\ C_3 &= (4n+2) - (10n+2)\varepsilon + 14n\varepsilon^2 + O(\varepsilon^3 n) \\ C_4 &= n \qquad - 2n\varepsilon + 2n\varepsilon^2 + O(\varepsilon^3 n). \end{aligned} \tag{3.5}$$

To obtain an asymptotic estimate for $\psi_n^{(1)}(r,\theta)$ in (3.2) we shall make use of the following lemma.

LEMMA 1. *Let* α_j, $j = 0, 1, \cdots, 5$, *be constants. Let* n *be a positive integer and let*

$$S = \sum_{j=0}^{5} (-1)^j \alpha_j \frac{\sin (n-1+j)\theta}{\sin \theta}$$

$$= \frac{\sin (n-1)\theta}{\sin \theta} \sum_{j=0}^{5} (-1)^j \alpha_j \cos j\theta + \cos (n-1)\theta \sum_{j=1}^{5} (-1)^j \alpha_j \frac{\sin j\theta}{\sin \theta} .$$

(a) *If* $\sum_{j=0}^{5} (-1)^j \alpha_j = 0$, *then*
$S = \{n(1-\cos\theta) + 1\} \cdot \max_j |\alpha_j| \cdot O(1)$ *as* $n \to \infty$.

(b) *If in addition to* (a), $\sum_{j=1}^{5} (-1)^j j^2 \alpha_j = 0$, *then*
$S = \{n(1-\cos\theta)^2 + 1\} \cdot \max_j |\alpha_j| \cdot O(1)$ *as* $n \to \infty$.

(c) *If in addition to* (a) *and* (b), $\sum_{j=1}^{5} (-1)^j j \alpha_j = 0$, *then*
$S = \{n(1-\cos\theta)^2 + (1-\cos\theta)\} \cdot \max_j |\alpha_j| \cdot O(1)$ *as* $n \to \infty$.

The lemma is easily obtained by considering the limits

$$\lim_{\theta\to 0} \frac{\sum_{j=0}^{5}(-1)^j \alpha_j \cos j\theta}{1-\cos\theta}, \qquad \lim_{\theta\to 0} \frac{\sum_{j=1}^{5}(-1)^j \alpha_j \sin j\theta}{\sin\theta - \frac{1}{2}\sin 2\theta} .$$

From (3.2) and (3.5) we obtain

$$\psi_n^{(1)}(r, \theta) = A + B(1-\cos\theta) - r^n[D_0 - D_1\varepsilon + D_2\varepsilon^2 - D_3\varepsilon^3] \tag{3.6}$$

where

$$D_0 = (n+2)\frac{\sin(n-1)\theta}{\sin\theta} - (4n+6)\frac{\sin n\theta}{\sin\theta} + (6n+6)\frac{\sin(n+1)\theta}{\sin\theta} - (4n+2)\frac{\sin(n+2)\theta}{\sin\theta} + n\frac{\sin(n+3)\theta}{\sin\theta} \tag{3.7}$$

$$= 4n\frac{\sin(n+1)\theta}{\sin\theta}(1-\cos\theta)^2 - 4\frac{\sin(n-1)\theta}{\sin\theta}(1+2\cos\theta)(1-\cos\theta)^2 - 4\cos(n-1)\theta(1-2\cos\theta)(1-\cos\theta)$$

$$= 4n\frac{\sin(n+1)\theta}{\sin\theta}(1-\cos\theta)^2 + (1-\cos\theta)O(n) .$$

$$(3.8)\quad D_1 = (4n+8)\frac{\sin(n-1)\theta}{\sin\theta} - (14n+18)\frac{\sin n\theta}{\sin\theta} + (18n+12)\frac{\sin(n+1)\theta}{\sin\theta} - (10n+2)\frac{\sin(n+2)\theta}{\sin\theta} + 2n\frac{\sin(n+3)\theta}{\sin\theta}$$
$$= \{n(1-\cos\theta)+1\}\cdot n\cdot O(1)\,.$$

$$(3.9)\quad D_2 = 8n\frac{\sin(n-1)\theta}{\sin\theta} - 26n\frac{\sin n\theta}{\sin\theta} + 30n\frac{\sin(n+1)\theta}{\sin\theta} - 14n\frac{\sin(n+2)\theta}{\sin\theta} + 2n\frac{\sin(n+3)\theta}{\sin\theta}$$
$$= \{n(1-\cos\theta)+1\}\cdot n\cdot O(1)\,.$$

$$(3.10)\quad D_3 = \sum_{j=0}^{4} O(n)\frac{\sin(n-1+j)\theta}{\sin\theta} = O(n^2)\,.$$

$$(3.11)\quad D_0 - D_1\varepsilon + D_2\varepsilon^2 - D_3\varepsilon^3 = 4n\frac{\sin(n+1)\theta}{\sin\theta}(1-\cos\theta)^2 + (1-\cos\theta)O(n) - \{n^2(1-\cos\theta)+n\}O(1)\frac{\log n}{n} + \{n^2(1-\cos\theta)+n\}O(1)\left(\frac{\log n}{n}\right)^2 + O(n^2)\left(\frac{\log n}{n}\right)^3$$
$$= 4n\frac{\sin(n+1)\theta}{\sin\theta}(1-\cos\theta)^2 + (1-\cos\theta)O(n\log n) + O(\log n)$$

$$(3.12)\quad \varphi_n^{(1)}(r,\theta) = 16\frac{(\log n)^3}{n^2} + 8\log n(1-\cos\theta) - \frac{\log n}{n^2}e^{-p}\left[4n\frac{\sin(n+1)\theta}{\sin\theta}(1-\cos\theta)^2 + (1-\cos\theta)O(n\log n) + O(\log n)\right]$$
$$= 16\frac{(\log n)^3}{n^2}(1+o(1)) + 8\log n(1+o(1))(1-\cos\theta) - 4\frac{\log n}{n}e^{-p}\frac{\sin(n+1)\theta}{\sin\theta}(1-\cos\theta)^2\,.$$

Thus the essential part of $\varphi_n^{(1)}(r,\theta)$ is the expression

$$(3.13)\qquad 4(1-\cos\theta)\log n\left[2 - \frac{e^{-p}}{n}\frac{\sin(n+1)\theta}{\sin\theta}(1-\cos\theta)\right].$$

When n is even, the minimum of the square bracket in (3.13) is reached for $\theta=\pi$. Thus $1-e^{-p}$ must be non-negative. If p denotes a bounded function of n, $p(n)$, we then have $\lim_{n\to\infty} p(2n) = 0$.

If n is odd, we let $-\mu = -0.217\cdots =$ the absolute minimum of $\sin h/h$, which occurs in $\pi < h < \frac{3\pi}{2}$. If c_0 is a sufficiently large constant it is easily seen that the square bracket in (3.13) is positive for

$$0 < \frac{c_0}{n+1} < \theta < \pi - \frac{c_0}{n+1}$$

and that its minimum occurs in the interval $\pi - \{(c_0)/n+1\} < \theta < \pi$ for large odd values of n. Let $\theta = \pi - \{(h)/n+1\}$. Then for n odd

$$\left[2 - \frac{e^{-p}}{n}\frac{\sin(n+1)\theta}{\sin\theta}(1-\cos\theta)\right] = 2\left(1 + e^{-p}\frac{\sin h}{h}\right) + O(1/n^2)$$
$$= 2(1-\mu e^{-p}) + O(n^{-2}) .$$

It follows that

$$\lim_{n\to\infty} p(2n+1) = \log\mu = -1.527\cdots .$$

It follows from the discussion above that we have

$$R_n^{(1)} = 1 - 2\frac{\log n}{n} + \frac{\log\log n}{n} + \frac{\beta}{n} + o(1/n) \tag{3.14}$$

where $\beta = 0 = -\log 1$, if n is even, and where

$$\beta = -\log\left\{\max_{\pi\leqq h\leqq\frac{3\pi}{2}}\left|\frac{\sin h}{h}\right|\right\} = -\log\mu = 1.527\cdots$$

when n is odd. This completes the proof of Theorem 1 for the case $k = 1$.

We note that for $0 \leqq r \leqq 1/2$, $\psi_n^{(1)}(r,\theta) \geqq 0$ for all n and all θ. Indeed, when $r = 1/2$, we obtain from (3.1) that $\psi_n^{(1)}(1/2,\theta) \geqq 0$ provided

$$\begin{aligned}(3.15)\quad &(30n+44)\sin\theta - (12n+32)\sin 2\theta - (2n+4)2^{-n}\sin(n-1)\theta\\ &+(13n+24)2^{-n}\sin n\theta - (30n+48)2^{-n}\sin(n+1)\theta\\ &+(28n+32)2^{-n}\sin(n+2)\theta - 8n\cdot 2^{-n}\sin(n+3)\theta\\ &\geqq 0,\ 0<\theta<\pi .\end{aligned}$$

Since $|\sin k\theta/\sin\theta| \leqq k$, $k = 1, 2, \cdots$, (3.15) is satisfied if

$$(6n-20)2^n \geqq 73n^2 + 192n + 108 . \tag{3.16}$$

It is easily verified that (3.17) is true for $n > 7$. For $1 \leqq n \leqq 7$ the author has verified that $\psi_n(1/2,\theta) \geqq 0$. The calculations are simple but somewhat tedious, and will be omitted.

4. Proof of Theorem 1 for $k = 2$. From (2.20) we have

$$\psi_n^{(2)}(r, \theta) = P_n(r, \theta) - r^n \sum_{j=0}^{5} (-1)^j C_j \frac{\sin (n-1+j)\theta}{\sin \theta} \tag{4.1}$$

where

$$\begin{aligned} P_n(r, \theta) = \{&n(n+1) + (2n^2+18n)r^2 - (2n^2-6n-36)r^4 \\ &- (n^2+5n+6)r^6\} \\ &- \{(2n^2+6n)r + (16n+24)r^3 - (2n^2+6n)r^5\}\frac{\sin 2\theta}{\sin \theta} \\ &+ \{(n^2+5n+6)r^2 - (n^2+n)r^4\}\frac{\sin 3\theta}{\sin \theta} \\ = {}& A + B(1-\cos\theta) + C(1-\cos\theta)^2 , \end{aligned} \tag{4.2}$$

$$\begin{aligned} A &= n^2 + n - (4n^2+12n)r + (5n^2+33n+18)r^2 - (32n+48)r^3 \\ &\quad - (5n^2-3n-36)r^4 + (4n^2+12n)r^5 - (n^2+5n+6)r^6 \\ &= -(1-r)^6(n^2+5n+6) + (1-r)^5(2n^2+18n+36) \\ &\quad - (1-r)^4(12n+54) + 24(1-r)^3 , \end{aligned} \tag{4.3}$$

$$\begin{aligned} B &= (4n^2+12n)r - (8n^2+40n+48)r^2 + (32n+48)r^3 \\ &\quad + (8n^2+8n)r^4 - (4n^2+12n)r^5 \\ &= -r(1-r)^4(4n^2+12n) + r(1-r)^3(8n^2+40n) \\ &\quad - r(1-r)^2(16n-48) - 48r(1-r) . \end{aligned} \tag{4.4}$$

$$\begin{aligned} C &= (4n^2+20n+24)r^2 - (4n^2+4n)r^4 \\ C_0 &= (2n+6)r^6 \\ C_1 &= (8n+24)r^5 + 2nr^7 \\ C_2 &= (12n+36)r^4 + 8nr^6 \end{aligned} \tag{4.5}$$

$$\begin{aligned} C_3 &= (8n+24)r^3 + 12nr^5 \\ C_4 &= (2n+6)r^2 + 8nr^4 \\ C^5 &= \qquad\qquad 2nr^3 . \end{aligned} \tag{4.6}$$

Letting

$$r = e^{-\varepsilon} , \quad \varepsilon = \frac{\log n}{n} - \frac{\log\log n}{n} + \frac{q}{n} ,$$

$$r^n = \frac{\log n}{n} e^{-q} , \quad 1 - r = 1 - \frac{\log n}{n} + \frac{\log\log n - q}{n} + O\left(\left(\frac{\log n}{n}\right)^2\right) ,$$

we obtain

$$A \cong 2\frac{(\log n)^5}{n^3} , \qquad B \cong 8\frac{(\log n)^3}{n} , \qquad C \cong 8n\log n , \tag{4.7}$$

$$(4.8)\quad \begin{aligned} C_0 &= (2n+6)-(12n+36)\varepsilon+(36n+108)\varepsilon^2 \\ &\qquad -(72n+216)\varepsilon^3+108n\varepsilon^4+O(n\varepsilon^5) \\ C_1 &= (10n+24)-(54n+120)\varepsilon+(149n+300)\varepsilon^2 \\ &\qquad -(281n+500)\varepsilon^3+\frac{4901}{12}n\varepsilon^4+O(n\varepsilon^5) \\ C_2 &= (20n+36)-(96n+144)\varepsilon+(240n+288)\varepsilon^2 \\ &\qquad -(416n+384)\varepsilon^3+560n\varepsilon^4+O(n\varepsilon^5) \\ C_3 &= (20n+24)-(84n+72)\varepsilon+(186n+108)\varepsilon^2 \\ &\qquad -(286n+108)\varepsilon^3+\frac{679}{2}n\varepsilon^4+O(n\varepsilon^5) \\ C_4 &= (10n+6)-(36n+12)\varepsilon+(68n+12)\varepsilon^2 \\ &\qquad -(88n+8)\varepsilon^3+\frac{260}{3}n\varepsilon^4+O(n\varepsilon^5) \\ C_5 &= 2n-6n\cdot\varepsilon+9n\cdot\varepsilon^2-9n\cdot\varepsilon^3+\frac{27}{4}n\varepsilon^4+O(n\varepsilon^5)\,. \end{aligned}$$

We now write

$$(4.9)\quad \varphi_n^{(2)}(r,\theta) = A + B(1-\cos\theta) + C(1-\cos\theta)^2 - r^n\cdot\sum_{j=0}^{5}(-1)^j D_j\varepsilon^j\,.$$

From (4.1), (4.2), and (4.8), we find

$$\begin{aligned} D_0\cdot\sin\theta &= (2n+6)\sin(n-1)\theta-(10n+24)\sin n\theta \\ &\qquad -(20n+36)\sin(n+1)\theta+(20n+24)\sin(n+2)\theta \\ &\qquad\qquad -(10n+6)\sin(n+3)\theta+2n\sin(n+4)\theta\,, \end{aligned}$$

$$D_0 = \left[-8n\cdot\frac{\cos(2n+3)\dfrac{\theta}{2}}{\cos\dfrac{\theta}{2}}+24\frac{\sin(n+1)\theta}{\sin\theta}\right](1-\cos\theta)^2.$$

$$\begin{aligned} D_1\cdot\sin\theta &= (12n+36)\sin(n-1)\theta-(54n+120)\sin n\theta \\ &\qquad +(96n+144)\sin(n+1)\theta-(84n+72)\sin(n+2)\theta \\ &\qquad\qquad +(36n+12)\sin(n+3)\theta-6n\sin(n+4)\theta\,. \end{aligned}$$

By Lemma 1, we obtain

$$\begin{aligned} D_1 &= [n(1-\cos\theta)^2+(1-\cos\theta]\cdot[O(n)+O(1)]\cdot O(1) \\ &= [n^2(1-\cos\theta)^2+n(1-\cos\theta)]\cdot O(1)\,. \end{aligned}$$

$$\begin{aligned} D_2\cdot\sin\theta &= (36n+108)\sin(n-1)\theta-(149n+300)\sin n\theta \\ &\qquad +(240n+288)\sin(n+1)\theta-(186n+108)\sin(n+2)\theta \\ &\qquad\qquad +(68n+12)\sin(n+3)\theta-9n\sin(n+4)\theta\,, \end{aligned}$$

$$D_2 = [n^2(1-\cos\theta)^2 + n(1-\cos\theta)]\cdot O(1) + [n(1-\cos\theta) + (1-\cos\theta)]\cdot O(1)\,.$$

$$D_3\cdot\sin\theta = (72n+216)\sin(n-1)\theta - (281n+500)\sin n\theta + (416n+384)\sin(n+1)\theta - (286n+108)\sin(n+2)\theta + (88n+8)\sin(n+3)\theta - 9n\sin(n+4)\theta\,,$$

$$D_3 = \{n(1-\cos\theta)+1\}\cdot n\cdot O(1) = [n^2(1-\cos\theta)+n]O(1)\,.$$

$$D_4\cdot\sin\theta = 108n\sin(n-1)\theta - \frac{4901}{12}n\sin n\theta + 560n\sin(n+1)\theta - \frac{679}{2}\sin(n+2)\theta + \frac{260}{3}n\sin(n+3)\theta - \frac{27}{4}n\sin(n+4)\theta\,,$$

$$D_4 = \{n(1-\cos\theta)+1\}\cdot n\cdot O(1) = [n^2(1-\cos\theta)+n]\cdot O(1)\,.$$

$$D_5 = O(n)\cdot\sum_{j=0}^{5}\left|\frac{\sin(n-1+j)\theta}{\sin\theta}\right| = O(n^2)\,.$$

$$(4.10)\qquad D_0 - D_1\varepsilon + D_2\varepsilon^2 - D_3\varepsilon^3 + D_4\varepsilon^4 - D_5\varepsilon^5$$
$$= \left[-8n\frac{\cos(2n+3)\theta/2}{\cos\theta/2} + 24\frac{\sin(n+1)\theta}{\sin\theta}\right](1-\cos\theta)^2$$
$$+ [n^2(1-\cos\theta)^2 + n(1-\cos\theta)]O(1)\frac{\log n}{n}$$
$$+ [n^2(1-\cos\theta)^2 + n(1-\cos\theta)]O(1)\left(\frac{\log n}{n}\right)^2$$
$$+ [n^2(1-\cos\theta) + n]O(1)\left(\frac{\log n}{n}\right)^3$$
$$+ [n^2(1-\cos\theta) + n]O(1)\left(\frac{\log n}{n}\right)^4 + O(n^2)\left(\frac{\log n}{n}\right)^5$$
$$= \left[-8n\frac{\cos(2n+3)\theta/2}{\cos\theta/2} + 24\frac{\sin(n+1)\theta}{\sin\theta} + O(n\log n)\right](1-\cos\theta)^2$$
$$+ (1-\cos\theta)\cdot O(\log n) + O\left(\frac{(\log n)^3}{n^2}\right)\,.$$

From (4.7), (4.9), and (4.10), we have

$$(4.11)\quad \psi_n^{(2)}(r,\theta) = \frac{2(\log n)^5}{n^3} + \frac{8(\log n)^3}{n}(1-\cos\theta) + 8n\log n(1-\cos\theta)^2$$
$$- \frac{\log n}{n}e^{-q}\left[-8n\frac{\cos(2n+3)\theta/2}{\cos\theta/2} + 24\frac{\sin(n+1)\theta}{\sin\theta} + O(n\log n)\right](1-\cos\theta)^2$$
$$- \frac{\log n}{n}e^{-q}\left[(1-\cos\theta)O(\log n) + O\left(\frac{(\log n)^3}{n^2}\right)\right]$$

$$= \left[2\frac{(\log n)^5}{n^3} - e^{-q}\cdot O\left(\frac{(\log n)^4}{n^3}\right)\right]$$

$$+ \left[8\frac{(\log n)^3}{n} - e^{-q}\cdot O\left(\frac{(\log n)^2}{n}\right)\right](1-\cos\theta)$$

$$+ 8\log n\left[n + e^{-q}\frac{\cos(2n+3)\theta/2}{\cos\theta/2} - \frac{24e^{-q}}{n}\frac{\sin(n+1)\theta}{\sin\theta}\right.$$

$$\left. + e^{-q}\cdot O(\log n)\right](1-\cos\theta)^2 .$$

From (4.11) it is seen that we must have the quantity $L \geqq 0$ where

$$L = 1 + e^{-q}\left[\frac{\cos(2n+3)\theta/2}{n\cos(\theta/2)} - \frac{24}{n^2}\frac{\sin(n+1)\theta}{\sin\theta}\right]. \tag{4.12}$$

For n even the minimum of L is attained for $\theta = \pi$ and equals

$$1 - e^{-q}\left(\frac{2n+3}{n} + \frac{24(n+1)}{n^2}\right) = 1 - 2e^{-q} + e^{-q}\cdot o\left(\frac{1}{n}\right).$$

Thus if $q = q(n)$, a bounded function of n, we require

$$\lim_{n\to\infty} q(2n) = \log 2 . \tag{4.13}$$

If n is odd, we let $\theta = \pi - (2h)/2n + 3$ and find that

$$L \cong 1 + 2e^{-q}\frac{\sin h}{h} \qquad \text{as } n \to \infty ,$$

and

$$\min_\theta L \cong 1 - 2\mu e^{-q}$$

where

$$\mu = 0.217\cdots = \max_{\pi \leqq h \leqq \frac{3\pi}{2}}\left|\frac{\sin h}{h}\right|, \quad \text{and} \quad \lim_{n\to\infty} q(2n-1) = \log(2\mu) .$$

It follows that we have

$$R_n^{(2)} = 1 - \frac{\log n}{n} + \frac{\log\log n}{n} - \frac{\gamma}{n} + o\left(\frac{1}{n}\right) \tag{4.14}$$

where

$$\gamma = \begin{cases} \log 2, & \text{if } n \text{ is even,} \\ \log(2\mu), & \text{if } n \text{ is odd.} \end{cases}$$

This completes the proof of Theorem 1 for the case $k = 2$.

In the case $k = 0$, which was investigated by Szász [8], if we employ procedures analogous to those above for $k = 1$ and 2, we are led to the expression

$$(4.15)\qquad 3\frac{\log n}{n} - \frac{\log n}{n^3}e^{-t}\cdot 2n(1-\cos\theta)\frac{\cos(2n+1)\theta/2}{\cos\theta/2}$$

when

$$r = 1 - 3\frac{\log n}{n} + \frac{\log\log n}{n} - \frac{t}{n}\,.$$

With arguments similar to those used above, we find that the "correct" value of t is $\log(4/3)$ when n is even (as Szász obtained [8]), and $\log(4\mu/3)$ when n is odd, the latter result being new.

5. Proof of Theorem 1 for $k = 3$. The theorem of Fejér [2], quoted in the introduction, states that

$$(5.1)\qquad R_n^{(3)} \geqq 1\,,\qquad n = 1, 2, 3, \cdots.$$

We shall give a new and simple proof of (5.1), and also give a demonstration of the sharpened result

$$(5.2)\qquad R_{2n}^{(3)} = 1\,,\quad R_{2n-1}^{(3)} > 1\,,\qquad n = 1, 2, 3, \cdots,$$

and

$$(5.3)\qquad \limsup_{n\to\infty}(2n-1)(R_{2n-1}^{(3)} - 1) \leqq \alpha_0 = 1.07\cdots$$

where α_0 is the positive root of the equation

$$(5.4)\qquad 3 - \alpha - 3\mu e^{\alpha} = 0$$

where

$$\mu = \max_{\pi\leqq h\leqq\frac{3\pi}{2}}\left|\frac{\sin h}{h}\right| = 0.217\cdots.$$

From (2.7) we have

$$(5.5)\qquad 6(1-z)^5 S_{n-2}^{(3)}(z) = n(n-1)(n-2)z - 3(n-1)(n^2-4)z^2 + 3(n+1)(n^2-4)z^3 - n(n+1)(n+2)z^4 + 6(n+2)z^{n+2} - 6(n-2)z^{n+3}\,.$$

Letting $z = e^{i\theta}$ in (5.5) we have for $n > 2$

$$(5.6)\qquad \Im S_{n-2}^{(3)}(z) = \frac{1}{32\sin^5(\theta/2)}\left[(n^2-4)\cos\frac{\theta}{2} - n^2\cos\frac{3\theta}{2} + (n+2)\cos(2n-1)\frac{\theta}{2} - (n-2)\cos(2n+1)\frac{\theta}{2}\right]$$
$$= \frac{\sin\theta}{16\sin^4(\theta/2)}\left[n^2 + n\frac{\sin n\theta}{\sin\theta} - 2\left(\frac{\sin n\theta/2}{\sin\theta/2}\right)^2\right].$$

In earlier papers we have shown [5], [6], that

$$(5.7)\qquad n^2 + n\frac{\sin n\theta}{\sin\theta} - 2\left(\frac{\sin n\theta/2}{\sin\theta/2}\right)^2 \geqq 0\,,$$

for all integers n and all θ. From (5.6) and (5.7) we have at once that

$$(5.8)\qquad \mathfrak{I}S_n^{(3)}(e^{i\theta}) \geqq 0\,,\quad 0 \leqq \theta \leqq \pi\,,\qquad n = 1, 2, \cdots .$$

However, the function

$$(5.9)\qquad F(z) = (z^{-1} - z)S_n^{(3)}(z)$$

is analytic in $|z| \leqq 1$, and $\Re F(e^{i\theta}) = 2\sin\theta\,\mathfrak{I}S_n^{(3)}(e^{i\theta}) \geqq 0$. Since the minimum of the harmonic function $\Re F(z)$ in $|z| \leqq 1$ occurs on $|z| = 1$ we have $\Re F(z) > 0$ for $|z| < 1$. From the representation (5.9) it follows from the work of Rogosinski [7] that $S_n^{(3)}(z)$ is typically-real in the unit circle, which is to say that

$$(5.10)\qquad \mathfrak{I}S_n^{(3)}(re^{i\theta}) \geqq 0\,,\qquad 0 \leqq \theta \leqq \pi,\ 0 \leqq r \leqq 1.$$

The theorem of Fejér, or inequality (5.1) follows from (5.10) and the remarks made in section two.

We now attack the problem from an alternative point of view for the case $k = 3$. From (2.17) and (5.6) we write

$$(5.11)\qquad \psi_n^{(3)}(r,\theta) = 384\sin^6\frac{\theta}{2}\left[(n+2)^2 + (n+2)\frac{\sin(n+2)\theta}{\sin\theta} - 2\left(\frac{\sin(n+2)\theta/2}{\sin\theta/2}\right)^2\right] + [\psi_n^{(3)}(r,\theta) - \psi_n^{(3)}(1,\theta)]\,.$$

Let $r = 1 + \alpha/n$ where $\alpha = \alpha(n) = O(1) > 0$. Then $r^k - 1 = k\alpha/n + O(\alpha^2/n^2)$, $k =$ positive integer independent of n, $r^{n+k} - 1 = e^\alpha - 1 + n^{-1}k\alpha e^\alpha + O(n^{-1}\alpha^2 e^\alpha)$. From (2.21) and (5.11) we then obtain for $r = 1 + \alpha/n$ asymptotically,

$$(5.12)\qquad \psi_n^{(3)}(r,\theta) - \psi_n^{(3)}(1,\theta) \cong 28n^2\alpha + 56n^2\alpha\cos\theta - 12n^2\alpha(4\cos^2\theta - 1) - 2n^2\alpha(4\cos\theta - 8\cos^3\theta) - \frac{(e^\alpha - 1)S}{\sin\theta}$$

$$= -128n^2\alpha\sin^6\frac{\theta}{2} - \frac{(e^\alpha - 1)S}{\sin\theta}\,,$$

where

$$S = 6n\sin(n-1)\theta - 36n\sin n\theta + 90n\sin(n+1)\theta - 120n\sin(n+2)\theta + 90n\sin(n+3)\theta - 36n\sin(n+4)\theta + 6n\sin(n+5)\theta$$

$$= -384n\sin(n+2)\theta\sin^6\frac{\theta}{2}\,,$$

$$\psi_n^{(3)}(r,\theta) - \psi_n^{(3)}(1,\theta) \cong 128\sin^6\frac{\theta}{2}\left[3n(e^{\alpha}-1)\frac{\sin(n+2)\theta}{\sin\theta} - \alpha n^2\right]. \tag{5.13}$$

Since for sufficiently large values of n we can only have

$$\psi_n^{(3)}(r,\theta) \geqq 0 \quad \text{or} \quad r = 1 + n^{-1}\alpha,\ \alpha = \alpha(n) > 0\ ,$$

provided

$$3\left[(n+2)^2 + (n+2)\frac{\sin(n+2)\theta}{\sin\theta} - 2\left(\frac{\sin(n+2)\theta/2}{\sin\theta/2}\right)^2\right] + 3(e^{\alpha}-1)n\frac{\sin(n+2)\theta}{\sin\theta} - n^2\alpha \geqq 0\ , \tag{5.14}$$

we see that, when n is even and $\theta = \pi$, we must have

$$-3(e^{\alpha}-1)n(n+2) - n^2\alpha \geqq 0\ . \tag{5.15}$$

(5.15) implies that α is non-positive, contrary to our assumption that $\alpha > 0$. Hence $\alpha = 0$ for n even and sufficiently large. However, it is easily seen that $\alpha = 0$ for all even n by the following argument. Since

$$S_n^{(3)}(z) = \sum_{\nu=1}^{n} \nu C_3^{n+3-\nu} z^{\nu}\ , \tag{5.16}$$

and because of the identity

$$\sum_{\nu=1}^{n}(-1)^{\nu}\nu^2(n+1-\nu)(n+2-\nu)(n+3-\nu) = 0\ , \quad n \text{ even}, \tag{5.17}$$

it follows that the derivative of $S_n^{(3)}(z)$ vanishes at $z = -1$ for n even. $S_n^{(3)}(z)$, typically-real in $|z| \leqq 1$, therefore cannot be typically-real in $|z| \leqq r$ for $r > 1$, n even. Thus $\alpha = 0$ for all even n, and $R_{2n}^{(3)} = 1$, $n = 1, 2, \cdots$.

The situation for n odd is not so simple. Fejér has pointed out [2] that $\Im S_n^{(3)}(e^{i\theta}) > 0$, $0 < \theta < \pi$, from which it follows that $\psi_n^{(3)}(1,\theta) > 0$ for all θ with the possible exception of the values $\theta = 0$ and π. When n is odd, however, it is easily seen from (5.6) and (2.17) that $\psi_n^{(3)}(1,\pi) > 0$. From (5.16) it also follows that

$$\lim_{\theta\to 0}\frac{\Im S_n^{(3)}(e^{i\theta})}{\sin\theta} = \sum_{v=1}^{n}\nu^2 C_3^{(n+3-\nu)} > 0\ .$$

Consequently $\Im S_n^{(3)}(e^{i\theta}) \cdot \operatorname{cosec}\theta > 0$ for all θ when n is odd, and so $R_{2n-1}^{(3)} > 1$, $n = 1, 2, \cdots$. To obtain an asymptotic upper bound for $R_{2n-1}^{(3)}$ we shall show that (5.14) is not verified, when n is sufficiently large, for all θ when α exceeds $\alpha_0 = 1.07\cdots$, α_0 being the positive root of the equation (5.4).

Letting $\theta = \pi - [h/(n+2)]$, n odd, we find that the left hand side of inequality (5.14) is asymptotically equal to the expression

$$3n^2\left[1+\frac{\sin h}{h}-O(n^{-2})\right]+3n^2(e^\alpha-1)\left(\frac{\sin h}{h}+O(n^{-1})\right)-n^2\alpha\ ,$$

from which (5.4) and (5.3) follow. It should be noticed that the constant α_0 in (5.3) could be replaced by a smaller one. Indeed, for $\theta = x/n$, the left-hand side of (5.14) is asymptotically equal to

$$n^2\left[3-\alpha+3e^\alpha\frac{\sin x}{x}-6\left(\frac{\sin x/2}{x/2}\right)^2\right]. \tag{5.18}$$

Calculation of the smallest positive α for which the expression (5.18) is non-positive for some $x \geqq \pi$ would lead to a smaller constant to replace α_0.

From (2.3) and (5.1) it follows at once that $R_n^{(k)} \geqq 1$ for $k \geqq 3$ and all positive integers n.

REFERENCES

1. R. K. Brown, *On the partial sums of a typically-real function*, Master of Science thesis (1950), Rutgers University Library.
2. L. Fejér, *Neue Eigenschaften der Mittelwerte bei den Fourierreihen*, J. London Math. Soc. **8** (1933), 53-62.
3. M. S. Robertson, *On the coefficients of a typically-real function*, Bull. Amer. Math. Soc. **41** (1935), 565-572.
4. ——, *Applications of a lemma of Fejér to typically-real functions*, Proc. Amer. Math. Soc. **1** (1950), 555-561.
5. ——, *Univalent power series with multiply monotonic sequences of coefficients*, Ann. of Math. **46** (1945), 533-55.
6. ——, *The coefficients of univalent functions*, Bull. Amer. Math. Soc. **51** (1945), 733-738.
7. W. Rogosinski, *Über positive harmonische Entwicklungen und typischreelle Potenzreihen*, Math. Z. **35** (1932), 93-121.
8. O. Szász, *On the partial sums of harmonic developments and of power series*, Trans. Amer. Math. Soc. **52** (1942), 12-21.
9. P. Turán, *Über die monotone Konvergenz der Cesàro-mittel bei Fourier-und Potenzreihen*, Proc. Cambridge Philos. Soc. **34** (1938), 134-143.

RUTGERS—THE STATE UNIVERSITY OF NEW JERSEY.

ON THE DETERMINATION OF NUMBERS BY THEIR SUMS OF A FIXED ORDER

J. L. SELFRIDGE AND E. G. STRAUS

1. Introduction. We wish to treat the following problem (suggested by a problem of L. Moser [2]):

Let $\{x\} = \{x_1, \cdots, x_n\}$ be a set of complex numbers (if one is interested in generality, one may consider them elements of an algebraically closed field of characteristic zero) and let $\{\sigma\} = \{\sigma_1, \cdots, \sigma_{\binom{n}{s}}\}$ be the set of sums of s distinct elements of $\{x\}$. To what extent is $\{x\}$ determined by $\{\sigma\}$ and what sets can be $\{\sigma\}$ sets?

In § 2 we answer this question for $s = 2$. In § 3 we treat the question for general s.

2. The case $s = 2$.

THEOREM 1. *If $n \neq 2^k$ then the first n elementary symmetric functions of $\{\sigma\}$ can be prescribed arbitrarily and they determine $\{x\}$ uniquely.*

Proof. Instead of the elementary symmetric functions we consider the sums of powers, setting

$$\sum\nolimits_k = \sum_{i=1}^{\binom{n}{2}} \sigma_i^k , \qquad S_k = \sum_{i=1}^{n} x_i^k .$$

Then

$$(1) \qquad \sum\nolimits_k = \sum_{i=1}^{\binom{n}{2}} \sigma_i^k = \sum_{1 \le i_1 < i_2 \le n} (x_{i_1} + x_{i_2})^k = \frac{1}{2} \sum_{\substack{i_1, i_2 = 1 \\ i_1 \neq i_2}}^{n} (x_{i_1} + x_{i_2})^k$$

$$= \frac{1}{2}\Big(\sum_{i_1, i_2 = 1}^{n} (x_{i_1} + x_{i_2})^k - \sum_{i=1}^{n} (2x_i)^k \Big) .$$

Expanding the binomials and collecting like powers we obtain

$$\sum\nolimits_k = \frac{1}{2}\Big(\sum_{l=0}^{k} \binom{k}{l} S_l S_{k-l} - 2^k S_k \Big)$$

$$= \frac{1}{2}(2n - 2^k) S_k + \frac{1}{2} \sum_{l=1}^{k-1} \binom{k}{l} S_l S_{k-l}$$

Thus, since the coefficient of S_k does not vanish, we can solve re-

Received May 16, 1958.

cursively for S_k in terms of $\sum_1, \cdots, \sum_k$. In particular $\sum_1, \cdots, \sum_n$ determine $S_1, \cdots, S_n$—and hence $x_1, \cdots, x_n$—uniquely.

THEOREM 2. *If $n = 2^k$ then $\sum_1, \cdots, \sum_{k+1}$ must satisfy a certain algebraic equation and $\{\sigma\}$ will not always determine $\{x\}$.*

Proof. Equation (1) for $\sum_{k+1}$ yields

$$\sum_{k+1} = \frac{1}{2} \sum_{l=1}^{k} \binom{k+1}{l} S_l S_{k+1-l} \tag{2}$$

where $S_1, \cdots, S_k$ are expressed by (1) as polynomials in $\sum_1, \cdots, \sum_k$.

To prove the second part of the theorem we proceed by induction. Assume there are two different sets $\{x_1, \cdots, x_{2^{k-1}}\}$, $\{y_1, \cdots, y_{2^{k-1}}\}$ which have the same $\{\sigma\}$. Then consider the two sets

$$\{X\} = \{x_1 + a, \cdots, x_{2^{k-1}} + a, y_1, \cdots, y_{2^{k-1}}\}$$
$$\{Y\} = \{x_1, \cdots, x_{2^{k-1}}, y_1 + a, \cdots, y_{2^{k-1}} + a\} .$$

Clearly every sum of two elements of $\{X\}$ is either σ_i or $\sigma_i + 2a$ or $x_i + y_j + a$ and the same holds for the sum of two elements of $\{Y\}$.

The sets $\{X\}$, $\{Y\}$ will clearly be different for some a. To show that they are different for any $a \neq 0$, rearrange $\{x\}$ and $\{y\}$ so that $x_i = y_i$; $i = 1, 2, \cdots, m$; $m \geq 0$, and $x_j \neq y_k$ for $j, k > m$. Then since $y_i + a = x_i + a$; $i = 1, 2, \cdots, m$, the sets $\{X\}$ and $\{Y\}$ will be different if $\{x_j \mid j > m\}$ is different from $\{x_j + a \mid j > m\}$. But this is clear for any $a \neq 0$.

Since $\{\sigma\}$ clearly does not determine $\{x\}$ for $n = 2$ the proof is complete.

In a sense we have completed the answer of the question raised in the introduction for $s = 2$, however there remain some unanswered questions in case $n = 2^k$.

1. *If $\{\sigma\}$ does not determine $\{x\}$ can there be more than two sets giving rise to same $\{\sigma\}$?*

The answer is trivially "yes" for $k = 0, 1$ and is "no" for $k = 2$. It seems probable that the answer is "no" for all $k \geq 2$, however we can see no simple way of proving this.

2. *For what values of n does there exist for all (real) $\{x\}$ a transformation $y_i = f_i(x_1, \cdots, x_n)$, different from a permutation, so that $\{x\}$ and $\{y\}$ give rise to the same $\{\sigma\}$?*

This question was suggested by T. S. Motzkin who gave the answer for $s = 2$.

LEMMA 1. *If $n > s$ and the above functions f_i exist then they are linear.*

Proof. The sets $\{y\}$, $\{x\}$ are connected by a system of equations

$$y_{i_1} + \cdots + y_{i_s} = x_{j_1} + \cdots + x_{j_s} .$$

Here the indices $i_1, \cdots, i_s$ are themselves functions of $\{x\}$. However, since they assume only a finite set of values, there exists a somewhere dense set of $\{x\}$ for which the indices are constant. We restrict our attention to that set. Let $\Delta_k^{(h)} y_i = f_i(x_1, \cdots, x_k + h, \cdots, x_n) - f_i(x_1, \cdots, x_k, \cdots, x_n)$ then we obtain

$$\Delta_k^{(h)} y_{i_1} + \cdots + \Delta_k^{(h)} y_{i_s} = 0 \text{ or } h . \tag{3}$$

If we let u_i be the difference of $\Delta_k^{(h)} y_i$ for two different sets of values of $\{x\}$ then, since the right-hand side of (3) is independent of the choice of $\{x\}$, we obtain

$$u_{i_1} + \cdots + u_{i_s} = 0 . \tag{4}$$

Summation over all sets $\{i_1, \cdots, i_s\} \subset \{1, \cdots, n\}$ yields

$$u_1 + u_2 + \cdots + u_n = 0 . \tag{5}$$

Now let t be the least positive integer so that $u_{i_1} + \cdots + u_{i_t} = 0$ for all $\{i_1, \cdots, i_t\} \subset \{1, \cdots, n\}$. Then $t \mid n$, for $n = mt + r$ with $0 < r < t$ implies

$$u_{i_1} + \cdots + u_{i_r} = u_1 + u_2 + \cdots + u_n - \sum(u_{j_1} + \cdots + u_{j_t}) = 0$$

for all $\{i_1, \cdots, i_r\} \subset \{1, \cdots, n\}$, contrary to hypothesis.

Since $n > s \geq t$ we must have $n \geq 2t$. If $t > 1$ then

$$u_j = -(u_{i_1} + \cdots + u_{i_{t-1}}) \text{ for every } j \notin \{i_1, \cdots, i_{t-1}\} .$$

But there are more than t such j, say $j_1, \cdots, j_t$. Hence

$$u_{j_1} + \cdots + u_{j_t} = -t(u_{i_1} + \cdots + u_{i_{t-1}}) = 0$$

or $u_{i_1} + \cdots + u_{i_{t-1}} = 0$ for every $\{i_1, \cdots, i_{t-1}\} \subset \{1, \cdots, n\}$ contrary to hypothesis. Thus $t = 1$ and

$$u_1 = u_2 = \cdots = u_n = 0 .$$

In other words $\Delta_k^{(h)} y_i = a_{ik}^{(h)} = \text{const.}$ Thus $\Delta_k^{(h_1)} y_i + \Delta_k^{(h_2)} y_i = \Delta_k^{(h_1+h_2)} y_i$ so that $a_{ik}^{(h)} = a_{ik} h$ and

$$y_i = \sum_k a_{ik} x_k .$$

THEOREM 3. *If $n > s$ and there exists a nontrivial transformation $y_i = f_i(x_1, \cdots, x_n)$ which preserves $\{\sigma\}$ then $n = 2s$ and the transformation is linear with matrix (up to permutations)*

$$\begin{pmatrix} -\frac{s-1}{s} & \frac{1}{s} & \cdots & \frac{1}{s} \\ \frac{1}{s} & -\frac{s-1}{s} & \cdots & \frac{1}{s} \\ \vdots & \vdots & & \vdots \\ \frac{1}{s} & \frac{1}{s} & \cdots & -\frac{s-1}{s} \end{pmatrix}$$

Proof. We know by Lemma 1 that the transformation must be linear. Let $y_i = \sum_k a_{ik} x_k$ then

$$(6) \qquad y_{i_1} + \cdots + y_{i_s} = \sum_k (a_{i_1 k} + \cdots + a_{i_s k}) x_k = x_{j_1} + \cdots + x_{j_s} \,.$$

Hence, for fixed k, we have

$$(7) \qquad a_{i_1 k} + \cdots + a_{i_s k} = \begin{cases} 0 \text{ for } \binom{n-1}{s} \text{ sets } \{i_1, \cdots, i_s\} \\ 1 \text{ for } \binom{n-1}{s-1} \text{ sets } \{i_1, \cdots, i_s\} \,. \end{cases}$$

Since $n > s$ two elements a_{ik}, a_{jk} in the same column satisfy

$$a_{ik} + a_{i_1 k} + \cdots + a_{i_{s-1} k} = 0 \text{ or } 1 \,; \; a_{jk} + a_{i_1 k} + \cdots + a_{i_{s-1} k} = 0 \text{ or } 1$$

where $\{i_1, \cdots, i_{s-1}\} \subseteq \{1, \cdots, n\} - \{i, j\}$.

Hence

$$(8) \qquad a_{ik} = a_{jk} \text{ or } a_{ik} = a_{jk} \pm 1 \,.$$

Let the two values assumed by terms in the kth column be a_k and $1 + a_k$. From (6) we see that both values must occur. On the other hand if both a_k and $1 + a_k$ would occur more than once then $\max(a_{i_1 k} + \cdots + a_{i_s k}) - \min(a_{i_1 k} + \cdots + a_{i_s k}) \geq 2$ in contradiction to (7).

If $1 + a_k$ is assumed only once, say $a_{kk} = 1 + a_k$, then $0 = s a_k$ or

$$(9) \qquad a_{ik} = \begin{cases} 1 & i = k \\ 0 & i \neq k \,. \end{cases}$$

According to (6) we have

$$(10) \qquad \sum_{k=1}^{n} (a_{i_1 k} + \cdots + a_{i_s k}) = s \qquad \{i_1, \cdots, i_s\} \subset \{1, \cdots, n\} \,.$$

We now repeat the argument that led to equation (8). Since $n > s$

we can write for any pair (i, j)

$$\sum_{k=1}^{n}(a_{i_1k} + \cdots + a_{i_{s-1}k}) + \sum_{k=1}^{n} a_{ik} = \sum_{k=1}^{n}(a_{i_1k} + \cdots + a_{i_{s-1}k}) + \sum_{k=1}^{n} a_{jk} = s$$

where $\{i_1, \cdots, i_{s-1}\} \subset \{1, \cdots, n\} - \{i, j\}$. Hence $\sum_{k=1}^{n} a_{ik} = \sum_{k=1}^{n} a_{jk}$ and according to (10), $s \sum_{k=1}^{n} a_{ik} = s$ so that

$$\sum_{k=1}^{n} a_{ik} = 1 \qquad i = 1, \cdots, n . \tag{11}$$

Combining (9) and (11) we obtain

$$a_{kj} = \begin{cases} 1 & j = k \\ 0 & j \neq k . \end{cases} \tag{12}$$

In other words, every column contains 0 and therefore $a_k = 0$ for $k = 1, \cdots, n$. Thus the transformation is a permutation.

The only nontrivial case arises therefore if the value a_k occurs only once, say $a_{kk} = a_k$. Then $s - 1 + sa_k = 0$ and

$$a_{ik} = \begin{cases} -(s-1)/s & i = k \\ 1/s & i \neq k . \end{cases} \tag{13}$$

Combining (11) and (13) we obtain

$$\sum_{k=1}^{n}\sum_{i=1}^{n} a_{ik} = n = \frac{n(n-1)}{s} - n\frac{s-1}{s} = \frac{n}{s}(n-s) \tag{14}$$

and hence $n = 2s$. It is now clear from (11) that each row and column contains exactly one term $-(s-1)/s$ and that the matrix (up to permutation) is the one given in the theorem.

3. General s. The procedure which led to Theorem 1 can be generalized. First we define, for every s, a function which is a polynomial in $n, 2^k, 3^k, \cdots, s^k$. Let

$$f(n, k) = \frac{1}{s}\sum_{P}(-1)^{s-t}n^{t-1}\sum_{i=1}^{r} a_i i^k \tag{15}$$

where the outer summation is over all permutations P on s marks, each permutation being composed of a_i i-cycles $i = 1, \cdots, r$, and $t = a_1 + \cdots + a_r$. Thus

$$\begin{aligned} f(n, k) = n^{s-1} &- \frac{1}{2}(s-1)(2^k + s - 2)n^{s-2} + (s-1)(s-2)\Big[\frac{1}{3}(3^k + s - 3) \\ &+ \frac{1}{8}(s-3)(2^{k+1} + s - 4)\Big]n^{s-3} - \cdots + (-1)^s(s-1)!\Big(\sum_{i=1}^{s-1}\frac{i^{k-1}}{s-i}\Big)n \\ &- (-1)^s(s-1)!\, s^{k-1} . \end{aligned} \tag{16}$$

THEOREM 4. *For every s consider the system of Diophantine equations* $f(n, k) = 0 \quad k = 1, 2, \cdots, n$. *If n satisfies none of these then the first n elementary symmetric functions of* $\{\sigma\}$ *can be prescribed arbitrarily and they determine* $\{x\}$ *uniquely. If* $f(n, k) = 0$, *then the first k elementary symmetric functions of* $\{\sigma\}$ *must satisfy an algebraic equation.*

Proof. In the notation of Theorem 1 we have

$$(17) \qquad \textstyle\sum_k = \sum_{1 \le i_1 < \cdots < i_s \le n} (x_{i_1} + x_{i_2} + \cdots + x_{i_s})^k = \frac{1}{s!} \sum_{D(s)} (x_{i_1} + \cdots + x_{i_s})^k$$

where by $D(t)$ is meant summation over all sets of subscripts i_j at least t of which are distinct. Hence

$$s! \textstyle\sum_k = \sum_{D(s-1)} (x_{i_1} + \cdots + x_{i_s})^k - \binom{s}{2} \sum_{D(s-1)} (2x_{i_1} + x_{i_2} + \cdots + x_{i_{s-1}})^k$$

$$= \textstyle\sum_{D(s-2)} (x_{i_1} + \cdots + x_{i_s})^k - \binom{s}{2} \sum_{D(s-2)} (2x_{i_1} + x_{i_2} + \cdots + x_{i_{s-1}})^k$$

$$+ 2\binom{s}{3} \textstyle\sum_{D(s-2)} (3x_{i_1} + x_{i_2} \cdots + x_{i_{s-2}})^k + 3\binom{s}{4} \sum_{D(s-2)} (2x_{i_1} + 2x_{i_2} + x_{i_3} + \cdots + x_{i_{s-2}})^k .$$

Continue cancelling terms until each summation is over $D(1)$. The coefficient of $\sum(m_1 x_{i_1} + \cdots + m_t x_{i_t})^k$ is just $(-1)^{s-t}$ times the number of permutations on s marks which are conjugate to one having cycles of length $m_1, \cdots, m_t$. This can be shown by a method quite similar to that used by Frobenius [1]. Hence we may write

$$(18) \qquad s! \textstyle\sum_k = \sum_P (-1)^{s-t} \sum_{D(1)} (m_1 x_{i_1} + \cdots + m_t x_{i_t})^k$$

where the outer summation is over all permutations P on s marks, and $m_1, \cdots, m_t$ are the lengths of the cycles of P. Now from the multinomial expansion we have

$$\sum_{D(1)} (m_1 x_{i_1} + \cdots + m_t x_{i_t})^k = \sum_{\substack{l_1 + \cdots + l_t = k \\ l_i \ge 0}} \frac{k!}{l_1! \cdots l_t!} m_1^{l_1} \cdots m_t^{l_t} S_{l_1} \cdots S_{l_t}$$

and the coefficient of S_k is $(m_1^k + \cdots + m_t^k) S_0^{t-1}$. Substituting in (18) and using (15), we obtain

$$(19) \qquad (s-1)! \textstyle\sum_k = f(n, k) S_k + \cdots$$

where the terms indicated by dots do not involve S_k. Thus, if $f(n, k) \neq 0$ for $k = 1, \cdots, n$, then (19) can be solved recursively for $S_1, \cdots, S_n$ in terms of $\sum_1, \cdots, \sum_n$.

On the other hand, if $f(n, k) = 0$ and $f(n, j) \neq 0$ for $j = 1, \cdots, k-1$ then (17) expresses $\sum_k$ as a polynomial in $S_1, \cdots, S_{k-1}$ which in turn are polynomials in $\sum_1, \cdots, \sum_{k-1}$.

COROLLARY. *If* $f(n, k) = 0$ *then* n *divides* $(s-1)!\, s^{n-1}$.

Thus $\{x\}$ will always be determined by $\{\sigma\}$ if s is less then the greatest prime factor of n.

EXAMPLE 1. $s = 3$. Here (18) becomes

$$6 \textstyle\sum_k = \sum_{i_1, i_2, i_3 = 1}^{n} (x_{i_1} + x_{i_2} + x_{i_3})^k - 3 \sum_{i_1, i_2 = 1}^{n} (2x_{i_1} + x_{i_2})^k + 2 \sum_{i=1}^{n} (3x_i)^k .$$

Expanding and collecting the coefficient of S_k, we get

$$f(n, k) = n^2 - (2^k + 1)n + 2 \cdot 3^{k-1} .$$

This has obvious zeros at $n = 1$, $k = 1$; $n = 2$, $k = 1, 2$; $n = 3$, $k = 2, 3$. Also, as we know from Theorem 3, there are zeros at $n = 6$, $k = 3, 5$. For all these values of n the set $\{\sigma\}$ does not, in general, determine $\{x\}$ uniquely.

In addition we observe that $f(n, k) = 0$ has the solutions $n = 27$, $k = 5, 9$ and $n = 486$, $k = 9$. We do not know whether for these values of n the set $\{\sigma\}$ determines $\{x\}$ uniquely or not. However we do know that these are the only cases left in doubt.

THEOREM 5. *If* $s = 3$ *then* $f(n, k) = 0$ *has solutions only for* $k = 1, 2, 3, 5, 9$.

Proof. If $f(n, k) = 0$ then

(20) $$n = 2^a \cdot 3^b \text{ with } a = 0 \text{ or } 1 .$$

Substituting (20) in $f(n, k) = 0$ we obtain

(21) $$2^a \cdot 3^b + 2^{1-a}\, 3^{k-b-1} = 2^k + 1 .$$

Let n be the smaller zero of $f(n, k)$ for a fixed k. Then the other zero is $n' = 2^{1-a}\, 3^{k-b-1}$ and $b \leq k - b - 1$. Hence

(22) $$2^k \equiv -1 \pmod{3^b}$$

and since 2 is a primitive root of 3^b,

(23) $$k \equiv 3^{b-1} \pmod{2 \cdot 3^{b-1}} .$$

But by (21) we have

$$3^{k-b-1} \leq 2^k < 3^{2k/3} \text{ or } k < 3(b+1)$$

so that

$$3^{b-1} \leq k < 3(b+1) \text{ and hence } b < 4 .$$

If $b = 3$ then $k \equiv 9 \pmod{18}$ and $k < 12$ so $k = 9$.
If $b = 2$ then $k \equiv 3 \pmod 6$ and $k < 9$ so $k = 3$.
If $b = 1$ then $k \equiv 1 \pmod 2$ and $k < 6$ so $k = 1, 3, 5$.
If $b = 0$ then $k < 3$.

EXAMPLE 2. $s = 4$. Here (18) becomes

$$\begin{aligned}24 \textstyle\sum_k = &\sum_{i_1, i_2, i_3, i_4} (x_{i_1} + x_{i_2} + x_{i_3} + x_{i_4})^k \\ &- 6 \sum_{i_1, i_2, i_3} (2x_{i_1} + x_{i_2} + x_{i_3})^k + 8 \sum_{i_1, i_2} (3x_{i_1} + x_{i_2})^k \\ &+ 3 \sum_{i_1, i_2} (2x_{i_1} + 2x_{i_2})^k - 6 \sum_i (4x_i)^k .\end{aligned} \tag{24}$$

Hence $f(n, k) = 0$ becomes

$$n^3 - 3(2^{k-1} + 1)n^2 + (2(3^k + 1) + 3 \cdot 2^{k-1})n - 3 \cdot 2^{2k-1} = 0 . \tag{25}$$

We first note that this has solutions $n = 1, k = 1$; $n = 2, k = 1, 2$; $n = 3, k = 1, 2, 3$; $n = 4, k = 2, 3, 4$; $n = 8, k = 3, 5, 7$. For these values of n, the set $\{\sigma\}$ does not generally determine $\{x\}$. When $n = 12, k = 6$ is a solution, and this case is left in doubt.

THEOREM 6. *If $s = 4$ then $f(n, k) = 0$ has solutions only for $n =$ 1, 2, 3, 4, 8, 12.*

Proof. Let $n = 3^a \cdot 2^b$ where $a = 0$ or 1. Now if $n \geq 3(2^{k-1} + 1)$ then $2 \cdot 3^k n > 3^{k+1} \cdot 2^k > 3 \cdot 2^{2k-1}$ and the left side of (25) is positive. Hence $n < 3(2^{k-1} + 1) < 2^{k+1}$ if $k > 3$ and so $b \leq k$. (For $k \leq 3$ we have listed all solutions of (25)). If k is even then $2(3^k + 1) \equiv 4 \pmod 8$ and if $k \geq 4$ then $8n$ divides the other terms unless $b \leq 2$. Similarly if k is odd then $2(3^k + 1) \equiv 8 \pmod{16}$ and if $k \geq 5$ then $b \leq 3$. So $b \leq 3$ in all cases. Now suppose $a = 1$. Then (25) becomes

$$2n - 3 \cdot 2^{2k-1} \equiv 0 \pmod 9$$

or

$$2^{b+1} \equiv 2^{2k-1} \equiv 2 \pmod 3$$

and b is even. Thus n must be 1, 2, 3, 4, 8, or 12. It is easy to check that none of these is a root for $k > 7$.

The corollary to Theorem 4 shows that exceptional pairs (s, n) are in a certain sense quite rare. Of course it is trivial to remark that if (s, n) is exceptional, then $(n - s, n)$ is exceptional. Hence the remarks for $s = 2$ apply equally well to $s = n - 2$ and we obtain the exceptional pairs (6, 8), (14, 16), (30, 32), $\cdots$. But there are other cases with $n > 2s$ which our method leaves in doubt.

THEOREM 7. *We can construct arbitrarily large values of s such that* $f(n, k) = 0$ *for some* $n > 2s$.

Proof. If $n < s$ then $\sum_k = 0$ but $S_1, \cdots, S_n$ may be prescribed arbitrarily. Hence the coefficient of S_k in the expansion of $\sum_k$ must be zero if $k \leq n$. If $n = s$ then $\sum_k = S_1^k$ but $S_2, \cdots, S_n$ may be prescribed arbitrarily. Hence $n = s$ is a zero of $f(n, k)$ for $k = 2, \cdots, n$. Thus $f(n, 1) = \prod_{i=1}^{s-1} (n - i)$; $f(n, 2) = \prod_{i=2}^{s} (n - i)$ and $f(n, 3) = (n - 2s) \prod_{i=3}^{s} (n - i)$. If we divide $f(n, 4)$ by its known factors then we obtain for $s > 2$

$$f(n, 4) = [n^2 - (6s - 1)n + 6s^2] \prod_{i=4}^{s} (n - i) \tag{26}$$

and the equation

$$n^2 - (6s - 1)n + 6s^2 = 0 \tag{27}$$

can be rewritten

$$(2n - 6s + 1)^2 - 3(2s - 1)^2 = -2 .$$

The Pell equation $u^2 - 3v^2 = -2$ has the general solution

$$u + v\sqrt{3} = \pm (1 + \sqrt{3})(2 + \sqrt{3})^r \qquad r = 0, \pm 1, \cdots .$$

Since u and v are odd, n and s are integers. It is interesting that all positive solutions are obtained in the following simple way. When $k = 4$, $(s, n) = (2, 8)$ is a solution. Hence $(6, 8)$ is a solution and putting $s = 6$ in (27) yields $(6, 27)$. Continuing in this way, we obtain $(21, 27)$, $(21, 98)$, $(77, 98)$, $(77, 363)$, $\cdots$.

In a similar manner we obtain for $s > 3$

$$f(n, 5) = [n^2 - (12s - 5)n + 12s^2](n - 2s) \prod_{i=5}^{s} (n - i) \tag{28}$$

and all integer roots of the quadratic factor may be obtained with the aid of the general solution of the Pell equation $u^2 - 6v^2 = 75$. Or we could start with $(2, 16)$ and obtain successively $(14, 147)$, $(133, 1444)$, $\cdots$. Starting with $(3, 27)$ yields $(24, 256)$, $(232, 2523)$, $\cdots$.

4. Concluding remarks. If we let $\{\tau\} = \{\tau_1, \cdots, \tau_{ns}\}$ be the set of sums of s not necessarily distinct elements of $\{x\}$, then $\{x\}$ is always determined by $\{\tau\}$. A method similar to the proof of Theorem 4 applies with the coefficient of S_k always positive. Alternatively, if the x_i are real, $x_1 \leq x_2 \leq \cdots \leq x_n$, we may determine them successively by a simple induction procedure.

Our method is applicable to the case of weighted sums $\sigma_{i_1 \cdots i_s} =$

$\sum_{j=1}^{s} a_j x_{i_j}$. The resulting Diophantine equations will however be of a rather different nature. Thus, if the a_j are all distinct then the analogue to $f(n, k) = 0$ is

$$(a_1^k + a_2^k + \cdots + a_s^k) n^{s-1} = 0 . \tag{29}$$

In other words the uniqueness condition is independent of n and depends on the a_i alone. For example if $a_1 + a_2 + \cdots + a_s = 0$ then $\{\sigma\}$ remains unchanged if we add the same constant to all x. It is not as easy to see what happens if (29) holds for some $k > 1$.

REFERENCES

1. G. Frobenius, *Ueber die Charaktere der symmetrischen Gruppe*, S.-B. Preuss. Akad. Wiss. Berlin, (1900), 516-534.
2. L. Moser, Problem E 1248, Amer. Math. Monthly (1957), 507.

UNIVERSITY OF CALIFORNIA, LOS ANGELES

A GENERAL SOLUTION FOR A CLASS OF APPROXIMATION PROBLEMS

ANNETTE SINCLAIR

1. Introduction. This paper generalizes a class of theorems showing the existence of an approximating function which may be required to satisfy certain auxiliary conditions.

Various theorems in analytic function theory which prove the existence of a function fulfilling specified conditions in an open set R have been proved by using a method of the following type. The set R is covered by an increasing sequence of sets $\{R_i\}$. Then the existence of a convergent sequence of functions $\{f_i(z)\}$ is shown such that each $f_n(z)$ behaves properly in R_n and such that $\{f_i(z)\}$ converges to a function satisfying the required conditions everywhere in R. Examples of theorems in which such a method of proof can be applied are furnished by the Mittag-Leffler Theorem, the Carleman Approximation Theorem [1], some rate of growth theorems proved by P. W. Ketchum [2], and the author's generalization of Runge's Theorem [5]. W. Kaplan considered certain problems of this type and remarked, [1], that Brelot has pointed out that this type of proof is valid for approximation to a function $Q(x_1, x_2 \cdots, x_n)$ continuous for all $(x_1, x_2, \cdots, x_n)$ by a function $u(x_1, x_2, \cdots, x_{n+1})$ harmonic for all $(x_1, x_2, \cdots, x_{n+1})$.

The present paper attempts to give an abstract solution for this general class of problems. Examples are also given of some new results obtainable by applying Theorem 1 and fundamental approximation theorems.

In Theorem 3 approximation by an analytic function is considered on a point set S consisting of an infinite number of circular discs tangent on the real axis. It is shown that a function $w(z)$ analytic at interior points of S and continuous on the closure of any finite number of the circular regions—hence, continuous at their points of tangency—can be approximated by an integral function $f(z)$. Moreover, $f(z)$ can be chosen so that the approximation is stronger than uniform approximation—so that corresponding to any $\{\varepsilon_i\}$ there exists $f(z)$ such that

$$|f(z) - w(z)| < \varepsilon_i \text{ on } S_i,$$

where S_i is the ith. circular region.

Theorem 2 combines some previously obtained results [5] by requiring that certain auxiliary conditions be satisfied simultaneously.

Received October 25, 1956.

Theorem 1 can sometimes also be used to show that when an approximation is known to be impossible in the infinite case that analogous results cannot hold in the finite case. An example of this usage is given.

In Part II a topological abstraction is made of Theorem 1 in which the sets of functions are topologized and the system so obtained interpreted as an inverse mapping system. It is then shown that Theorem 1 can be regarded as a special case of Theorem 5.

PART I

2. Fundamental theorem. Let R be an open subset of a topological space. A sequence of sets $\{R_i\}$ which satisfy the following conditions will be called an *increasing sequence of R-covering sets.*

(1) $R_i \subset R$;

(2) $\overline{R}_i$ *is interior to* R_{i+1};

(3) $\bigcup_{i=1}^{\infty} R_i = R$.

$W_1 \cup W_2 \cup \cdots$ such that $W_k \cap W_m = \phi$ for $k \neq m$ is said to be a *decomposition* of a set S if $S = W_1 \cup W_2 \cup \cdots$. An *R-covering sequence* $\{R_i\}$ for R *and a decomposition* $W_1 \cup W_2 \cup \cdots$ of a set $S \subset R$ are said to *correspond* if, for every n, $W_n \subset R_n$, but $W_{n+1} \cap R_n = \phi$.

For a given set $S \subset R$ suppose that an increasing sequence $\{R_i\}$ of R-covering sets and a decomposition $W_1 \cup W_2 \cup \cdots$ of S correspond. Let there be defined classes of functions $\mathscr{W}_n$ and $\mathscr{R}_n$ transforming W_n and R_n respectively into the complex plane, $n = 1, 2, \cdots$. Suppose that each function of $\mathscr{R}_n$ defines a function of $\mathscr{R}_{n-1}$, $n = 2, 3, \cdots$.

THEOREM 1. *Let* $S, R, R_n, W_n, \mathscr{R}_n$, *and* $\mathscr{W}_n$, $n = 1, 2, \cdots$, *be defined as above. Suppose that*

(1) *If* $\{g_i(X)\}$ *is a sequence of functions of* $\mathscr{R}_{n+1}$ *which converges on* R_{n+1} *and uniformly on any closed subset of* R_{n+1}, $\lim_{i\to\infty} g_i(X)$ *defines a function of* $\mathscr{R}_n$;

and (2) *Any function defined on* R_n *by an arbitrary function of the class* $\mathscr{R}_n$ *and on* W_{n+1} *by a function of* $\mathscr{W}_{n+1}$ *can be uniformly approximated arbitrarily closely on* $R_n \cup W_{n+1}$ *by a function of* $\mathscr{R}_{n+1}$, $n = 0, 1, 2, \cdots$ (*where* R_0 *is the null set.*).

Let $w(X)$ *be a function defined on* S *in such a way as to determine a function of* $\mathscr{W}_i$ *for each* i. *Then, corresponding to any* $\{\varepsilon_i\}$, *there exists* $r(X)$ *defined on* R *which determines a function of* $\mathscr{R}_n$ *for each* n *such that*

$$|r(X) - w(X)| < \varepsilon_i \text{ when } X \in W_i, \qquad i = 1, 2, \cdots.$$

Proof. Suppose $\{\varepsilon_i\}$ and $w(X)$ preassigned. When n is taken as 0, (2) implies the existence of $r_1(X)$ of $\mathscr{R}_1$ such that

$$|r_1(X) - w(X)| < \varepsilon_1/2^2 \qquad \text{when } X \in W_1 .$$

In general, for $n = 1, 2, \cdots$, choose $r_n(X)$ of $\mathscr{R}_n$ so that

$$|r_n(X) - r_{n-1}(X)| < \frac{\varepsilon^{(n)}}{2^{n+1}} \qquad \text{on } R_{n-1}$$

and

$$|r_n(X) - w(X)| < \frac{\varepsilon^{(n)}}{2^{n+1}} \qquad \text{on } W_n ,$$

where $\varepsilon^{(n)} = \min_\infty \{\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\}$.

Since $\{r_i(X)\}_{i=k+1}$ converges in R_{k+1}, for arbitrary k, and uniformly on any closed subset of R_{k+1}, it follows from (1) that $\lim_{i\to\infty} r_i(X) = r(X)$ defines a function of $\mathscr{R}_k$ for $k = 1, 2, \cdots$.

It remains to show that $r(X)$ satisfies the assigned approximation conditions. For any k, there exists $m > k$ so that

$$|r(X) - r_m(X)| < \frac{\varepsilon^{(k)}}{2^k} \qquad \text{when } X \in \overline{R}_k .$$

Now

$$|r(X) - w(X)| \leq |r(X) - r_m(X)| + |r_m(X) - w(X)|$$
$$\leq |r(X) - r_m(X)| + \sum_{j=k+1}^{m} |r_j(X) - r_{j-1}(X)| + |r_k(X) - w(X)|$$
$$< \frac{\varepsilon^{(k)}}{2^k} + \sum_{j=k+1}^{m} \frac{\varepsilon^{(j)}}{2^{j+1}} + \frac{\varepsilon^{(k)}}{2^{k+1}} \qquad \text{when } X \in W_k \subset R_k .$$

Thus, $|r(X) - w(X)| < \varepsilon^{(k)} \leqq \varepsilon_k$ when $X \in W_k$. This completes the proof of the theorem.

3. Applications to specific problems. In Theorem 2 we consider approximation on a Q-set of the complex plane having an infinite number of components. A set S is a *Q-set* if its component are closed and its set of *sequential limit points* lie in $C(S)$, the complement of S. (A *sequential limit point of* S is a limit point of a set of points chosen one from each component of S. We note incidentally that a Q-set in the complex plane has at most a denumerable number of components and that its set of sequential limit points may separate the plane [5].)

A set in $C(S)$ is called a *$B^*(S)$-set* if it contains the set B of sequential limit points of S and exactly one point of each component $I_k(S)$ of $C(S)$ such that $I_k(S) \cap B = \phi$.

The author has shown [5] that if S is any Q-set of the complex plane and B^* any $B^*(S)$-set there exists an increasing sequence $\{R_i\}$ of closed covering sets for $C(B)$ such that

(1) If S_t is any component of S and if $S_t \cap R_n \neq \phi$, then $S_t \subset R_n$;

(2) If $I_r(R_n, S)$ is any component of $C(R_n \cup S)$,

$$I_r(R_n, S) \cap B^* \neq \phi .$$

When we set $W_i = S \cap R_i \cap C(R_{i-1})$, we obtain a decomposition $W_1 \cup W_2 \cup \cdots$ of S which corresponds to the increasing sequence $\{R_i\}$ of covering sets for $C(B)$.

A function is *meromorphic on a set* if it is single-valued and analytic in a neighborhood of each point of the set except for poles.

THEOREM 2. *Suppose S is a Q-set, B its set of sequential limit points, and B^* any $B^*(S)$-set. Let $\{R_i\}$ be an increasing sequence of covering sets for $C(B)$ as described above which determines the corresponding decomposition $W_1 \cup W_2 \cup \cdots$ of S. Suppose $w(z)$ is meromorphic on S and that I denotes the set of points of S at which $w(z)$ has poles. Then, corresponding to any sequence $\{\varepsilon_i\}$ of positive constants, there exists $r(z)$ meromorphic in $C(B)$ and analytic in $C(B^* \cup I)$ such that*

$$|r(z) - w(z)| < \varepsilon_i \text{ when } z \in (W_i - I), \qquad i = 1, 2, \cdots$$

It can be required that

(1) The poles of $r(z)$ at points of I have the same principal parts as $w(z)$;

and (2) If K is an isolated interior subset of S such that $K \cap I = \phi$, $r(z)$ can be chosen so that $r(k) = w(k)$ at each point k of K. If B^* has no limit point on S, $r(z)$ can be required to have the same multiplicities at points of K as $w(z)$.

Proof. Define $\mathscr{W}_i$ as the set of those functions meromorphic on W_i and analytic on $(W_i - I)$ which have poles with the same principal parts as $w(z)$ on $(I \cap W_i)$ and k-points with the same multiplicities as $w(z)$ on $(K \cap W_i)$. In $\mathscr{R}_i$ include those functions meromorphic on R_i and analytic in $R_i - (I \cup B^*)$ which have poles with the same principal parts as $w(z)$ on $(I \cap R_i)$ and k-points with the same multiplicities as $w(z)$ on $(K \cap R_i)$, also those functions which are identically constant on a component of R_i which contains no point of I.

Suppose $\{g_i(z)\}$ is a sequence of functions of $\mathscr{R}_{n+1}$ which converges in R_{n+1} and uniformly on any closed subset of R_{n+1} (where any points of $(I \cup B^*)$ are deleted from a closed subset which contains them). Then $\lim g_i(z)$ is meromorphic on R_n and analytic in $R_n - (I \cup B^*)$ with poles and k-points identical with those of $w(z)$ at points of I and K, except that $\lim g_i(z)$ may be identically constant on a component of R_n which contains no points of I. Thus, $\lim g_i(z) \in \mathscr{R}_n$ and (1) of Theorem 1 is satisfied.

Before applying Theorem 1 it remains to show that for any $g(z)$ of $\mathscr{R}_n$ and $v(z)$ of $\mathscr{W}_{n+1}$, corresponding to arbitrary $\varepsilon > 0$, there exists $f(z)$ of $\mathscr{R}_{n+1}$ such that

$$|f(z) - g(z)| < \varepsilon \text{ when } z \in R_n,$$
$$\text{and } |f(z) - v(z)| < \varepsilon \text{ when } z \in W_{n+1} .$$

This follows from Walsh's generalization of Runge's Theorem [7, p. 15] and from another theorem of Walsh [7, p. 313] after it is noted that a finite number of poles on $R_n \cup W_{n+1}$ cause no real difficulty. Just apply the general Mittag-Leffler Theorem [4] to show the existence of a function $h(z)$ meromorphic in $C(B)$ whose poles coincide with those of $g(z)$ and $v(z)$ on R_n and W_{n+1} respectively with the same principal parts. Then define $F(z) = \begin{cases} g(z) - h(z) \text{ on } R_n. \\ v(z) - h(z) \text{ on } W_{n+1}. \end{cases}$

Since $F(z)$ is analytic on $R_n \cup W_{n+1}$, by Walsh's generalization of Runge's Theorem [7, p. 15], there exists a rational function $k(z)$ whose poles lie in B^* such that $|F(z) - k(z)| < \varepsilon$ when $z \in (R_n \cup W_{n+1})$. Another theorem by Walsh [7, p. 313] implies that $k(z)$ can be chosen so that $k(z) = F(z)$ at points of K and so that $k(z)$ has the same multiplicities at these points as $F(z)$. Set $f(z) = h(z) + k(z)$.

Now $f(z)$ is meromorphic on R_{n+1} and its poles on R_{n+1} lie at points of $I \cup (R_{n+1} \cap B^*)$ with those on I having the same principal parts as $h(z)$, hence as $g(z)$ or $v(z)$, and so the same as $w(z)$. Also

$$\begin{aligned} |g(z) - f(z)| &= |[g(z) - h(z)] - k(z)| \\ &= |F(z) - k(z)| < \varepsilon \text{ when } z \in R_n \end{aligned}$$

and similarly $|v(z) - f(z)| = |F(z) - k(z)| < \varepsilon$ when $z \in W_{n+1}$. Since $k(z) = F(z)$ at points of K,

$$\begin{aligned} f(z) &= h(z) + k(z) = h(z) + F(z) = h(z) + g(z) - h(z) \\ &= g(z) \text{ on } R_n \cap K \end{aligned}$$

and, similarly, $v(z)$ on $W_{n+1} \cap K$.

This completes the proof that the hypothesis of Theorem 1 is satisfied. Hence, by Theorem 1, there is a function $r(z)$ defined on $C(B)$ (where ∞ is allowed as a functional value) which determines a function of $\mathscr{R}_n$ for each n such that

$$|r(z) - w(z)| < \varepsilon_i \text{ when } z \in W_i, \qquad i = 1, 2, \cdots .$$

Thus, $r(z)$ is meromorphic in $C(B)$, analytic in $C(I \cup B^*)$, and has poles and k-points of $w(z)$ on S as specified and also satisfies the required approximation condition. (*In* general, $r(z)$ is not identically constant on a component of $C(B)$.)

In Theorem 3 S consists of circular discs tangent on the real axis. More precisely, let $W_i = \{z / |z - i| \leq 1/2\}$, except that $z = i - 1/2$ is deleted, and define S as $\bigcup_{i=1}^{\infty} W_i$. Set $R_i = \{z / |z| \leq i + 1/2\}$ and let R be the finite plane. Then $\{R_i\}$ and the decomposition $W_1 \cup W_2 \cup \cdots$ of S correspond.

THEOREM 3. *Suppose S defined as in the preceding paragraph. Let $w(z)$ be any function analytic at interior points of S and continuous on the boundary except at infinity. Then, corresponding to any $\{\varepsilon_i\}$, there exists an integral function $r(z)$, such that $|r(z) - w(z)| < \varepsilon_i$ when $z \in W_i$ and $r(i + 1/2) = w(i + 1/2), i = 1, 2, \cdots$.*

Proof. Let $\mathscr{R}_i$ be the set of all functions $f(z)$ analytic on R_i such that $f(k + 1/2) = w(k + 1/2)$ for $k = 1, 2, \cdots, i$. Let $\mathscr{W}_i$ be the set of all functions $f(z)$ analytic at interior points of W_i and continuous on W_i such that $f(i + 1/2) = w(i + 1/2)$ and $\lim_{z \to i} f(z - 1/2) = w(i - 1/2)$, $(z - 1/2) \in W_i \, i = 1, 2, \cdots$.

If $\{g_i(z)\}$, where $g_i(z)$ is a member of $\mathscr{R}_{n+1}$, converges on R_{n+1}, uniformly on any closed subset of R_{n+1}, $\lim g_i(z)$ gives a function of $\mathscr{R}_n$,

By a theorem of Walsh [7, p. 47] a function $g(z)$ analytic interior to and continuous on a closed set C which does not separate the plane and which is bounded by a finite number of Jordan curves, as is the case if $C = R_n \cup W_{n+1}$, can be uniformly approximated on C by a polynomial $p(z)$. Then by another theorem of Walsh [7, p. 310], $p(z)$ can be chosen so that $p(k + 1/2) = g(k + 1/2)$, $k = 1, 2, \cdots, n + 1$. If $g(k + 1/2) = w(k + 1/2)$ then $p(z) \in \mathscr{R}_{n+1}$. Thus, the hypothesis of Theorem 1 is satisfied and the required conclusion follows.

The next theorem is an extension of the Carleman approximation theorem in that values of ahe approximating function are preassigned at certain points.

THEOREM 4. (*Carleman Approximation Theorem*). *Let $w(x)$ be a continuous complex-valued function of x for $-\infty < x < \infty$. Then, corresponding to any $\{\varepsilon_i\}$, there exists an integral function $f(z)$ such that*

$$|f(x) - w(x)| < \varepsilon_i \text{ when } i - 1 < |x| \leq i, \qquad i = 1, 2, \cdots,$$

and such that $f(i) = w(i), i = \pm 1, \pm 2, \cdots$.

Proof. The proof is like that of Theorem 3 when W_i is defined as $\{x/i - 1 < |x| \leq i\}$ (except that W_1 also includes the origin); R_i as $\{z/|z| \leq i\}$; R as the finite plane; $\mathscr{W}_i$ as those functions continuous on W_i such that $f(\pm i) = w(\pm i)$ and $\lim_{x \to (\pm i \mp 1)} f(x) = w(\pm i \mp 1)$; and $\mathscr{R}_i$ as those functions $f(z)$ analytic on R_i such that $f(k) = w(k)$, $k = \pm 1, \pm 2, \cdots, \pm i$.

Theorem 1 can sometimes be used to show that certain requirements on the approximating function cannot, in general, be made, even when the approximation is on a set having only a finite number of components. Next an application of this type is indicated.

When approximating a function analytic and simple on each component of a closed set C by a function $f(z)$ analytic in a preassigned finite region D containing C, one cannot, in general, require that $f(z)$ be simple in D. To verify this we consider a Q-set S whose components are simply connected and which has infinity as its only sequential limit point. Suppose $\{R_i\}$ is an increasing sequence of R-covering sets, as described for Theorem 2, which gives the corresponding decomposition $W_1 \cup W_2 \cup \cdots$ of S. Let $\mathscr{R}_i$ (and $\mathscr{W}_i$) consist of all functions analytic and simple on $R_i(W_i)$, also all constants. We note that (1) of Theorem 1 is satisfied [6, p, 203]. If (2) were also satisfied, Theorem 1 would imply that arbitrary $w(z)$ simple on S could be approximated on S by a function simple in the whole finite plane. Since $w(z)$ can be chosen so that $f(z)$ would necessarily have an essential singularity at ∞, this does not hold. We conclude that (2) is not, in general, satisfied.

Theorems 2, 3, and 4 and the illustration just stated are examples of some of the applications which can be made of Theorem 1.

PART II

4. Topological abstraction of Theorem 1. Theorem 1 can be interpreted as a density result for a Cartesian product space. The author's original version treated the $\mathscr{R}_i$'s of Theorem 1 with the respective topologies induced by the metrics

$$d_i(f, g) = \sup_{x \in R_i} |f(X) - g(X)|$$

as a *nested sequence of spaces*. The interpretation given in Theorem 5 as an inverse mapping system was suggested by Prof. Hans Samelson of the University of Michigan. In addition to having the advantage of conforming to convention, this formulation applies to classes of functions other than analytic functions.

If $\{W_i\}$ is any sequence of topological spaces, W^∞ denotes the Cartesian product space $W_1 \times W_2 \times \cdots$. We shall be concerned with the *box topology* for W^∞ in which a neighborhood of $w = (w_1, w_2, \cdots)$ is defined as $\mathrm{N}_{w_1}(W_1) \times \mathrm{N}_{w_2}(W_2) \times \cdots$.

If $\{R_i\}_{i=1}^{\infty}$ is a denumerable system of T_2-spaces and if for $n = 2,3,\cdots$, there is defined a continuous transformation $\prod_{n-1}^{n}$ of R_n into R_{n-1}, the system $\sum = \{R_i, \prod^i\}$ of the R_i's and $\prod$'s is an *inverse mapping system*, [3, p. 31]. The subset R of $R^{\infty} = R_1 \times R_2 \times \cdots$ of all those points $x = \{x_i\}$ such that $\prod_i^{i+1} x_{i+1} = x_i$ is called the *limit space of the inverse mapping system* $\sum$.

In Theorem 5 we suppose that $R_1, R_2, \cdots$ are given sets and that for each i, and arbitrary points $p, q \in R_i$, there is defined a metric $d_i(p, q)$, where ∞ is allowed as a possible value. Then R_i with the neighborhood system induced by $d_i(p, q)$ is a T_2-space. If, for $i = 2,3,\cdots$, a transformation $\prod_{i-1}^{i}$ of R_i into R_{i-1} is defined which is a contraction (that is, $d_{i-1}(\prod_{i-1}^{i} p, \prod_{i-1}^{i} q) \leq d_i(p, q)$), then the $\prod$'s are continuous and $\{R_i, \prod_{i-1}^{i}\}$ is an inverse mapping system.

Before stating Theorem 5 we note that the R_i's of this theorem are analogous to the $\mathscr{R}_i$'s and the W_i's to the $\mathscr{W}_i$'s of Theorem 1.

THEOREM 5. *Let $\{W_i\}_{i=1}^{\infty}$ be a system of topological spaces and let $\{R_i, \prod^i\}$ be an inverse mapping system as described in the preceding paragraphs. Suppose that for each i there is defined a continuous transformation f_i which maps R_i into W_i. Suppose also that the following conditions are satisfied:*

(1) *If $\{p_j^{(n)}\}_{j=1}^{\infty}$ is a Cauchy sequence in R_n, its image $\{\prod_{n-1}^{n} p_j^{(n)}\}_{j=1}^{\infty}$ is convergent in R_{n-1};*

(2) *$f_1(R_1)$ is dense in W_1 and, when $n > 1$, $\prod_{n-1}^{n} \times f_n(R_n)$ is dense in $R_{n-1} \times W_n$.*

Then under the transformation $\{\times f_i\}$ the image of the limit space R of the inverse mapping system $\sum$ is dense in W^{∞} by the box topology.

Proof. Let $w = (w_1, w_2, \cdots)$ be any point of W^{∞} and let $\mathrm{N}_w = \mathrm{N}_{w_1}(W_1) \times \mathrm{N}_{w_2}(W_2) \times \cdots$ be an arbitrary neighborhood of w.

Since $f_1(R_1)$ is dense in W_1, there is a point $r_1 \in R_1$ such that $f_1(r_1) \in \mathrm{N}_{w_1}(W_1)$. There exists $N_{r_1}(R_1) \subset N_{r_1}^{(1)}(R_1)$, where $N_{r_1}^{(\alpha)} = \{p \in R_1 / d_1(p, r_1) < \alpha\}$. Since f_1 is continuous and R_1 is regular, we can suppose $N_{r_1}(R_1)$ chosen so that $f_1(\overline{N}_{r_1}) \subset N_{w_1}(W_1)$.

In general, since $\prod_{n-1}^{n} \times f_n(R_n)$ is dense in $R_{n-1} \times W_n$, there exists $r_n \in R_n$ so that $\prod_{n-1}^{n} \times f_n(r_n) \in N_{r_{n-1}}(R_{n-1}) \times N_{w_n}(W_n)$. There exists $N_{r_n}(R_n) \subset N_{r_n}^{(1/2^{n-1})}(R_n)$. Since f_n and $\prod_{n-1}^{n}$ are continuous and R_n is regular, we can suppose $N_{r_n}(R_n)$ chosen so that $N_{r_n}(R_n) \subset N_{r_n}^{(1/2^{n-1})}(R_n)$ and so that

$$\prod_{n-1}^{n} \times f_n(\overline{N}_{r_n}) \subset N_{r_{n-1}}(R_{n-1}) \times N_{w_n}(W_n) .$$

The sequence $\{\prod_n^{n+i} r_{n+i}\}_{i=0}^{\infty}$, where $\prod_n^{n+i} = \prod_n^{n+1} \cdots \prod_{n+i-1}^{n+i}$, is a Cauchy sequence in R_n. For, corresponding to any $\varepsilon > 0$, there exists $m > n$ so that $1/(2^{m-2}) < \varepsilon$; then

$$d_n(\textstyle\prod_n^{m+k} r_{m+k}, \prod_n^m r_m) \le \sum_{i=0}^{k-1} d_n(\prod_n^{m+i+1} r_{m+i+1}, \prod_n^{m+i} r_{m+i})$$
$$\le \sum_{i=0}^{k-1} d_{m+i}(\textstyle\prod_{m+i}^{m+i+1} r_{m+i+1}, r_{m+i}) < \sum_{i=0}^{k-1} \frac{1}{2^{m+i-1}} < \frac{1}{2^{m-2}} < \varepsilon .$$

The first inequality follows from the triangle inequality, the second from the fact that the $\prod$'s are contractions, and the third holds since $\prod_{m+i}^{m+i+1} r_{m+i+1} \in N_{r_{m+i}}(R_{m+1}) \subset N_{r_{m+i}}^{((1/2)m+i-1)}$.

By (1) the image of the Cauchy sequence above is convergent in R_{n-1}. Hence, we let $r^{(n-1)}$ denote $\lim_{i\to\infty} \prod_{n-1}^{n+i} r_{n+i}$ in R_{n-1}. Now $r^{(n-1)} \in \overline{N}_{r_{n-1}}(R_{n-1})$ and $f_{n-1}(\overline{N}_{r_{n-1}}(R_{n-1})) \subset N_{w_{n-1}}(W_{n-1})$. Hence, $f_{n-1} r^{(n-1)} \in N_{w_{n-1}}(W_{n-1})$.

To complete the proof of the theorem it is sufficient to show that $\{r^{(i-1)}\}_{i=2}^{\infty}$ belongs to the limit space, that is, that $r^{(n-2)} = \prod_{n-2}^{n-1} r^{(n-1)}$ for $n = 3, 4, \cdots$. Since $\{\prod_{n-1}^{n+i} r_{n+i}\}$ converges to $r^{(n-1)}$ in R_{n-i}, corresponding to any $\delta > 0$, there exists k such that $i > k$ implies $d_{n-1}(\prod_{n-1}^{n+i} r_{n+1}, r^{(n-1)}) < \delta$. Then, since the π's are contractions, $d_{n-2}(\prod_{n-2}^{n+i} r_{n+i}, \prod_{n-2}^{n-1} r^{(n-1)}) < \delta$ for all $i > k$. Now $\lim \prod_{n-2}^{n+i} r_{n+i}$ is unique in R_{n-2}, and so $\prod_{n-2}^{n-1} r^{(n-1)} = \lim_{i\to\infty} \prod_{n-2}^{n+i} r_{n+i} = r^{(n-2)}$. This completes the proof of Theorem 5.

If a function of $\mathscr{R}_n$ defines a function of $\mathscr{W}_n$, $n = 1, 2, \cdots$, Theorem 1 can be obtained from Theorem 5. Since each function of $\mathscr{R}_n$ in Theorem 1 defines a function of $\mathscr{R}_{n-1}$ (and in the case just specified, also $\mathscr{W}_n$), transformations $\prod_{n-1}^n$ and f_n are determined of $\mathscr{R}_n$ into $\mathscr{R}_{n-1}$ and $\mathscr{W}_n$. Let us define a metric $d_n(f, g)$ for each $\mathscr{R}_n$ (also $\mathscr{W}_n$) as $\sup_{X \in R_n (\text{or } W_n)} |f(X) - g(X)|$. Thus, T_2-topologies are determined for $\mathscr{R}_n$ and $\mathscr{W}_n$ respectively. If $f, g \in \mathscr{R}_n$ then,

$$\sup_{X \in R_{n-j} \text{ or } W_n} |f(X) - g(X)| \le \sup_{X \in R_n} |f(X) - g(X)|$$

and so $\prod_{n-1}^n$ and f_n are contractions and hence continuous. We note that $\{\mathscr{R}_i, \prod_{i-1}^i\}$ is an inverse mapping system $\sum$. The hypotheses (1) and (2) of Theorem 1 correspond to (1) and (2) of Theorem 5. By Theorem 5 the image of the limit space $\mathscr{R}$ of the inverse mapping system $\sum$ is dense in $\mathscr{W}^\infty$. This is just a statement that corresponding to any function $w(X)$ which defines a point w of $\mathscr{W}^\infty$ and to any $\{\varepsilon_i\}$ there exists $r(X)$ which determines a function of $\mathscr{R}_n$ for each n such that $|r(X) - w(X)| < \varepsilon_i$ when $X \in W_i$. In this way Theorem 1 can be regarded as a special case of Theorem 5.

References

1. W. Kaplan, *Approximation by entire functions*, Mich. Math. J. **3** (1955-6), 43-52.
2. P. W. Ketchum, *On the possible rate of growth of an analytic function*, Trans. Amer. Math. Soc. **49** (1941), 211-228.
3. S. Lefschetz, *Algebraic topology*, Amer. Math. Soc. Colloquium Publications, vol. 27, New York, 1942.
4. G. Mittag-Leffler, *Sur la représentation analytique des fonctions monogènes uniformes d'une variable indépendante*, Acta. Math. **4** (1884), 1-79.
5. A. Sinclair, *Generalization of Runge's theorem to approximation by analytic functions*, Trans. Amer. Math. Soc. **72** (1952), 148-164.
6. W. J. Thron, *The theory of functions of a complex variable*, John Wiley and Sons, Inc., 1953.
7. J. L. Walsh, *Interpolation and approximation by rational functions in the complex domain*, Amer. Math. Soc. Colloquium Publications, vol. 20, New York, 1935.

Southern Illinois University
Carbondale, Illinois

(C, ∞) AND (H, ∞) METHODS OF SUMMATION

G. SZEKERES AND A. JAKIMOVSKI

1. Introduction. Let

$$T: \qquad \beta_n = \sum_{m=1}^{\infty} \tau_{nm} \cdot \alpha_m , \qquad n = 1, 2, \cdots$$

be a linear transformation of the sequence $\alpha = \{\alpha_n\}$ into the sequence $\beta = \{\beta_n\}$; we write

$$\beta = T\alpha , \quad \beta_n = (\beta)_n = (T\alpha)_n .$$

α is said to be *summable* T *to the value* a if

$$(1) \qquad \lim_{n\to\infty} (T\alpha)_n = a ;$$

and U is said to contain T if every sequence summable T to the value a is also summable U to a. In particular T is called regular if it contains the identity transformation I.

We shall generalize the concept of regularity in several directions. A sequence of transformations $\{T_k\}$ $(k \geqq 0, T_0 = I)$ will be called *regular* if each T_k is included in T_{k+1}. As an example of a regular sequence we mention the iterates of a regular transformation T ; they are defined

$$T^0\alpha = \alpha , \quad T^{k+1}\alpha = T(T^k\alpha) , \qquad k = 0, 1, 2, \cdots .$$

Given a regular sequence of transformations, $\{T_k\}$, we say that α is *summable* T_∞ *to* a if

$$\lim_{k\to\infty} (T_k\alpha)_n = a_n$$

exists for $n \geqq n_0$ and $\lim_{n\to\infty} a_n = a$. For T_∞ to be significant, it is plainly desirable that it shall contain each T_k. A regular sequence $\{T_k\}$ will be called *strongly regular* if, whenever for some k $\lim_{n\to\infty}(T_k\alpha)_n = a$, then α is summable T_∞ to a. With trivial modifications these definitions also apply to families of transformations T_λ which depend on a continuous real parameter $\lambda \geqq 0$.

Of particular interest are sequences of transformations of the type

$$(2) \qquad (T_k\alpha)_n = \frac{(T^k\alpha)_n}{(T^k\varepsilon)_n}$$

where ε is the unit sequence $\varepsilon_n = 1$, $n = 1, 2, \cdots$, and T is a transformation such that $T^k\varepsilon$ exists for $k \geqq 0$. By an extension of the concept of regularity we say that T is *strongly regular* if the particular sequence (2) is strongly regular, and summability T_∞ for this sequence will be denoted

Received Jaunary 27, 1958.

by summability (T, ∞).

To examine the usefulness of these concepts and in particular the possibility of 'infinite iteration', let us consider the Hölder process

$$(H\alpha)_n = \frac{1}{n}\sum_{m=1}^{n}\alpha_m .$$

It follows from

$$(H^{k+1}\alpha)_n = \frac{1}{n}\sum_{m=1}^{n}(H^k\alpha)_m = \frac{n-1}{n}(H^{k+1}\alpha)_{n-1} + \frac{1}{n}(H^k\alpha)_n ,$$

by induction on n, that

$$\lim_{k\to\infty}(H^k\alpha)_n = \alpha_1 ,$$

for every n. Thus we find the disconcerting result that every sequence is summable (H, ∞) to the first term of the sequence.

Similary if $(C\alpha)_n = \sum_{m=1}^{n}\alpha_m$, we have

$$(C^k\alpha)_n = \sum_{m=1}^{n}\binom{n-m+k-1}{n-m}\alpha_m , \quad (C^k\varepsilon)_n = \binom{n+k-1}{n-1} ,$$

and hence

$$\lim_{k\to\infty}\frac{(C^k\alpha)_n}{(C^k\varepsilon)_n} = \alpha_1$$

for every n. As in the case of (H, ∞), we find that the (C, ∞) limit always exists and is equal to the first term of the sequence. Thus neither the Hölder, nor the Cesàro process is strongly regular, and infinite iteration gives nothing useful.

In the next section we shall reconsider the problem of (H, ∞) and (C, ∞) from the point of view of generalized limits of functions; this will lead quite naturally to strongly regular transformations. Here we mention an interesting example of a strongly regular family of transformations, known as the 'circle methods' of Hardy and Littlewood. For $\lambda > 0$ define

$$(3) \qquad (T_\lambda\alpha)_n = \rho^{n+1}\sum_{m=n}^{\infty}\binom{m}{n}(1-\rho)^{m-n}\alpha_m , \qquad n \geqq 0 ,$$

where $\rho = e^{-\lambda}$; (3) certainly exists if

$$(4) \qquad \overline{\lim_{n\to\infty}}\,|\alpha_n|^{1/n} \leqq 1 .$$

If α satisfies this condition and $\mu > \lambda \geqq 0$, then T_μ contains T_λ; this follows from the regularity of T_λ and the formula

$$(5) \qquad T_\mu\alpha = T_{\mu-\lambda}(T_\lambda\alpha)$$

which is valid for sequences satisfing (4) (See [4, p. 218]). It is seen directly from the definition of T_λ that if $\lim_{n\to\infty} \alpha_n = a$ then for every fixed $n \geqq 0$,

$$\lim_{\lambda\to\infty} (T_\lambda \alpha)_n = a ;$$

it follows, by (5), that $\lim_{n\to\infty}(T_\lambda\alpha)_n = a$ for some $\lambda > 0$ implies $\lim_{\mu\to\infty}(T_\mu\alpha)_n = a$. Hence T_∞ contains T_λ, and the family T_λ is strongly regular, at least for sequences which satisfy the condition (4).

2. (C, ∞) and (H, ∞) limits. Let us consider the problem of infinite iteration from the point of view of generalized limits of functions. For the sake of definitness we consider limits at $x = \infty$. We say $f(x)$ has a (generalized) limit a at $x = \infty$ by the process

$$(6) \qquad T: \quad Tf(x) = \int_0^\infty \tau(x, t) f(t) dt$$

if $\lim_{x\to\infty} Tf(x)/Te(x) = a$, where $e(x)$ is the unit function $e(x) = 1$ for all $x > 0$; we assume that

$$T^k e(x) = \int_0^\infty \tau(x, t_1)dt_1 \int_0^\infty \tau(t_1, t_2)dt_2 \cdots \int_0^\infty (t_{k-1}, t_k)dt_k$$

exists for every $k > 0$.[1] Integrals of the form $\int_0^\infty \varphi(t)dt$ are understood to be improper Lebesgue integrals in the following sense: $\varphi(t)$ is assumed to be L-integrable in every interval $0 < t_1 \leqq t \leqq t_2 < \infty$ and

$$\int_0^\infty \varphi(t)dt = \lim_{\varepsilon \downarrow 0} \int_\varepsilon^1 \varphi(t)dt + \lim_{x \uparrow \infty} \int_1^x \varphi(t)dt .$$

The *domain* of T is the class of functions $f(x)$ for which the integral (6) exists; but since we are interested in the limit of $f(x)$ when $x \to \infty$, it is convenient also to consider the subclass of these functions in the domain of T for which $f(x) = 0$ for $x < x_0$ (where x_0 is not necessarily the same number for every $f(x)$). This subclass will be called the *essential domain* of T.

The definitions given in § 1 apply equally well to transformations of the form (6); in particular, T is called strongly regular if the sequence

$$(7) \qquad T_k f(x) = \frac{T^k f(x)}{T^k e(x)}, \qquad k \geqq 0,$$

is regular and $\lim_{x\to\infty} T_k f(x) = a$ for some $k \geqq 0$ implies

$$\lim_{x\to\infty} \lim_{k\to\infty} T_k f(x) = a .$$

The natural analogues of the Cesàro and Hölder limits at $x = \infty$ are obtained by the transformations

[1] Unless the contrary is stated, letters $k, m, n, \cdots$ denote non-negative integers.

$$Cf(x) = \int_0^x f(t)dt$$

and

$$Hf(x) = \frac{1}{x}\int_0^x f(t)dt = \int_0^1 f(xt)dt\ ;$$

their domain is the class of functions L-integrable over every interval $0 < t_1 \leqq t \leqq t_2 < \infty$ for which $\int_0^1 f(t)dt$ (as an improper integral) exists. Denote this class by $\Phi_{1,0}$. Clearly $f(x) \in \Phi_{1,0}$ implies the existence and continuity of $Cf(x)$ for $x > 0$ and $\lim_{x\downarrow 0} Cf(x) = 0$; therefore $C^k f(x)$ exists for every $k > 0$. On the other hand $H^k f(x)$ does not always exist when $f(x) \in \Phi_{1,0}$; for example $f(x) = (x \log^2 x)^{-1}$ is in $\Phi_{1,0}$, but not $f(x) = (x \log x)^{-1}$, so that $H^2 f(x)$ does not exist ; We denote by $\Phi_{k,0}$ the class of functions for which $H^k f(x)$ exists ; $\Phi_{\infty,0}$ denotes the intersection of all classes $\Phi_{k,0}$; $\Phi_{0,0}$ denotes the class of functions L-integrable over every interval $0 < t_1 \leqq t \leqq t_2 < \infty$. For later use we also define : $\Phi_{0,m}$, the class of functions $s(x)$ such that $f(x) \equiv s(1/x)$ is in $\Phi_{m,0}$; $\Phi_{k,m} = \Phi_{k,0} \cap \Phi_{0,m}$. If $s(x) \in \Phi_{k,m}$ then $f(x) \equiv s(1/x) \in \Phi_{m,k}$. Finally Φ_I shall denote the class of functions for which $\int_0^\infty f(t)dt$ exists, and Φ_B is the subclass of bounded functions of $\Phi_{0,0}$; clearly Φ_I is a subclass of $\Phi_{1,1}$ and Φ_B is a subclass of $\Phi_{\infty,\infty}$.

The examination of the infinite iteration of the C and H methods for functions leads to a result which is analogous in some respects to the corresponding result for sequences. It turns out that the limit by (C, ∞) or (H, ∞), if it exists at all, depends on the behaviour of the function in the neighbourhood of zero rather than infinity. If in particular $\lim_{x\downarrow 0} f(x) = a$ exists then

$$\lim_{k\to\infty} H_k f(x) = \lim_{k\to\infty} C_k f(x) = a$$

for all $x > 0$. More generally we shall show :

THEOREM 1. *Suppose that* $f(x) \in \Phi_{1,0}$ *and*

$$\lim_{x\downarrow 0} C_k f(x) = a \tag{8}$$

for some $k \geqq 0$. *Then*

$$\lim_{n\to\infty} C_n f(\xi) = a \tag{9}$$

for every fixed $\xi > 0$.

THEOREM 1*. *Suppose that* $f(x) \in \Phi_{k,0}$ *for some* $k \geqq 0$ *and*

$$\lim_{x\downarrow 0} H_k f(x) = a\ , \tag{8*}$$

Then $f(x) \in \Phi_{\infty,0}$ *and*

$$\lim_{n\to\infty} H_n f(\xi) = a \tag{9*}$$

for every fixed $\xi > 0$.

Theorem 1 and Theorem 1* show that although C and H are not strongly regular with respect to $x \uparrow \infty$, they are strongly regular with respect to limits of the form

$$\lim_{x\downarrow 0} C_k f(x) \quad \text{and} \quad \lim_{x\downarrow 0} H_k f(x) . \tag{10}$$

Generalized limits of the type (10) were first considered by Hardy and Littlewood in connection with the summability problem of Fourier series [5]; in the present context they appear as natural extensions of the Cesàro and Hölder processes distinguished by the property that they admit infinite iteration.

Proof of Theorem 1. For $k > 0$ we find by repeated integration by parts

$$C^k f(x) = \frac{1}{(k-1)!}\int_0^x (x-t)^{k-1} f(t)dt , \tag{11}$$

$$C^k e(x) = \frac{1}{k!} x^k , \qquad x > 0 \tag{11*}$$

and

$$C_k f(x) = k\int_0^1 (1-t)^{k-1} f(xt)dt , \qquad k > 0 . \tag{12}$$

The relations (11) and (12) define C^k and C_k also for non-integral $k > 1$. The existence of $C^\sigma f(x)$ for $\sigma > 1$ can be seen from

$$\int_\varepsilon^x (x-t)^{\sigma-1} f(t)dt = [(x-t)^{\sigma-1} Cf(t)]_{t=\varepsilon}^{t=x} - (\sigma-1)\int_\varepsilon^x (x-t)^{\sigma-2} Cf(t)dt ;$$

the expressions on the right have a limit when $\varepsilon \downarrow 0$, since $\lim_{x\downarrow 0} Cf(x) = 0$. Clearly $C^\sigma f(x)$ is continuous for $\sigma \geqq 1$, $x > 0$, and $\lim_{x\downarrow 0} C^\sigma f(x) = 0$. By partial integration we obtain, for fixed $\xi > 0$ and $\sigma > k$ [2]

$$C_\sigma f(\xi) = \frac{\Gamma(\sigma+1)}{\Gamma(k+1)\Gamma(\sigma-k)}\int_0^1 (1-t)^{\sigma-k-1} t^k C_k f(\xi t)dt . \tag{13}$$

This can be regarded (for fixed ξ and k) as a transformation from $C_k f(\xi t)$ to $C_\sigma f(\xi)$; and Theorem 1 is proved if we can show that this transformation is regular. Now regularity follows immediately from a remark to § 3.5(3) in [4, p. 61], since the following three conditions are satisfied:

---

[2] This is a well-known identity; see for example [2, p. 3].

(1) $$\frac{\Gamma(\sigma+1)}{\Gamma(k+1)\Gamma(\sigma-k)}(1-t)^{\sigma-k-1}t^k \geqq 0 \qquad \text{for } \sigma > k,\ 0 \leqq t \leqq 1\,.$$

(2) $$\frac{\Gamma(\sigma+1)}{\Gamma(k+1)\Gamma(\sigma-k)}\int_0^1 (1-t)^{\sigma-k-1}t^k dt = 1\,.$$

(3) For a fixed x, $0 < x < 1$, and a fixed k,

$$\lim_{\sigma\uparrow\infty} \frac{\Gamma(\sigma+1)}{\Gamma(k+1)\Gamma(\sigma-k)}\int_x^1 (1-t)^{\sigma-k-1}t^k dt = 0$$

since

$$\begin{aligned} 0 &\leqq \frac{\Gamma(\sigma+1)}{\Gamma(k+1)\Gamma(\sigma-k)}\int_x^1 (1-t)^{\sigma-k-1}t^k dt \\ &\leqq \frac{\Gamma(\sigma+1)}{\Gamma(k+1)\Gamma(\sigma-k)}\int_x^1 (1-t)^{\sigma-k-1} dt \\ &= \frac{\Gamma(\sigma+1)}{\Gamma(k+1)\Gamma(\rho-k+1)}(1-x)^{\sigma-k} \to 0 \qquad \text{as} \quad \sigma\uparrow\infty\,. \end{aligned}$$

Proof of Theorem 1*. We note first that $H^k e(x) = 1$ for every $k \geqq 0$, $x > 0$, and conditon (8*) implies that $H^k f(x)$ is bounded for $0 < x \leqq M + \infty$. Therefore $H^m f(x)$ exists for $m \geqq k$ and hence $f(x) \in \Phi_{\infty,0}$. Condition (8*) implies

(14) $$H^k f(x) = a + \varphi(x)$$

where $\lim_{x\downarrow 0}\varphi(x) = 0$; we have to show that $\lim_{n\to\infty} H^n\varphi(x) = 0$. For bounded functions repeated partial integration gives

(15) $$H^{k+1}f(x) = \frac{1}{k\,!}\int_0^1 \left(\log\frac{1}{t}\right)^k f(xt)dt\,, \qquad k \geqq 0\,.$$

Thus the statement to be proved is, that for any fixed $\xi > 0$,

(16) $$\lim_{n\uparrow\infty}\int_0^1 \frac{1}{n\,!}\left(\log\frac{1}{t}\right)^n \varphi(\xi t)dt = 0\,.$$

Choose $\delta > 0$ so that $|\varphi(\xi t)| < \varepsilon/2$ for $0 < t < \delta$ and let $|\varphi(\xi t)| < K$ for $0 < t \leqq 1$; then

$$\left|\frac{1}{n\,!}\int_0^\delta \left(\log\frac{1}{t}\right)^n \varphi(\xi t)dt\right| < \frac{1}{2}\,\varepsilon\, \frac{1}{n\,!}\int_0^1 \left(\log\frac{1}{t}\right)^n dt = \frac{1}{2}\,\varepsilon\,,$$

$$\left|\frac{1}{n\,!}\int_\delta^1 \left(\log\frac{1}{t}\right)^n \varphi(\xi t)dt\right| < K\,\frac{1}{n\,!}\left(\log\frac{1}{\delta}\right)^n < \frac{1}{2}\,\varepsilon\,, \qquad \text{for } n \geqq n_0\,,$$

which proves (16) and our theorem.

The proof of Theorem 1 suggests that $\lim_{n\to\infty} C^n f(\xi) = a$ implies $\lim_{\sigma\uparrow\infty} C_\sigma f(\xi) = a$. The following lemma shows that this is in fact so.

LEMMA 1. *Suppose that* $g(x) \in \Phi_{1,0}$ *and*

(17) $$\lim_{n\to\infty} n \cdot \int_0^1 (1-t)^{n-1} g(t)dt = a\ .$$

Then

(18) $$\lim_{\sigma\uparrow\infty} \sigma \cdot \int_0^1 (1-t)^{\sigma-1} g(t)dt = a\ .$$

Proof. The proof follows easily on integration by parts and by noticing that the function $G(x)$, defined by $G(x) = \int_0^x g(t)dt$, is bounded for $0 < x < 1$ and that $\lim_{x\downarrow 0} G(x) = 0$.

In the formulation of Theorem 1 and Theorem 1* we have made a distinction between the limits (8) and (8*). This is not really necessary: the two limits are equivalent.

THEOREM 2. *$f(x) \in \Phi_{k,0}$ and*

(19) $$\lim_{x\uparrow\infty} C_k f(x) = a$$

for some $k > 0$ imply

(20) $$\lim_{x\uparrow\infty} H_k f(x) = a\ .$$

Conversely, $f(x) \in \Phi_{k,0}$ and (20) imply (19).

THEOREM 2*. *$f(x) \in \Phi_{1,0}$ and*

(19*) $$\lim_{x\downarrow 0} C_k f(x) = a$$

for some $k > 0$ imply that $f(x) \in \Phi_{k,0}$ and

(20*) $$\lim_{x\downarrow 0} H_k f(x) = a\ .$$

Conversely, $f(x) \in \Phi_{k,0}$ and (20*) *imply* (19*).

Theorem 2 and Theorem 2* are the continuous analogues of the well-known Knopp-Schnee equivalence theorem for sequences. Note that in the first statement of Theorem 2 it is necessary to assume $f(x) \in \Phi_{k,0}$; otherwise it may happen that $H^k f(x)$ does not exist, for example $f(x) = (x^2 \log x)^{-1}$. This is not a serious restriction, though, since the assumption only affects the behaviour of $f(x)$ in the neighbourhood of 0 and is obviously satisfied in the essential domain of (C, k). There is no restriction of this kind in Theorem 2* where the assumption $f(x) \in \Phi_{1,0}$ and the existence of the limit (19*) automatically ensures that $f(x) \in \Phi_{k,0}$.

Theorem 2 is due to Landau [7]; Theorem 2 is stated (without proof and without specifiying the precise conditions of $f(x)$) by Hardy and Littlewood [5, p. 96]. It can be proved by an argument similar to the one used in [4, p. 112].

Once we have established the strong regularity of the limits (10), there is no difficulty in constructing strongly regular methods for $x \uparrow \infty$ and hence for sequences. We first note that limits at 0 by any method T can be converted into limits at ∞ by a 'reciprocal' method T^* which is defined as follows.

Suppose that T is given by

$$Tf(x) = \int_0^\infty \tau(x, t)f(t)dt .$$

To indicate clearly the function that is transformed and the point where the transform is taken, we shall use the notation

$$Tf(x) = T[f(t)](x) .$$

By an obvious change of variable

$$\begin{aligned} T[f(t)]\left(\frac{1}{x}\right) &= \int_0^\infty \tau\left(\frac{1}{x}, t\right)f(t)dt \\ &= \int_0^\infty \tau^*(x, t)f\left(\frac{1}{t}\right)dt \\ &= T^*\left[f\left(\frac{1}{t}\right)\right](x) \end{aligned}$$

where

$$\tau^*(x, t) = t^{-2}\tau\left(\frac{1}{x}, \frac{1}{t}\right) .$$

Clearly

$$T^k[f(t)]\left(\frac{1}{x}\right) = T^{*k}\left[f\left(\frac{1}{t}\right)\right](x) , \qquad k = 1, 2, \cdots .$$

Therefore

$$\lim_{x \downarrow 0} T_k[f(t)](x) = a$$

implies

$$\lim_{x \uparrow \infty} T_k^*\left[f\left(\frac{1}{t}\right)\right](x) = a ,$$

and conversely. Also

$$\lim_{k \to \infty} T_k[f(t)]\left(\frac{1}{x}\right) = \lim_{k \to \infty} T_k^*\left[f\left(\frac{1}{x}\right)\right](x) ,$$

in the sense that if one of the expressions exists then so does the other and the two are equal. It follows therefore that if T is strongly regular with respect to $x \downarrow 0$ then T^* is strongly regular with respect to $x \uparrow \infty$. In particular

$$C^*s(x) = \int_x^\infty t^{-2}s(t)dt , \tag{25}$$

and

$$H^*s(x) = x\int_x^\infty t^{-2}s(t)dt \equiv \int_1^\infty t^{-2}s(xt)dt \tag{26}$$

are strongly regular for $x \uparrow \infty$. Note that the domain of C^* and H^* is $\Phi_{0,1}$.

The processes (25) and (26) can easily be converted, if we wish, into strongly regular methods for sequences; for instance

$$(H^*\alpha)_n = n\sum_{m=n}^{\infty} \frac{\alpha_m}{m(m+1)} \tag{27}$$

is such a method. Its strong regularity is proved if it is shown that $\beta_n \to 0$ implies $\lim_{k\to\infty}(H^{*k}\beta)_n = 0$ for every fixed n. This can be shown for instance by comparing the sequence $(H^{*k}\beta)_n$ with suitable integrals and applying Theorem 1*.

Altough the method (27) is equivalent (in the ordinary sense) to $(H, 1)$[3], the two methods behave very differently from the point of view of iteration. The Hölder process has no useful infinite iterate whereas the process (27) has an infinite iterate which contains (and is compatible with) every finite (H^*, k). There exist in fact sequences (both bounded and unbounded) which are summable (H^*, ∞), but not summable by any finite (H^*, k) and (H, k)[4]. On the other hand, there exist (unbounded) sequences which are summable (H, k), but not by any (H^*, k), $k > 0$; for instance $(H^*, 1)$ is not even applicable to $\alpha_n = (-1)^n n(n+1)$. This raises the question of the relative strength of (H^*, k) and (H, k); we shall consider the problem only for the continuous case.

The following theorem is due to R. P. Agnew [**1**]; it is the continuous analogue of Knopp's equivalence theorem and asserts the equivalence of H and H^* for functions.

THEOREM 3. *$f(x) \in \Phi_{0,1}$ and*

$$\lim_{x\downarrow 0} x \cdot \int_x^\infty t^{-2}f(t)dt = a \tag{28}$$

imply $f(x) \in \Phi_{1,1}$ and

$$\lim_{x\downarrow 0} x^{-1} \cdot \int_0^x f(t)dt = a\,. \tag{29}$$

Converesely, $f(x) \in \Phi_{1,0}$ and the existence of the limit (29) *imply $f(x) \in \Phi_{1,1}$ and* (28).

A similar statement (with $\Phi_{0,1}$ and $\Phi_{1,0}$ interchanged) holds when $x \downarrow 0$ is replaced by $x \uparrow \infty$.

---

[3] A proof is given in [**6**, p. 487].

[4] Examples for the continuous case will be given below.

In the general case of (C, k) and (C^*, k) or (H, k) and (H^*, k) a stronger assumption on $f(x)$ is necessary. Because of the Theorem 2 and Theorem 2* it is sufficient to consider one of the possible combinations, say C and C^*.

THEOREM 3.* *Let* $f(x) \in \Phi_{1,1}$ *and suppose that for some* $n > 0$,

(30) $$\lim_{x \downarrow 0} C_n^* f(x) = a$$

then

(31) $$\lim_{x \downarrow 0} C_n f(x) = a\ .$$

Conversely, the existence of the limit (31) *and* $f(x) \in \Phi_{1,1}$ *imply* (30). A similar statement holds when $x \downarrow 0$ is replaced by $x \uparrow \infty$.

Theorem 3* shows that (C, n) and (C^*, n) are equivalent within $\Phi_{1,1}$, that is within the class of functions to which both methods are applicable. However, if we disregard the difficulty that $f(x)$ may behave badly at ∞ (0) when we are interested in the limit at 0 (∞), that is, if we restrict ourselves to the essential domain of the two methods, then it appears that C^* includes C for limits at 0, and C includes C^* for limits at ∞. C^* is actually stronger than C for $x \downarrow 0$, as shown by the example $f(x) = 2x^{-3} \sin x^{-2}$. In fact $C[2t^{-3} \sin t^{-2}](x)$ does not exist since the function is not integrable down to 0, but

$$C^*[2t^{-3} \sin t^{-2}](x) = -\frac{1}{x^2} + \sin \frac{1}{x^2}$$

$$C^*\left[-t^{-2} \cos \frac{1}{t^2} + \sin \frac{1}{t^2}\right](x) = O\left(\frac{1}{x}\right),$$

hence

$$\lim_{x \downarrow 0} x^2 C^{*2}\left[\frac{2}{t^3} \sin \frac{1}{t^2}\right](x) = 0\ .$$

This example shows that the condition $f(x) \in \Phi_{1,1}$ cannot be relaxed and for instance $f(x) \in \Phi_{0,1}$ and the existence of $\lim_{x \downarrow 0} C_2^* f(x) = a$ does not imply $f(x) \in \Phi_{1,1}$.

For the proof of Theorem 3 we need the following lemma.

LEMMA 2. *Given* $\infty \geqq a_1 \geqq a_2 \geqq \cdots \geqq a_n > 0$, $n > 0$, *and* $f(x) \in \Phi_{1,1}$. *Let* $f_k(x)$ *for* $k = 0, 1, \cdots, n$ *be defined by*

$$f_0(x) = f(x)\ ,\ \ f_k(x) = \int_x^{a_k} t^{-2} f_{k-1}(t)dt \quad \textit{for } k > 0\ .$$

Then

$$\lim_{x \downarrow 0} x^{k+1} f_k(x) = 0 \qquad \textit{for } k = 1, 2, \cdots, n\ .$$

Proof. $\int_0^x f(t)dt$ exists by assumption. Therefore given $\varepsilon > 0$ we can choose a positive δ such that

$$\left| \int_\xi^\eta f(t)dt \right| < \varepsilon$$

for every $0 < \xi < \eta < \delta$. Let $\xi < \delta$. By the second mean value theorem

$$\xi^2 \int_\xi^\delta t^{-2} f(t)dt = \int_\xi^\eta f(t)dt$$

for some η in the interval (ξ, δ). Hence

$$\xi^2 \left| \int_\xi^\delta t^{-2} f(t)dt \right| < \varepsilon \qquad \text{for every } 0 < \xi < \delta .$$

Also

$$\xi^2 \left| \int_\delta^{a_1} t^{-2} f(t)dt \right| < \varepsilon \qquad \text{for every } 0 < \xi < \xi_0 \leqq \delta$$

provided that ξ_0 is sufficiently small. Therefore

$$\left| \xi^2 \int_\xi^{a_1} t^{-2} f(t)dt \right| < 2\varepsilon \qquad \text{for every } \xi \text{ in } (0, \xi_0).$$

This proves the lemma for $k = 1$. Suppose now that $k > 1$ and $f_{k-1}(x) = o(x^{-k})$, as $x \downarrow 0$. Given $\varepsilon > 0$ choose $\delta \leqq a_k$ so that $|f_{k-1}(t)| < \varepsilon t^{-k}$ for $0 < t < \delta$. For $0 < x < \delta$

$$|f_k(x)| \leqq \left| \int_x^\delta t^{-2} f_{k-1}(t)dt \right| + \left| \int_\delta^{a_k} t^{-2} f_{k-1}(t)dt \right| .$$

But

$$\left| \int_x^\delta t^{-2} f_{k-1}(t)dt \right| < \varepsilon \int_x^\delta t^{-k-2} dt < \frac{\varepsilon}{k+1} x^{-k-1}$$

and

$$\left| \int_\delta^{a_k} t^{-2} f_{k-1}(t)dt \right| < \frac{\varepsilon k}{k+1} x^{-k-1}$$

provided that x is sufficiently small, $0 < x < x_0 \leqq \delta$, say. Therefore $|f_k(x)| < \varepsilon x^{-k-1}$ for all x, $0 < x < x_0 = x_0(\varepsilon)$.

COROLLARY. *For $k > 0$ and $f(x) \in \Phi_{1,1}$ we have $C^{*k} f(x) = o(x^{-k-1})$ as $x \downarrow 0$.*

Proof of Theorem 3*. For simplicity we shall write $D = C^*$ throughout the proof. It is convenient to prove the first statement of the theorem in the following more general form: Suppose that $f(x) \in \Phi_{1,1}$ and for some $n > 0$ and $p \geqq 0$,

$$\lim_{x \downarrow 0} (n+p)!\, x^{n+p} D^{n+p} f(x) = a , \tag{32}$$

where $Df(x) = \int_x^\infty t^{-2}f(t)dt$; then for every $r \geqq 0$, $s \geqq 0$

$$\lim_{x \downarrow 0} \frac{(s+n)!\,(p+r)!}{s!} x^{-(n+s)} C^n[t^{p+r+s}D^{p+r}f(t)](x) = a \,. \tag{33}$$

The proof is by induction on n. We first note that (32) implies

$$D^m f(x) = \frac{1}{m!} \cdot x^{-m} \cdot a + o(x^{-m}) \,, \quad \text{as } x \downarrow 0 \,, \tag{34}$$

for every $m \geqq n + p$. Now let $k \geqq 0$ and $m \geqq n + p$. We have, by partial integration,

$$\begin{aligned} k(m-1)!\,x^{-k}&\int_0^x t^{m+k-2}D^{m-1}f(t)dt \\ &= -k(m-1)!\,x^{-k}[t^{m+k}D^m f(t)]_0^x \\ &\qquad + k(m-1)!\,(m+k)x^{-k}\int_0^x t^{m+k-1}D^m f(t)dt \,. \end{aligned}$$

The first expression on the right is $-k(m-1)!\,x^m D^m f(x)$ which, by (34), tends to $-(k/m)a$ when $x \downarrow 0$; similary, the second term tends to $[(m+k)/m]\cdot a$. Hence

$$\lim_{x \downarrow 0} k(m-1)!\,x^{-k} \cdot \int_0^x t^{m+k-2}D^{m-1}f(t)dt = a \,. \tag{35}$$

This proves (33) for $n = 1$ (with $k = s + 1$, $m = p + r + 1$). Note that $f(x) \in \Phi_{0,1}$ and the existence of the limit (32) implies the existence of the integral in (35) ; therefore in particular $f(x) \in \Phi_{0,1}$ and (28) in Theorem 3 implies $f(x) \in \Phi_{1,1}$ and (29). Suppose now that $n > 1$, and write $m = n + p + r$, $r \geqq 0$. We have

$$\begin{aligned} k(m-1)!\,x^{-x} &\cdot \int_0^x t^{m+k-2}D^{m-1}f(t)dt \\ &= k(m-1)!\,x^{-k}\int_0^x t_1^{m+k-2}dt_1\int_{t_1}^\infty t_2^{-2}dt_2 \cdots \int_{t_{n-2}}^\infty t_{n-1}^{-2}dt_{n-1}\int_{t_{n-1}}^\infty t_n^{-2}D^{p+r}f(t_n)dt_n \\ &= k(m-1)!\,x^{-k}\Big\{D^{m-1}f(x)\int_0^x t_1^{m+k-2}dt_1 + D^{m-2}f(x)\int_0^x t_1^{m+k-2}dt_1\int_{t_1}^x t_2^{-2}dt_2 \\ &\qquad + \cdots + D^{p+r+1}f(x)\int_0^x t_1^{m+k-2}dt_1 \cdots \int_{t_1}^x t_2^{-2}dt_2\int_{t_{n-2}}^x t_{n-1}^{-2}dt_{n-1} \\ &\qquad\qquad + \int_0^x t_1^{m+k-2}dt_1\int_{t_1}^x t_2^{-2}dt_2 \cdots \int_{t_{n-1}}^x t_n^{-2}D^{p+r}f(t_n)dt_n\Big\} \\ &= \frac{k(m-1)!}{(m+k-1)!}\Big\{\sum_{j=1}^{n-1}(j+k+p+r-1)!\,x^{j+p+r}D^{j+p+r}f(x) \\ &\qquad + (k+p+r)!\,x^{-k}\int_0^x t^{k+p+r+1}D^{p+r}f(t)dt\Big\} \,. \end{aligned} \tag{36}$$

The last expression in the brackets is obtained by repeated partial integration, and using Lemma 2 Equations (35) and (36) give

$$x^{-k}\int_0^x t^{k+p+r-1}D^{p+r}f(t)dt = \frac{(m+k-1)!}{k(m-1)!(k+p+r)!}a$$
$$-\sum_{j=1}^{m-1}\frac{(j+k+p+r-1)!}{(k+p+r)!}x^{j+p+r}D^{j+p+r}f(x)+o(1)\,.$$

Hence

$$\frac{(n+k-1)!}{(k-1)!}x^{-n-k+1}C^n[t^{k+p+r-1}D^{p+r}f(t)](x)$$
$$=\frac{(n+k-1)!(p+r)!}{(k-1)!}x^{-n-k+1}C^{n-1}\left[\int_0^t u^{k+p+r-1}D^{p+r}f(u)du\right](x)$$
$$=\frac{(p+r)!(n+k+p+r-1)!}{(k+p+r)!(n+p+r-1)!}a$$
$$-\sum_{j=1}^{n-1}\frac{(n+k-1)!(k+j+p+r-1)!(p+r)!}{(k+p+r)!(k-1)!}x^{-n-k+1}$$
$$\times C^{n-1}[t^{j+k+p+r}D^{j+p+r}f(t)](x)+o(1)\,.$$

Here we have, by the induction hypothesis (33), applied to $Df(x)$ instead of $f(x)$ and $n-1$, $p+1$,

$$\frac{(n+k-1)!(j+p+r)!}{k!}x^{-n-k+1}C^{n-1}[t^{j+k+p+r}D^{j+p+r}f(t)](x)=a+o(1)\,.$$

Hence by (37)

$$\frac{(n+k-1)!(p+r)!}{(k-1)!}x^{-n-k+1}C^n[t^{k+p+r-1}D^{p+r}f(t)](x)$$
$$=\frac{(p+r)!k!}{(k+p+r)!}\left\{\binom{n+k+p+r-1}{k}-\sum_{j=1}^{n-1}\binom{j+k+p+r-1}{k-1}\right\}a+o(1)$$
$$=a+o(1)\,.$$

This proves (33). The proof of the converse is very similar. With the notation $s(x)\equiv f(1/x)$ the converse statement can be formulated as follows:

$$\lim_{x\uparrow\infty} n!\cdot x^nD^ns(x)=a \tag{30*}$$

implies

$$\lim_{x\uparrow\infty} n!x^{-n}C^ns(x)=a\,. \tag{31*}$$

The proof is identical with the derivation of (31) from (30) except that $f(x)$ has to be replaced everywhere by $s(x)$ (which is also in $\Phi_{1,1}$) and $x\downarrow 0$ by $x\uparrow\infty$.

3. The relative strengh of the (H, ∞) and (C, ∞) methods. So far we did not consider the relative strengh of the (H, ∞) and (C, ∞) methods. We know from Theorem 1 and Theorem 1* that both these methods include the finite (H, k) and (C, k) methods for $x \downarrow 0$, more precisely, if

$$\lim_{x \downarrow 0} C_k f(x) = \lim_{x \downarrow 0} H_k f(x) = a$$

exists for some $k \geqq 0$ then

$$H_\infty f(x) = \lim_{m \uparrow \infty} H_m f(x) \quad \text{and} \quad C_\infty f(x) = \lim_{m \uparrow \infty} C_m f(x)$$

exist for every $x > 0$ and are in fact the constant function $H_\infty f(x) = C_\infty f(x) = a$.

Now the following theorems show that this is always so: whenever $C_\infty f(x)$ and $H_\infty f(x)$ exist at all, they are a constant.

THEOREM 4. *Let $f(x) \in \Phi_I$ and suppose that for some fixed $\xi > 0$*

(38) $$\lim_{n \to \infty} C_n f(\xi) = a$$

then

(39) $$\lim_{\sigma \uparrow \infty} \sigma \int_0^\infty e^{-\sigma t} f(t) dt = a \, .$$

Conversely, $f(x) \in \Phi_I$ and (39) imply

$$\lim_{n \to \infty} C_n f(x) = a$$

for every $x > 0$.

THEOREM 4*. *Let $f(x) \in \Phi_{\infty,0}$ and suppose that for a fixed $\xi > 0$,*

(40) $$\lim_{n \to \infty} H^n f(\xi) = a \, ;$$

then for every $x > 0$

$$\lim_{n \to \infty} H^n f(x) = a \, .$$

Theorem 4 shows that (C, ∞) is essentially equivalent to the Abel-Poisson method L:

(41) $$Lf(x) = \int_0^\infty e^{-t} f(xt) dt$$

in the sense that $\lim_{x \downarrow 0} C_\infty f(x) = a$ if, and only if, $\lim_{x \downarrow 0} Lf(x) = a$ provided that $f(x)$ is in the essential domain of the two methods. As a corollary we find that L includes every (C, k)[5]; but we know of no example to show that (C, ∞) or L is actually stronger than the collection of every (C, k). For bounded functions (C, ∞) is equivalent to $(C, 1)$; more generally the following is true.

[5] A dual of this statement, referring to $x \to \infty$, is proved by G. Doetsch in [3, p. 204].

THEOREM 5. *$f(x) \in \Phi_{1,0}$, $f(x) = 0$ for $x \geqq x_0$, $f(x) = O_L(1)$ and $C_\infty f(x) = a$, $x > 0$, imply*

$$\lim_{x \downarrow 0} C_1 f(x) = a .$$

Proof of Theorem 4. By Lemma 1 we can replace the integral variable n by the continuous variable σ. The assumption $f(x) \in \Phi_I$ implies that $F(x) \equiv Cf(x)$ is bounded for $x > 0$ and $\lim_{x \downarrow 0} F(x) = 0$. Therefore we obtain for $\sigma > 1$

$$\begin{aligned}(42)\quad C_\sigma f(\xi) &= \sigma\int_0^1 (1-t)^{\sigma-1} f(\xi t)dt \\ &= \frac{\sigma-1}{\xi}\int_0^\sigma \left(1-\frac{u}{\sigma}\right)^{\sigma-2} F\left(\frac{\xi u}{\sigma}\right)du \\ &= \frac{\sigma-1}{\xi}\int_0^{\sigma^{1/2}} \left(1-\frac{u}{\sigma}\right)^{\sigma-2} F\left(\frac{\xi u}{\sigma}\right)du + O(\sigma e^{-\sigma/2}) \\ &= \frac{\sigma-1}{\xi}\int_0^{\sigma^{1/2}} e^{-u} F\left(\frac{\xi u}{\sigma}\right)du + O\left(\int_0^{\sigma^{1/2}} u^2 e^{-2}\left|F\left(\frac{\xi u}{\sigma}\right)\right|du\right) \\ &\qquad + O(\sigma e^{-\sigma^{1/2}}) \\ &= \frac{\sigma(\sigma-1)}{\xi}\int_0^\infty e^{-\sigma t}F(\xi t)dt + o(1) \\ &= (\sigma-1)\int_0^\infty e^{-\sigma t} f(\xi t)dt + o(1)\end{aligned}$$

where all O and o symbols refer to fixed ξ and $\sigma \uparrow \infty$. Hence $\lim_{\sigma\uparrow\infty} C_\sigma f(\xi) = a$ implies $\lim_{\sigma\uparrow\infty}\sigma\int_0^\infty e^{-\sigma t} f(\xi t)dt = a$. Therefore

$$\lim_{\rho\uparrow\infty} \rho\int_0^\infty e^{-\rho u} f(xu)du = a$$

for any fixed $x > 0$. By (42) we see that $\lim_{\rho\uparrow\infty} C_\rho f(x) = a$.

Proof of Theorem 5. Theorem 5 is an immediate consequence of Theorem 4 and the following lemma, which is a special case of a well-known Tauberian theorem for the Laplace transform (see [3, p. 210, Satz 3]).

LEMMA 3. *Suppose that $g(x) = O_L(1)$, $\int_0^\infty e^{-\sigma t} g(t)dt$ converges for all $\sigma > 0$ and*

$$\lim_{\sigma\uparrow\infty} \sigma\int_0^\infty e^{-\sigma t} g(t)dt = a .$$

Then

$$\lim_{x\downarrow 0} \frac{1}{x}\int_0^x g(t)dt = a .$$

The statement remains true if $\sigma \uparrow \infty$ *is replaced by* $\sigma \downarrow 0$ *and* $x \downarrow 0$ *by* $x \uparrow \infty$.

Proof of Theorem 4*. Without loss of generality we may assume $a = 0$; otherwise consider $f(x) - a$ instead of $f(x)$. Let x_1, x_2 be two fixed positive numbers $x_1 < \xi < x_2$. Write $g(x) = Hf(x)$. Then $H^n f(x) = H^{n-1}g(x)$, for $n > 0$, and $\lim_{n\to\infty} H^n g(\xi) = 0$. We shall show that

$$\text{(40*)} \qquad \lim_{n\to\infty} H^n g(x) \equiv \lim_{n\to\infty} H^{n+1}f(x) = 0, \text{ uniformly for } x_1 \leqq x \leqq x_2.$$

Denote

$$\text{(43)} \qquad \begin{cases} \varepsilon_0 = \dfrac{x_1}{\xi} \cdot \text{upper bound } \{|g(x)|;\ x_1 \leqq x \leqq x_2\} < +\infty , \\ \varepsilon_k = |H^k g(\xi)| \qquad \text{for } k > 0 . \end{cases}$$

We prove that

$$\text{(44)} \qquad |H^n g(x)| \leqq \frac{\xi}{x_1} \sum_{p=0}^{n} \frac{1}{p\,!} \left| \frac{x-\xi}{x_1} \right|^p \varepsilon_{n-p} \qquad \text{for } x_1 \leqq x \leqq x_2 .$$

For $n = 0$ the statement follows from (43); suppose therefore that $n > 0$ and that (44) is true for $n - 1$.

$$\begin{aligned} |H^n g(x)| &\leqq \left| \frac{1}{x}\int_0^{\xi} H^{n-1}g(t)dt \right| + \left| \frac{1}{x}\int_{\xi}^{x} H^{n-1}f(t)dt \right| \\ &\leqq \left| \frac{\xi}{x_1} \cdot \frac{1}{\xi}\int_0^{\xi} H^{n-1}g(t)dt \right| + \frac{\xi}{x_1^2} \sum_{p=0}^{n-1} \int_{\xi}^{p} \frac{1}{p\,!} \left| \frac{t-\xi}{x_1} \right|^p \varepsilon_{n-p-1}dt \\ &\leqq \frac{\xi}{x_1} \left\{ \varepsilon_n + \sum_{p=0}^{n-1} \frac{1}{(p+1)\,!} \left| \frac{x-\xi}{x_1} \right|^{p+1} \varepsilon_{n-p-1} \right\} \end{aligned}$$

which proves the statement for n. From (44) we obtain, by writing $\lambda = \max\{\xi - x_1, x_2 - \xi\}$,

$$\text{(45)} \qquad |H^n g(x)| \leqq \frac{\xi}{x_1} \sum_{p=0}^{n} \frac{1}{(n-p)\,!} \lambda^{n-p} \varepsilon_p , \qquad x_1 \leqq x \leqq x_2 .$$

But for any $\lambda > 0$,

$$e^{-\lambda} \sum_{p=0}^{n} \frac{1}{(n-p)\,!} \lambda^{n-p} s_p$$

is a regular transform of the sequence $\{s_k\}$, $k \geqq 0$; it follows therefore that the expression on the right side of (45) tends to zero when $n \to \infty$.

By Theorem 5 (C, ∞) does not extend the range of $(C, 1)$ for bounded functions. This is in striking contrast with (H, ∞) which is decidedly more powerful for bounded functions than (H, 1). An example is furnished by $\cos \log x$, or more conveniently by $e^{-i \log x}$. We find by induction

$$H^k[e^{-i\log t}](x) = \left(\frac{1+i}{2}\right)^k e^{-i\log x}$$

which has no limit at $x = 0$; by Theorem 2* and Theorem 5 it has therefore no limit by (C, ∞). On the other hand it has limit zero by (H, ∞):

$$\lim_{k\to\infty} H^k[e^{-i\log t}](x) = 0 \qquad \text{for every } x > 0 .$$

Also for unbounded functions (H, ∞) appears to be more effective than (C, ∞); a suitable example is $-x^{-1/2}\cdot\cos\log x$ or $x^{-1/2}e^{-i\log x}$ which can be shown to have no limit by (C, ∞), and limit zero by (H, ∞). These examples reveal a remarkable difference (in favour of the Hölder process) between the Cesàro and Hölder processes which remains completely hidden when finite iterations alone are considered.

For bounded functions repeated partial integration gives

$$H^{k+1}f(x) = \frac{1}{k\,!}\int_0^1\left(\log\frac{1}{t}\right)^k f(xt)dt\ ; \tag{46}$$

and we find the following analogue of the equivalence of (C, ∞) and L for (H, ∞):

THEOREM 6. *Let $f(x) \in \Phi_B$ and*

$$I_0(x) \equiv J_0(ix) \equiv \sum_{n=0}^{\infty}\frac{1}{(n\,!)^2}\left(\frac{t}{2}\right)^{2n} .$$

Then

$$\lim_{n\to\infty}\int_0^1\frac{1}{n\,!}\left(\log\frac{1}{t}\right)^n f(t)dt = \lim_{v\to\infty} e^{-v}\int_0^1 f(t)I_0\left(2\left(v\log\frac{1}{t}\right)^{1/2}\right)dt \tag{47}$$

in the sense that if one side exists then the other side exists too and they are equal.

By making use of the well-known asymptotic expression

$$I_0(x) = (2\pi)^{-1/2}x^{-1/4}e^x\left(1 + O\left(\frac{1}{x}\right)\right), \qquad \text{as } x\uparrow\infty ,$$

Theorem 6 can be put in a more convenient form. For bounded functions we have

$$\begin{aligned}
&\lim_{v\uparrow\infty} e^{-v}\int_0^1 f(t)I_0\left(2\left(v\log\frac{1}{t}\right)^{1/2}\right)dt\\
&\quad = \lim_{v\uparrow\infty}\frac{1}{2}\pi^{-1/2}e^{-v}\int_0^1 f(t)\left(v\log\frac{1}{t}\right)^{-1/4}\exp\left[2\left(v\log\frac{1}{t}\right)^{1/2}\right]dt\\
&\quad = \lim_{\sigma\uparrow\infty}\pi^{-1/2}\int_0^\infty f(e^{-u^2})\cdot e^{-(u-\sigma)^2}\cdot\left(\frac{u}{\sigma}\right)^{1/2}du
\end{aligned}$$

(by the substitution $v = \sigma^2$, $t = e^{-u^2}$), and the latter is easily seen to be equal to

$$\lim_{\sigma\uparrow\infty} \pi^{-1/2}\int_0^\infty f(e^{-u^2})e^{-(u-\sigma)^2}du .$$

This gives

THEOREM 6*. *Let $f(x) \in \Phi_B$; then $\lim_{x\downarrow 0} f(x) = a$ by (H, ∞) if, and only if,*

$$\lim_{\sigma\uparrow\infty} \pi^{-1/2}\int_0^\infty f(e^{-u^2})e^{-(u-\sigma)^2}du = \lim_{v\uparrow\infty} \frac{1}{2}\pi^{-1/2}e^{-v}\int_0^1 f(t)\frac{\exp(2(v\log 1/t)^{1/2}}{(\log 1/t)^{1/2}}dt = a .$$

The following estimate of (H, ∞) for bounded functions is weaker, but it has the advantage of great formal simplicity.

THEOREM 7. *Le $f(x) \in \Phi_B$ and suppose that for a fixed $\xi > 0$*

$$\lim_{n\to\infty} H_n f(\xi) = a ,$$

then

$$\lim_{x\downarrow 0} H[f(e^{-1/t})](x) = \lim_{x\uparrow\infty} H[f(e^{-t})](x) = a .$$

Theorem 6, Theorem 6* and Theorem 7 do not remain valid for unbounded functions; a suitable counter example is $x^{-1/2}\cos\log(1/x)$. Also the converse of Theorem 7 is not true; a counter example is furnished by $f(x) = \exp(i(\log 1/x)^{1/2})$. Clearly

$$\frac{1}{x}\int_0^x f(e^{-1/t})dt = \frac{1}{x}\int_0^x \exp(it^{-1/2})dt = O(x^{1/2}) \to 0 \quad \text{when } x\downarrow 0 .$$

On the other hand

$$\lim_{n\to\infty}\int_0^\infty \frac{1}{n!}\left(\log\frac{1}{t}\right)^n \exp\left[i\left(\log\frac{1}{t}\right)^{1/2}\right]dt$$

does not exist; for otherwise by Theorem 6*

$$\lim_{\sigma\uparrow\infty} \pi^{-1/2}\int_0^\infty \exp(iu-(u-\sigma)^2)du$$
$$= \lim_{\sigma\uparrow\infty} \pi^{-1/2}\exp\left(-\frac{1}{u}-i\sigma\right)\int_0^\infty \exp\left[-\left(u-\sigma-\frac{1}{2}\right)^2\right]du$$

existed, But the last expression is asymptotically equal to $e^{-1/4-i\sigma}$ when $\sigma\uparrow\infty$.

In the proof of Theorem 6 we use

LEMMA 5. *Let $f(x) \in \Phi_B$. Then for every fixed $\xi > 0$,*

$$H^n f(\xi) - H^{n+1}f(\xi) = O(n^{-1/2}) \qquad \text{as } n\to\infty .$$

Proof. Observing the relations

$$\frac{1}{(n+1)!}\int_x^\infty u^{n+1}e^{-u}du - \frac{1}{n!}\int_x^\infty u^n e^{-u}du$$
$$= \frac{1}{n!}\int_0^x u^n e^{-u}du - \frac{1}{(n+1)!}\int_0^x u^{n+1}e^{-u}du$$
$$= \frac{x^{n+1}e^{-x}}{(n+1)!},$$

$$n! \sim \sqrt{2\pi}\, n^{n+1/2}e^{-n}, \qquad \text{as } n \to \infty,$$

$$|f(t)| \leqq K \qquad \text{for a suitable constant } K > 0,$$

we obtain

$$\left|\int_0^1 f(\xi t)\frac{1}{n!}\left(\log\frac{1}{t}\right)^n dt - \int_0^1 f(\xi t)\frac{1}{(n+1)!}\left(\log\frac{1}{t}\right)^{n+1} dt\right|$$
$$\leqq K\int_0^1\left|\frac{1}{n!}\left(\log\frac{1}{t}\right)^n - \frac{1}{(n+1)!}\left(\log\frac{1}{t}\right)^{n+1}\right|dt$$
$$= K\int_0^\infty\left|\frac{u^n}{n!} - \frac{u^{n+1}}{(n+1)!}\right|e^{-u}du$$
$$= 2K\frac{(n+1)^{n+1}}{(n+1)!}e^{-(n+1)}$$
$$\sim \frac{2K}{\sqrt{2\pi}}\cdot\frac{1}{\sqrt{n+1}}.$$

Proof of Theorem 6. By the regularity of the Borel transform, $\lim_{n\to\infty}s_n = a$ implies

$$\lim_{v\uparrow\infty} e^{-v}\sum_{n=0}^\infty \frac{s_n}{n!}\cdot v^n = a.$$

Hence

$$\lim_{n\to\infty}\int_0^1 f(t)\frac{1}{n!}\left(\log\frac{1}{t}\right)^n dt = a \tag{48}$$

implies

$$\lim_{v\uparrow\infty} \mathrm{e}^{-v}\sum_{n=0}^\infty\int_0^1 f(t)\frac{1}{(n!)^2}\left(\log\frac{1}{t}\right)^n v^n dt$$
$$= \lim_{v\uparrow\infty} e^{-v}\int_0^1 f(t)\sum_{n=0}^\infty\frac{1}{(n!)^2}\left(v\log\frac{1}{t}\right)^n dt$$
$$= \lim_{v\uparrow\infty} e^{-v}\int_0^1 f(t)I_0\left(2\left(v\log\frac{1}{t}\right)^{1/2}\right)dt = a; \tag{49}$$

the interchange of the order of summation and integration is clearly permissible if $f(t)$ is bounded. Conversely, from Lemma 5 and the Tauberian theorem of Hardy-Littlewood for the Borel transform [4, p.

220, Theorem 156] we conclude that (49) implies (47).

Proof of Theorem 7. If in the proof of Theorem 6 we use the Abel transform instead of the Borel transform, Theorem 7 is obtained. First

$$\lim_{n\to\infty}\int_0^1 f(t)\frac{1}{n!}\left(\log\frac{1}{t}\right)^n dt = a$$

implies

$$\begin{aligned} a &= \lim_{v\uparrow 1}(1-v)\sum_{n=0}^{\infty}\int_0^1 f(t)\frac{1}{n!}\left(v\log\frac{1}{t}\right)^n dt \\ &= \lim_{v\uparrow 1}(1-v)\int_0^1 f(t)t^{-v}dt \\ &= \lim_{\sigma\downarrow 0}\int_1^\infty f\left(\frac{1}{u}\right)u^{-\sigma+1}du \\ &= \lim_{\sigma\downarrow 0}\sigma\int_0^\infty f(e^{-t})e^{-\sigma t}dt \end{aligned}$$

and this implies by Lemma 3

$$\lim_{x\uparrow\infty}\frac{1}{x}\int_0^x f(e^{-t})dt = \lim_{x\uparrow\infty} H[f(e^{-t})](x) = a\ .$$

By Theorem 3 the last equation is equivalent to

$$\lim_{x\downarrow 0}\frac{1}{x}\int_0^x f(e^{-1/t})dt = \lim_{x\downarrow 0} H[f(e^{-1/t})](x) = a\ .$$

REFERENCES

1. R. P. Agnew, *Mean values and Frullani integrals*, Proc. Amer. Math. Soc. **2** (1951), 237-241.
2. K. Chandrasekharan and S. Minakshisundaram, *Typical means*, Oxford, 1952.
3. G. Doetsch, *Theorie und Anwendung der Laplace-Transformation*, Berlin, 1937.
4. G. H. Hardy, *Divergent series*, Oxford, 1949.
5. G. H. Hardy and J. E. Littlewood, *Solution of the Cesàro summability problem for power series and Fourier series*, Math. Z. **19** (1942), 67-96.
6. K. Knopp, *Theory and application of infinite series*, London, 1928.
7. E. Landau, Leipziger Ber. **65** (1913), 131-138.

THE UNIVERSITY OF ADELAIDE, AUSTRALIA AND
THE HEBREW UNIVERSITY, JERUSALEM

APPROXIMATION OF SEMI-GROUPS OF OPERATORS

H. F. TROTTER

1. Introduction. The usual methods for numerically computing the solution of a partial differential equation consist in replacing the differential operators by difference operators which approximate them, and taking the solution of the resulting difference equation as an approximation to the solution of the original equation. The question of *convergence* then arises; that is, when will a sequence of difference equations have the property that their solutions converge to the solution of a given differential equation? We treat this question in an operator-theoretic fashion, and our discussion has much in common with that of Lax and Richtmyer [17], as is pointed out in more detail below. The reader is referred to the bibliography of [17] for a list of the principal papers dealing with this question of convergence.

Our discussion will be limited to the initial value problem (Cauchy problem) for linear equations in the form

$$\frac{\partial}{\partial t}u(t, x) = \Omega u(t, x)\,;\, u(0, x) = f(x) \tag{1.1}$$

in which Ω is linear and constant in time. Here t is a non-negative real variable, and x is a point in some space S. Equation (1.1) is formally of parabolic type, but, as is shown in [14, Chap. XX], the initial value problem for hyperbolic equations can also be put into this form. Lateral conditions (i.e., boundary conditions such as are needed for the heat equation on a finite interval) are considered to be incorporated into the definition of Ω as restrictions on its domain [8, 17].

The process of setting up a sequence of finite difference approximations to (1.1) may be described in the following general terms. For each n, take a positive number h_n and a set $S_n \subset S$ whose points form a suitable grid. The solution to the nth approximating equation is defined inductively, for t an integral multiple of h_n and x a point of S_n by the following system of equations:[1]

[1] Sometimes only the space variable is made discrete, so that u_n is defined by a finite set of simultaneous *differential* equations (cf. [13, p. 233]). Theorem 5.2 can be applied to this situation just as Theorem 5.3 can be applied to the case in which the time variable is made discrete and the u_n are defined by (1.2).

This paper was originally accepted by the Transactions of the American Mathematical Society, Received July 3, 1957, by Trans. Amer. Math. Soc. Research done at Princeton University and supported in part by the Office of Ordnance Research U.S. Army.

$$(1.2) \qquad u_n((k+1)h_n, x) = T_n u(kh_n, x) \qquad k = 0, 1, 2, \cdots$$
$$u_n(0, x) = f_n(x)$$

where T_n is a linear operator and f_n is a function defined on S_n which suitably approximates f. For example, f_n may be simply the restriction of f to S_n. (Other possible ways of defining f_n are discussed in § 2.) It will be convenient to extend the definition of u_n to all t by setting

$$(1.3) \qquad u_n(t, x) = u_n(kh_n, x) \text{ for } kh_n \leqq t < (k+1)h_n .$$

In some numerical methods, the relation between the values of u_n at steps k and $k+1$ is given by a set of simultaneous linear equations, so that T_n is defined implicitly rather than by explicit formulae; the particular way in which T_n may be defined will be irrelevant to our discussion.

If the operators Ω_n defined by

$$(1.4) \qquad \Omega_n = h_n^{-1}(T_n - I)$$

converge to Ω in some suitable sence, then (1.2) converges formally to (1.1) as $n \to \infty$, and it is plausible that under certain conditions the functions u_n will converge to the solution of (1.1).

It was observed by von Neumann [20] that a system like (1.2) may be *unstable* in the sense that small errors in the initial data may lead to errors in $u_n(t)$ which become unbounded as $n \to \infty$. The definitions of stability given in the literature vary slightly in detail. We adopt essentially the same definition as that used in [17]. We suppose that the space of functions on S_n is normed as a Banach space. (Examples of how such a norm may be defined are given in § 2.) Then the norm of T_n as an operator on this space is defined. We shall say that the system (1.2) is *stable* if

$$(1.5) \qquad \| T_n^k \| \leqq Me^{Kkh_n}$$

for some constants K and M independent of n and k. The simpler condition $\| T_n \| \leqq 1$ is satisfied in many applications, and clearly implies (1.5).[2]

Although it is possible to find an example of an unstable system whose solutions converge to the correct result if the approximating functions f_n are appropriately chosen [18], Lax and Richtmyer [17] have shown that in general an unstable system cannot converge. On the other hand, they have shown that stability, together with some reasonable

[2] It should be pointed out that we are concerned only with the behaviour of the *exact* solutions of (1.2). In actual computation the effect of round-off errors must be considered, and the situation becomes more complicated.

assumptions on the limiting behaviour of the Ω_n, is sufficient to imply convergence. Our main result is very similar; however, our hypotheses differ in two respects from those of [17].

In the first place, we do not require that the limit function u and the approximating functions u_n all belong to the same Banach space. Frequently, the T_n arise most naturally as operators on functions which are defined only at a grid of points in the space S. In most practical applications the T_n can be modified so as to become operators on functions defined on the whole space, and this is assumed in [17]. Such modification, however, is usually unnatural, and in the case of random walks and diffusion processes which we discuss in § 6 it is not necessarily possible. Consequently it seems worth while to eliminate this assumption. We introduce the notion of an *approximating sequence of Banach spaces* and define associated concepts of convergence of vectors and operators. Section 2 is devoted to setting up the definitions and giving examples; it also includes some lemmas on the convergence of operators. Section 3 contains some remarks on the adjoint spaces of an approximating sequence which have application in § 6. Kantorovich [16] uses a similar sort of approximation of one Banach space by another, but requires the approximating space to be isomorphic to a subspace of the approximated space. This requirement is unnecessarily restrictive for our purposes.

To explain the second difference, we must describe more precisely what we mean by a solution of (1.1). We give an abstract formulation in terms of a semi-group of operators [14, chap. 20; 15]. Let X be some suitable Banach space of functions on S, and let Ω be a linear operator on X. Suppose that Ω is densely defined and has the property that for every f in its domain there exists a unique function $u(t, x)$ satisying

(1.6) (i) $u(t, x)$ and $\frac{\partial}{\partial t} u(t, x)$ are in X for all $t \geqq 0$

(ii) $\frac{\partial}{\partial t} u(t, x) = \Omega\, u(t, x)$ for $t \geqq 0$

(iii) $\| u(t, x) - f(x) \| \to 0$ as $t \to 0$

(iv) $\| u(t, x) \| \leqq M \| f(x) \|$ for $0 \leqq t \leqq 1$

Where M is a constant independent of f. Then for each t, setting

$$[T(t)f](x) = u(t, x) \tag{1.7}$$

defines an operator on the domain of Ω. Since Ω is densely defined, $T(t)$ (which is bounded by $1 + M^{t+1}$) can be extended to all of X by continuity. The operators $T(t)$ then form a semi-group with Ω as infinitesimal generator, which is (hypothesis (iii)) strongly continuous at the origin. The conditions (1.6) express the requirement that the initial value problem (1.1) be " well-posed " [cf. **17**].

In [**17**] it is assumed that an operator Ω is given which leads to a well-posed problem, and the operators Ω_n are required to satisfy a consistency condition which may be translated into our terminology as follows :

There exists a dense set of functions f such that $\|(\Omega_n - \Omega)T(t)f\|$ tends to zero uniformly for t restricted to a bounded interval.

Our condition is that $\lim_{n\to\infty} \Omega_n$ be densely defined, and that for some positive λ, $\lim_{n\to\infty} (\lambda - \Omega_n)$ have a dense range. (The precise meaning we attach to " limit of a sequence of operators " is given in § 2.) It is part of our conclusion that the operator to which the Ω_n converge gives rise to a well-posed problem. Although our condition appears to be quite different from that of Lax and Richtmyer, we have found no example in which we could show that one condition was satisfied and the other was not. Our condition seems to be easier to verify in the case of the applications made in § 6.

Our proof of convergence is based on the relation between a semi-group and its resolvent which is made explicit in the Hille-Yosida theorem [**14**, **22**]. In § 4 we develop the relevant facts in a form convenient for our discussion. All the results in this section are well-known ; the proof of Theorem 4.1, however, is new.

Section 5 is concerned with the convergence problem proper, and the theorems of this section represent the principal results of the paper.

In § 6 we consider in some detail the convergence of random walks to one-dimensional diffusion processes, and discuss several examples.

This paper is based on a thesis submitted to the Department of Mathematics of Princeton University in June 1956. I wish to thank Professor W. Feller for his guidance in the writing of the thesis, and the National Research Council of Canada for fellowship support during the academic year 1955–56.

2. Approximating sequences of Banach spaces. Let X be a Banach space. A sequence of Banach spaces $\{X_n\}$ together with a sequence of linear maps $P_n\colon X \to X_n$ is called a *sequence*[3] *of Banach spaces approximating* X if

[3] We could equally well define a system of Banach spaces approximating X using an arbitrary directed set as index set. The proofs of all the main theorems of this paper require only trivial changes of language to adapt them to the more general situation. Unless X is assumed to be separable, it is necessary to use nets in discussing some of the properties of the adjoint approximating sequence (§ 3).

(2.1) $$\| P_n \| \leqq 1$$

and

(2.2) $$\lim_{n\to\infty} \| P_n f \| = \| f \| \text{ for every } f \in X .$$

Condition (2.2) obviously implies that in some sense the maps P_n become isomorphisms "in the limit". The following lemma gives a precise expression to this idea.

LEMMA 2.1. *Let X' be any finite-dimensional subspace of X, and let Q_n be the restriction of P_n to X'. Then Q_n is one-to-one if n is sufficiently large, and* $\lim_{n\to\infty} \| Q_n^{-1} \| = 1$.

Proof. Take any $\varepsilon > 0$, and let $\{f_i\}$ be a finite set of vectors of unit length in X' such that the ε-neighbourhoods of the f_i cover the unit sphere in X'. Now take N sufficiently large that for all $n \geqq N$, $\max_i(\| f_i \| - \| P_n f_i \|) < \varepsilon$. Then for any g on the unit sphere in X', $\| P_n g \| > 1 - 2\varepsilon$ for all $n \geqq N$. Hence Q_n is one-to-one and

$$1 \leqq \| Q_n \|^{-1} \leqq \| Q_n^{-1} \| \leqq (1 - 2\varepsilon)^{-1} .$$

We now define convergence for sequences of vectors and operators. We shall use the following terminology and notation. By "operator on X" we shall mean "linear transformation defined on a linear subset of X and taking values in X". If A is an operator on X, the linear subset of X on which A is defined is the *domain* of A, written $\mathbf{D}(A)$. The *range* of A, denoted by $\mathbf{R}(A)$, is the linear subset consisting of all $g \in X$ for which there exists an $f \in \mathbf{D}(A)$ with $g = Af$. If A and B are two operators such than $\mathbf{D}(A) \subset \mathbf{D}(B)$ and $Af = Bf$ for all $f \in \mathbf{B}(A)$ then we call B an *extension* of A and write $A \subset B$. The identity operator on any space will be denoted by I.

A sequence $\{f_n\}$, where $f_n \in X_n$, *converges* to $f \in X$ if $\lim \| f_n - P_n f \| = 0$. It is easy to see that (2.2) implies that a sequence cannot converge to more than one $f \in X$. A sequence is *convergent* if there exists an $f \in X$ to which it converges; we call f the *limit* of $\{f_n\}$ and write $f = \lim_{n\to\infty} f_n$.

The *limit of a sequence of operators* $\{A_n\}$, where A_n is an operator on X_n, is the operator on X whose domain consists of those $f \in X$ for which $\{A_n P_n f\}$ converges and whose value for such an f is $\lim_{n\to\infty} A_n P_n f$.

EXAMPLES.

(1) Let X be an arbitrary Banach space, and for every n let $X_n = X$ and $P_n = I$. Then the convergence of vectors is ordinary

convergence, and convergence of operators is the usual strong convergence. Lemmas 2.2, 2.3, and 2.4 below then become results on the strong convergence of operators.

(2) Let X be the space of bounded continuous functions on some topological space S, with the uniform norm. For each n, let S_n be a subset of S and let X_n be some Banach space of functions on S_n (with the uniform norm) which contains all the restrictions of elements of X. For $f \in X$, define $P_n f$ to be the restriction of f to S_n. (Note that there is no requirement that the projection P_n map X *onto* X_n.) Condition (2.1) is obviously satisfied, and if the sets S_n become dense in S in the sense that every open set $U \subset S$ contains points of S_n for n sufficiently large, then (2.2) is also satisfied. In this case $\{f_n\}$ converges to f if and only if

$$\lim_{n\to\infty} \sup_{x\in S_n} |f_n(x) - f(x)| = 0 .$$

(3) Let S be a region of Euclidean space with Lebesgue measure, and let X be $L^p(S)$. For each n, let S be partitioned into measureable sets $S_{n,i}$, each with finite measure $m_{n,i} > 0$. Let X_n be the subspace of X consisting of functions constant on the $S_{n,i}$. For $f \in X$, define $P_n f$ to have the value $m_{n,i}^{-1}\int_{S_{n,i}} f(x)dx$. Condition (2.1) is always satisfied. If each partition is a refinement of the preceding one, and if the partitions become sufficiently fine in the sense that the Borel field generated by the collection of all the $S_{n,i}$ contains all Borel-measureable subsets of S, then (2.2) will be satisfied. It is clear that a similar procedure can be followed to get an approximating sequence to the L^p-space over any measure space. Essentially this type of approximation (with $p = 2$) has been used by Douglas [4].

The following lemmas will be needed for later use. Lemma 2.2 is a generalization of the Banach-Steinhaus lemma [1, p. 79], and Lemma 2.3 similarly generalizes an obvious fact about the strong convergence of operators. Lemma 2.4 is a more special result which is used in the proof of Theorem 5.2.

LEMMA 2.2. *For each n, let A_n be an operator on X_n, with domain all of X_n. If there exists a constant M such that $\|A_n\| \leqq M$ for all n, and if $A = \lim_{n\to\infty} A_n$ is densely defined, then A is defined on all of X and $\|A\| \leqq M$.*

Proof. By hypothesis, for any $f \in X$ it is possible find a sequence $\{f^j\}$ converging to f, with every f^j in $\mathbf{D}(A)$. This, by definition, means that for each f^j there exists a $g^j \in X$ such that

$\| A_n P_n f^j - P_n g^j \| \to 0$ as $n \to \infty$. For any i, j,

$$\| g^i - g^j \| = \lim_{n\to\infty} \| P_n(g^i - g^j)\| = \lim_{n\to\infty} \| A_n P_n(f^i - f^j) \| \leqq M \| f^i - f^j \| .$$

Since the f^i form a convergent sequence, it follows that the g^i form a Cauchy sequence and consequently have a limit g. Now

$$\| A_n P_n f - P_n g \| \leqq \| A_n P_n (f - f^i) \| + \| A_n P_n f^i - P_n g^i \| + \| P_n(g^i - g) \| .$$

The first and last terms on the right may be made arbitrarily small by taking i sufficiently large, independent of n, and the middle term goes to zero as $n \to \infty$. Hence Af is defined and equal to g. It is obvious that $\| A \| \leqq M$.

LEMMA 2.3. *Let $\{A_n\}$ be a sequence of operators satisfying the hypotheses of Lemma 2.2, and converging to A. For each n, let B_n be an operator on X_n and let $B = \lim_{n\to\infty} B_n$. Then $AB \subset \lim_{n\to\infty} A_n B_n$.*

Proof. Suppose $f \in \mathbf{D}(AB)$. Then

$$P_n ABf - A_n B_n P_n f = (P_n ABf - A_n P_n Bf) + A_n(P_n B - B_n P_n)f.$$

The first term on the right tends to zero because $A = \lim_{n\to\infty} A_n$. The second term on the right is dominated in norm by $M \| P_n Bf - B_n P_n f \|$ and this tends to zero because $f \in \mathbf{D}(AB) \subset \mathbf{D}(B)$.

COROLLARY. *Let $\{A_n\}$ be a sequence of operators satisfying the hypotheses of Lemma 2.2 and converging to A. Then for any positive integer k, $\lim_{n\to\infty} A_n^k = A^k$.*

LEMMA 2.4. *For each n, let A_n be an operator on X_n with an inverse B_n defined on all of X_n. Suppose that $\| B_n \| \leqq M$ for all n, and that both the domain and range of $A = \lim_{n\to\infty} A_n$ are dense in X. Then $B = \lim_{n\to\infty} B_n$ is defined on all of X, and has a dense range. B has an inverse if and only if A has a closed extension, and then B^{-1} is the closure of A.*

Proof. Consider an arbitrary $g \in \mathbf{R}(A)$, $g = Af$.
Then

$$\| B_n P_n g - P_n f \| = \| B_n(P_n g - A_n P_n f) \| \leqq M \| P_n g - A_n P_n f \| .$$

Since $g = \lim_{n\to\infty} A_n P_n f$, the term on the right goes to zero, and consequently $B_n P_n g \to f$. Since $\mathbf{R}(A)$ is dense, it follows from Lemma 2.2 that $B = \lim_{n\to\infty} B_n$ is defined on all of X. Since $Bg = f$ if $g = Af$, $BA \subset I$. Hence $\mathbf{R}(B) \supset \mathbf{D}(A)$ and is dense in X.

If B^{-1} exists it is a closed operator because B is closed; since $BA \subset I$, it is an extension of A. Conversely, suppose A has a closed extension A'. Then, since B is bounded, $A'B$ is a closed extension of AB. But for $g = Af \in \mathbf{R}(A)$, we have $ABg = AB(Af) = A(BA)f = Af = g$, so that AB coincides with I on a dense set. Hence $A'B = I$; therefore B has an inverse, and A' is an extension of it. This is true for *any* closed extension of A, and consequently B^{-1} is the closure of A.

3. Adjoint spaces. If $\{X_n\}$ is a sequence approximating X, with associated projection operators P_n, then the adjoint spaces X_n^* are connected with X_n^* by the adjoint operators $P_n^* : X_n^* \to X^*$. Condition (2.1) of course implies that $\|P_n^*\| \leqq 1$. Condition (2.2) clearly should imply that in some sense the images of the P_n^* become dense in X^* in the limit. This is not generally true in terms of the norm topology on X^*—it is easy to give an example in which the union of the images of the P_n^* is nowhere dense in X^* with respect to this topology. It is more appropriate to consider the weak topology [1] (often called the weak* topology) on X^*. For any $f^* \in X^*, f_1, \cdots, f_m \in X$, and $\varepsilon > 0$ we write $V_{f^*}(f_1, \cdots, f_m; \varepsilon)$ for the set $\{g^* \in X^* : |f^*(f_i) - g^*(f_i)| < \varepsilon, i = 1, 2, \cdots, m\}$. The collection of the "cubical neighourhoods" forms a base for the neighbourhoods of f^* in the weak topology.

LEMMA 3.1. *Let f^* be any element of X^* and $V = V_{f^*}(f_1, \cdots, f_m; \varepsilon)$ a cubical neighbourhood of it. Then for all sufficiently large n, there exists an $f_n^* \in X_n^*$, with $\|f_n^*\| \leqq 2\|f^*\|$ and $f_n^* P_n \in V$.*[4]

Proof. Let X' be the subspace of X spanned by $f_1, \cdots, f_m$, and as in Lemma 2.1, let Q_n be the restriction of P_n to X'. Let $f^{*\prime}$ be the restriction of f^* to X'. Let n be sufficiently large that Q_n^{-1} exists, and define $f_n^* = f^{*\prime} Q_n^{-1}$. It follows directly from the definitions that $f_n^* P_n f_i = f^*(f_i)$ for any one of the f_i, so that $f_n^* P_n \in V$. Since $\| Q_n^{-1} \| \to 1$, the condition on the norm of f_n^* is satisfied for sufficiently large n.

If $\{f_\alpha^*\}$ is a net with $f_\alpha^* \in X_{n_\alpha}^*$ for each α, we say that $\{f_\alpha^*\}$ *converges weakly to* f^* if $n_\alpha \to \infty$ and $\{f_\alpha^* P_{n\alpha}\}$ converges weakly to f^*. The net $\{f_\alpha^*\}$ is said to be *bounded* if $\|f_\alpha^*\|$ is bounded uniformly with

[4] We write adjoint operators on the right when applied to a vector (dropping the "star" since the position of the operator indicates that it is the adjoint).

respect to α.

From Lemma 3.1 we obtain at once the proposition: for every $f^* \in X^*$ there exists a bounded net $\{f^*_\alpha\}$ with $f^*_\alpha \in X^*_{n_\alpha}$ which converges weakly to it. To construct such a net, let the index set consist of the cubical neighbourhoods of f^*, ordered by inclusion. For each such neighbourhood α, we can pick an $f^*_\alpha \in X^*_{n_\alpha}$, $n_\alpha > \varepsilon^{-1}$ (where ε is the positive number used to define the cubical neighbourhood α), with $\|f^*_\alpha\| \leqq 2\|f^*\|$ and $f^*_\alpha P_{n_\alpha} \in \alpha$. This net converges to f^*. If X is *separable* there exists a uniformly bounded *sequence* $\{f^*_n\}$ with $f^*_n \in X^*_n$ which converges weakly to f^*. To show this, let $f_1, f_2, \cdots$ be a sequence which is dense in X, and let X'_m be the subspace of X spanned by the first m vectors of the sequence. Using the construction of Lemma 3.1 we can define N_m inductively so that $N_{m+1} > N_m$, and for $N_m \leqq n < N_{m+1}$ there exists an $f^*_n \in X^*_n$ with $\|f^*_n\| \leqq 2\|f^*\|$ and $f^*_n P_n f_i = f^*(f_i)$ for $i < m$. Then $\|f^*_n P_n\|$ is uniformly bounded and $\lim_{n\to\infty} f^*_n P_n f_i = f^*(f_i)$ for all the f_i. Hence by a theorem of Banach [1, Theorem 2 of Chapter 7], $\{f^*_n P_n\}$ converges weakly to f^*.

For each n let B_n be an operator on X^*_n. The *weak limit* of the sequence $\{B_n\}$ is defined to be an operator B on X^*, with domain consisting of all $f^* \in X^*$ such that for every bounded net $\{f^*_\alpha\}$ with $f^*_\alpha \in X^*_{n_\alpha}$ which converges weakly to f^*, the net $\{f^*_\alpha B_{n_\alpha} P_{n_\alpha}\}$ converges weakly to a unique limit. For $f^* \in \mathbf{D}(B)$, f^*B is defined to be this limit.

LEMMA 3.2. *For each n let A_n be a bounded operator on X_n with domain all of X_n, and suppose that $A = \lim A_n$ is defined on all of X. Then the adjoint operators A^*_n converge weakly to A^*.*

Proof. Take any $f^* \in X^*$, and let $\{f^*_\alpha\}$ be any bounded net converging to it weakly. We must show that for every $f \in X$. $\{f^*_\alpha A_{n_\alpha} P_{n_\alpha} f\}$ converges to f^*Af. We have

$$f^*_\alpha A_{n_\alpha} P_{n_\alpha} f - f^*Af = f^*_\alpha(A_{n_\alpha} P_{n_\alpha} - P_{n_\alpha} A)f + (f^*_\alpha P_{n_\alpha} - f^*)Af .$$

The first term on the right goes to zero because $n_\alpha \to \infty$, $A = \lim_{n\to\infty} A_n$, and the f^*_α are uniformly bounded. The second term on the right goes to zero because $\{f^*_\alpha\}$ converges weakly to f^*.

4. Semi-groups and resolvents. Throughout this section we shall be dealing with a fixed Banach space X. Convergence of operators is to be interpreted as strong convergence.

A (*one-parameter*) *semi-group of operators* is a family $\{T(t)\}$ of bounded operators on X, t ranging over the non-negative real numbers, which satisfies the relation

$$T(t+s) = T(t)T(s) \quad t, s \geqq 0 , \tag{4.1}$$

with $T(0)$ defined to be I. We shall consider only semi-groups for which

$$\lim_{t\to 0} T(t) = I \tag{4.2}$$

and

$$\| T(t) \| \leqq M \qquad \text{all } t \tag{4.3}$$

for some constant M. Semi-groups satisfying (4.2) and (4.3) will be called *proper*.[5]

The operator

$$\Omega = \lim_{t\to 0} t^{-1}(T(t) - I) \tag{4.4}$$

is the *infinitesimal generator* of the semi-group and is always closed and densely defined. One has $g = \Omega f$ if and only if

$$T(t)f = f + \int_0^t T(s)g\, ds . \tag{4.5}$$

For every $\lambda > 0$ there is a bounded operator

$$J(\lambda) = (\lambda - \Omega)^{-1} = \int_0^\infty e^{-\lambda t}T(t)\, dt . \tag{4.6}$$

The family $\{J(\lambda)\}$ is called *the resolvent family of* $T(t)$. The operators $J(\lambda)$ satisfy the relations

$$J(\lambda) - J(\mu) = (\mu - \lambda)J(\lambda)J(\mu) \tag{4.7}$$

$$\| \lambda^m J^m(\lambda) \| \leqq M \qquad \lambda > 0, m = 1, 2, \cdots \tag{4.8}$$

$$\lim_{\lambda\to\infty} \lambda J(\lambda) = I . \tag{4.9}$$

A family of operators satisfying (4.7) will be called a *resolvent family*, and one which satisfies (4.8) and (4.9) as well, a *proper* resolvent family.

If $\{J(\lambda)\}$ is a resolvent family it is clear that any f annihilated by one of the $J(\lambda)$ is annihilated by all of them, and also that $\mathbf{R}(J(\lambda))$ is independent of λ. These remarks, together with (4.9), show that *the operators of a proper resolvent family are one-to-one transformations with*

[5] Condition (4.3) is not a serious restriction. If a semi-group satisfies (4.2) then $\| T(t) \|$ must be bounded for t near zero. It is then well-known **[14]** that $\| T(t) \| \leqq Me^{Kt}$ for some K, so that the closely related semi-group $T'(t) = e^{-Kt}T(t)$ will be proper. We use this trick in proving Theorems 5.2 and 5.3.

dense range. Hence for each λ, $\{J(\lambda)\}^{-1}$ is a densely defined operator. It is easy to show from (4.7) that

$$\Omega = \lambda - \{J(\lambda)\}^{-1} \tag{4.10}$$

is independent of λ; we call it *the infinitesimal generator associated with the resolvent family* $\{J(\lambda)\}$.

A proper semi-group is uniquely determined by its resolvent family. Suppose $\{T(t)\}$ and $\{T'(t)\}$ are two proper semi-groups with the same resolvent family $\{J(\lambda)\}$. For any $f \in X$, $f^* \in X^*$ the function $f^*(J(\lambda)f)$ will be the Laplace transform of both $f^*(T(t)f)$ and $f(T'(t)f)$. The classical uniqueness theorem for the Laplace transform [**21**, p. 63] then implies the identity of the last two functions (since both are bounded and continuous), and since f and f^* are arbitrary it follows that $T(t) = T'(t)$. This fact shows that $\{J(\lambda)\}$ is the resolvent family of the semi-group $T(t)$ if and only if the operator Ω defined by (4.10) is the same as that defined by (4.4).

The question still remains as to whether every proper resolvent family is the resolvent family of some proper semi-group. The Hille-Yosida theorem [**14**, **22**] provides an affirmative answer. Our next theorem gives an expression for the semi-group in terms of the resolvent. (cf. [**14**, p. 234]; the proof given there depends on the Post-Widder inversion formula for the Laplace transform.)

THEOREM 4.1. *Let $\{J(\lambda)\}$ be a proper resolvent family. Then the operators*

$$T(t) = \lim_{\lambda\to\infty} \{\lambda J(\lambda)\}^{[\lambda t]} \qquad t \geqq 0$$

are defined on all of X and form a proper semi-group which has $\{J(\lambda)\}$ as its resolvent family.

We first prove several lemmas.

LEMMA 4.1 *Let A and B be two operators which commute and have the property that $\| A^i \|, \| B^i \| \leqq M$ for all positive integers i.*

Then, for any f,

$$\| (A^n - B^n)f \| \leqq nM^2 \| (A - B)f \| .$$

Proof. Since A and B commute

$$(A^n - B^n)f = \sum_{i=0}^{n-1} A^{n-i-1}B^i(A - B)f .$$

The right hand side contains n terms, each with norm less than or equal to $M^2 \| (A - B)f \|$.

LEMMA 4.2. *For any* $f \in \mathbf{D}(\Omega)$ (where Ω is defined by (4.10))

$$\| (\lambda J(\lambda) - I)f \| \leqq \lambda^{-1} M \| \Omega f \| .$$

Proof. This follows from (4.8) since for $f \in \mathbf{D}(\Omega)$,

$$(\lambda J(\lambda) - I)f = J(\lambda)\, \Omega f . \tag{4.11}$$

LEMMA 4.3. *For any* $f \in \mathbf{D}(\Omega^2)$

$$\| \{\lambda J(\lambda) - (2\lambda J(2\lambda))^2\} f \| \leqq (2\lambda)^{-2} M^2 \| \Omega^2 f \| .$$

Proof. Note that

$$J(\lambda) - J(2\lambda) = \lambda J(\lambda) J(2\lambda) \tag{4.12}$$

is a special case of (4.7). Now

$$\begin{aligned}
\lambda J(\lambda)f - (2\lambda J(2\lambda))^2 f &= (\lambda J(\lambda) - I)f - (2\lambda J(2\lambda) + I)(2\lambda J(2\lambda) - I)f \\
&= J(\lambda)\Omega f - (2\lambda J(2\lambda) + I) J(2\lambda)\Omega f && \text{by (4.11)} \\
&= \{J(\lambda) - J(2\lambda) - 2\lambda J^2(2\lambda)\}\Omega f \\
&= J(2\lambda)\{\lambda J(\lambda) - 2\lambda J(2\lambda)\}\Omega f && \text{by (4.12)} \\
&= J(2\lambda)\{\lambda J(\lambda) - I - 2\lambda J(2\lambda) + I\}\Omega f \\
&= J(2\lambda)\{J(\lambda) - J(2\lambda)\}\Omega^2 f && \text{by (4.11)} \\
&= \lambda J(\lambda) J^2(2\lambda)\Omega^2 f && \text{by (4.12)} .
\end{aligned}$$

The conclusion follows from this identity and (4.8).

Proof of Theorem 4.1. Let r be an arbitrary positive number which will be assumed fixed throughout the following discussion. Write r_n as an abbreviation for $2^n r$. Define

$$T(n, t) = \{r_n J(r_n)\}^{[t r_n]} . \tag{4.13}$$

From the definition we have

$$\begin{aligned}
T(n+1, t) &= \{2r_n J(2r_n)\}^{[2t r_n]} \\
&= \{2r_n J(2r_n)\}^{2[t r_n] + \varepsilon}
\end{aligned}$$

where $\varepsilon = [2tr_n] - 2[tr_n] = 0$ or 1. Thus

$$\begin{aligned}
T(n+1, t)f - T(n, t)f = {} & \{2r_n J(2r_n)\}^{2[t r_n]}(\{2r_n J(2r_n)\}^{\varepsilon} - I)f \\
& + (\{2r_n J(2r_n)\}^{2[t r_n]} - \{r_n J(r_n)\}^{[t r_n]})f .
\end{aligned}$$

Estimating the first term on the right by Lemma 4.2 and the second term by Lemmas 4.1 and 4.3, we obtain, for $f \in \mathbf{D}(\Omega^2)$,

$$(4.14)\quad \|\{T(n+m,t)-T(n,t)\}f\| \leqq \sum_{k=0}^{m-1} \|\{T(n+k+1,t)-T(n+k,t)\}f\| \leqq \sum_{k=0}^{\infty} r^{-1}2^{-n-k}(M\|\Omega f\| + tM^4\|\Omega^2 f\|) \leqq 2r_n^{-1}(M\|\Omega f\| + tM^4\|\Omega^2 f\|) .$$

Since the right-hand side goes to zero as $n \to \infty$, the vectors $T(n,t)f$ form a Cauchy sequence and therefore converge. It should be remarked that the convergence is uniform over every bounded t-interval. $\mathbf{D}(\Omega^2)$ is dense in X since it includes $\mathbf{R}(J^2(\lambda))$, and the latter is dense because of (4.9). From (4.8), $\|T(n,t)\| \leqq M$ for all n. Hence, by the Banach-Steinhaus theorem

$$(4.15)\qquad T(t) = \lim_{n\to\infty} T(n,t)$$

is everywhere defined and satisfies (4.3). Taking $n=0$ in (4.14) and letting $m \to \infty$, we obtain the estimate

$$(4.16)\qquad \|\{rJ(r)\}^{[tr]}f - T(t)f\| \leqq 2r^{-1}(M\|\Omega f\| + tM^4\|\Omega^2 f\|$$

for any $f \in \mathbf{D}(\Omega^2)$.

For any n,

$$T(n, s+t) - T(n,s)T(n,t) = T(n,s)T(n,t)(\{r_n J(r_n)\}^{\epsilon} - I) ,$$

where $\varepsilon = [r_n(s+t)] - [r_n s] - [r_n t] = 0$ or 1.
Consequently (Lemma 4.2)

$$(4.17)\qquad \|\{T(n,s+t) - T(n,s)T(n,t)\}f\| \leqq r_n^{-1}M\|\Omega f\|$$

for any $f \in \mathbf{D}(\Omega)$. Taking the limit as $n \to \infty$ in (4.17) shows that the operators defined by (4.15) satisfy (4.1) on a dense subset of X, and hence on all of X, by continuity.

From Lemmas 4.1 and 4.2 we get

$$(4.18)\qquad \|\{T(n,t) - I\}f\| \leqq r_n^{-1}[r_n t]M^3\|\Omega f\| \leqq tM^3\|\Omega f\| .$$

This inequality must hold also in the limit and shows that $\lim_{t=0} T(t)f = f$ for all f in the dense set $\mathbf{D}(\Omega)$. By the Banach-Steinhaus theorem it follows that the operators defined by (4.15) satisfy (4.2).

This completes the proof that the operators $T(t)$ which we have constructed form a proper semi-group. The construction, however, depended on the choice of a number r, and we still have to show that the result is independent of this choice. We shall show that the semi-group has the given family $\{J(\lambda)\}$ as its resolvent family, and since the resolvent family of a semi-group determines it uniquely, it will follow that the result of our construction is independent of r.

We show that $\{T(t)\}$ has the original $\{J(\lambda)\}$ as resolvent family by demonstrating that the operator Ω defined by (4.10) is the infinitesimal generator of the semi-group as defined by (4.4). Suppose $f \in \mathbf{D}(\Omega)$ and $g = \Omega f$. For convenience, let t be such that $r_n t$ is an integer for sufficiently large n. Such values of t are dense in the line. Then

$$\begin{aligned}\int_0^t T(n, s)g\,ds &= \int_0^t \{r_n J(r_n)\}^{[r_n s]}\,\Omega f\,ds \\ &= r_n^{-1} \sum_{k=1}^{k=tr_n} \{r_n J(r_n)\}^{k-1}\Omega f \\ &= \sum_{k=1}^{k=tr_n} \{r_n J(r_n)\}^{k-1} J(r_n)\Omega f \\ &\quad + r_n^{-1}\{\Omega f - T(n, t)\Omega f\} \\ &= \sum_{k=1}^{k=tr_n} \{r_n J(r_n)\}^{k-1}\{r_n J(r_n) - I\} f \\ &\quad + r_n^{-1}\{\Omega f - T(n, t)\Omega f\} \\ &= T(n, t)f - f + r_n^{-1}\{\Omega f - T(n, t)\Omega f\}\ .\end{aligned}$$

Letting $n \to \infty$ we obtain formula (4.5). (Passage to the limit under the integral sign is justified since $T(n, t)f$ converges uniformly on every bounded t-interval.) The number t was any one of a dense set, and by continuity, (4.5) must in fact hold for all t. This shows that $g = \Omega f$ with Ω defined by (4.4), and the proof of Theorem 4.1 is complete.

5. Convergence of Semi-groups. Throughout this section $\{X_n\}$ will be a sequence of Banach spaces approximating X, with associated projections P_n. We shall use the notational convention that vectors with subscript n are elements of X_n, and operators with subscript n are operators on X_n; vectors and operators without subscript will be associated with X.

A sequence of semi-groups $\{T_n(t)\}$ or resolvent families $\{J^n(\lambda)\}$ will be said to be *uniformly proper* if each member of the sequence is proper, and the constant M in conditions (4.3) or (4.8) may be taken independent of n.

THEOREM 5.1. *Let $\{T_n(t)\}$ be a uniformly proper sequence of semi-groups, and $\{J_n(\lambda)\}$ the sequence of associated resolvent families. Then if the operators $J(\lambda) = \lim_{n\to\infty} J_n(\lambda)$ form a proper resolvent family, the sequence $\{T_n(t)\}$ converges to $T(t)$, the proper semi-group having $\{J(\lambda)\}$ as resolvent family.*

Proof. Since the $T_n(t)$ are uniformly bounded, it will be sufficient, by Lemma 2.2, to prove that $T_n(t)f$ converges to $T(t)f$ for a set of f which is dense in X. It has already been remarked (in the paragraph following (4.9)) that any operator in a proper resolvent family has a dense range. By similar considerations it is easy to show that the square of any operator in a proper resolvent family has a dense range. Hence we need only consider f such that $f = J^2(\mu)g$ for some g and μ. Define $g_n = P_n g$ and $f_n = J_n^2(\mu)g_n$. Letting Ω_n be the infinitesimal generator of $T_n(t)$ we have

$$\begin{aligned}(5.1)\qquad \| \Omega_n f_n \| &= \| (\mu J_n(\mu) - I)J_n(\mu)g_n \| \\ &\leqq \mu^{-1}(M+1) \| g_n \| \leqq \mu^{-1}(M+1) \| g \| \leqq K\end{aligned}$$

and

$$\begin{aligned}(5.2)\qquad \| \Omega_n^2 f_n \| &= \| (\mu J_n(\mu) - I)^2 g_n \| \\ &\leqq (M+1)^2 \| g_n \| \leqq (M+1)^2 \| g \| \leqq K\end{aligned}$$

for some sufficiently large constant K; we also have $\| \Omega f \|, \|\Omega^2 f \| \leqq K$. Now

$$\begin{aligned}P_n T(t)f - T_n(t)P_n f = P_n(T(t) &- \{rJ(r)\}^{[rt]})f \\ &+ P_n\{rJ(r)\}^{[rt]}f - \{rJ_n(r)\}^{[rt]}P_n f \\ &+ \{rJ_n(r)\}^{[rt]}(P_n f - f_n) \\ &+ (\{rJ_n(r)\}^{[rt]} - T_n(t))f_n \\ &+ T_n(t)(f_n - P_n f).\end{aligned}$$

Applying (4.16), (5.1), (5.2) and the uniform boundedness of the $T_n(t)$, this yields

$$\begin{aligned}\| P_n T(t)f - Tn(t)P_n f \| &\leqq 4r^{-1}K(M + tM^4) + 2M \| P_n f - f_n \| \\ &+ \| P_n\{rJ(r)\}^{[rt]}f - \{rJ_n(r)\}^{[rt]}P_n f \| .\end{aligned}$$

For any fixed r, the last two terms go to zero as $n \to \infty$, because $f_n \to f$ and $\{rJ(r)\}^k = \lim_{n\to\infty} \{rJ_n(r)\}^k$ for any fixed k. Thus

$$\limsup_{n\to\infty} \| P_n T(t)f - T_n(t)P_n f \| \leqq 4r^{-1}K(M + tM^4) .$$

Since r may be taken arbitrarily large it follows that $T(t)f = \lim_{n\to\infty} T_n(t)P_n f$.

LEMMA 5.1. *Let $\{J_n(\lambda)\}$ be a uniformly proper sequence of resolvent families, such that for some positive μ, $\lim_{n\to\infty} J_n(\mu)$ is densely defined and has a dense range. Then for every λ, $J(\lambda) = \lim_{n\to\infty} J_n(\lambda)$ is defined on all of X, and $\{J(\lambda)\}$ is a proper resolvent family.*

Proof. That $J(\mu)$ is everywhere defined follows immediately from Lemma 2.2. To show that $\lim_{n\to\infty} J_n(\lambda)$ is everywhere defined we make use of the relation [14, p. 119]

$$J_n(\mu - \nu) = \sum_{k=1}^{\infty} \nu^{k-1} J_n^k(\mu) \tag{5.3}$$

where, provided that $|\nu| < \mu$, the series converges in the uniform operator norm, uniformly in n. This follows from the formula

$$J_n(\mu - \nu) = \nu^m J_n^m(\mu) J_n(\mu\, \nu) + \sum_{k=1}^{m} \nu^{k-1} J_n^k(\mu)$$

which is easily derived from (4.7) by induction on m, condition (4.8), and the assumption that the $J_n(\lambda)$ are uniformly proper. For each k, $\lim_{n\to\infty} J_n^k(\mu)$ is everywhere defined, by the corollary to Lemma 2.3, and the uniform convergence of (5.3) implies that $\lim_{n\to\infty} J_n(\lambda)$ is everywhere defined for $2\mu > \lambda > 0$. Repetition of the argument, replacing μ by, say, $3\mu/2$, shows the convergence of $\{J_n(\lambda)\}$ for $3\mu > \lambda > 0$. By further repetitions of the argument it can be shown that $J(\lambda) = \lim_{n\to\infty} J_n(\lambda)$ is everywhere defined for all $\lambda > 0$. Relation (4.7) holds for each n, and by Lemma 2.3 it continues to hold in the limit. Condition (4.8) is clearly satisfied by every $J(\lambda)$. To complete the proof it is only necessary to demonstrate (4.9). Since any $f \in \mathbf{R}(J(\mu))$ is in $\mathbf{D}(\Omega)$, where Ω is defined by (4.10), it follows from Lemma 4.2 that (4.9) holds on $\mathbf{R}(J(\mu))$, which is dense by hypothesis. Since the operators $\lambda J(\lambda)$ are uniformly bounded, the conclusion follows by the Banach-Steinhaus theorem.

THEOREM 5.2. *Let $\{T_n(t)\}$ be a sequence of semi-groups satisfying* (4.2) *and the stability condition*

$$\| T_n(t) = \| \leqq Me^{Kt}$$

where M and K are independent of n and t. Let Ω_n be the infinitesimal generator of $T_n(t)$ and define $\Omega = \lim_{n\to\infty} \Omega_n$.

Suppose that

(i) *Ω is densely defined*

(ii) *for some $\lambda > K$, $\mathbf{R}(\lambda - \Omega)$ is dense in X.*

Then the closure of Ω is the infinitesimal generator of a semi-group $T(t)$ which satisfies (4.2), *and $T(t) = \lim_{n\to\infty} T_n(t)$.*

Proof. Define $T'_n(t) = e^{-Kt}T_n(t)$ and $\Omega'_n = \Omega_n - K$. Then Ω'_n is the infinitesimal generator of $T'_n(t)$ and the semi-groups $T'_n(t)$ from a uniformly proper sequence. Also $\Omega' = \lim_{n\to\infty} \Omega'_n$ is densely defined and $\mathbf{R}(\lambda - K - \Omega')$ is dense in X. The sequence $J'_n(\lambda - K) = (\lambda - K - \Omega'_n)^{-1}$ is uniformly bounded, From Lemma 2.4 it follows that $J'(\lambda - K)$ is everywhere defined and has a dense range. By Lemma 5.1, the operators $J'(\lambda) = \lim_{n\to\infty} J'_n(\lambda)$ form a proper resolvent family, which by Theorem 5.1 is the resolvent family of a semi-group $T'(t)$ such that $T'(t) = \lim_{n\to\infty} T'_n(t)$. Since $J'(\lambda)$ has an inverse, this inverse (by Lemma 2.4) is the closure of $\lambda - K - \Omega'$, and it follows that the closure of Ω' is the infinitesimal generator of $T'(t)$. The results stated for $T(t)$ follow immediately from what we have proved about $T'(t)$.

LEMMA 5.2. *Let h be a given positive number, and T an operator such that $\| T^k \| \leqq M$ for all n. Then $\Omega = h^{-1}(T - I)$ is the infinitesimal generator of a semi-group $S(t)$ such that $\| S(t) \| \leqq M$.*

Proof. Define

$$S(t) = \sum_{k=0}^{\infty} (k\,!)^{-1}(t\Omega)^k = e^{-th^{-1}} \sum_{k=0}^{\infty} (k\,!)^{-1}(th^{-1}T)^k .$$

That Ω is the infinitesimal generator of $S(t)$ can be verified by term-by-term differentiation of the first expression given. From the second expression we obtain

$$\| S(t) \| \leqq e^{-th^{-1}} \sum_{k=0}^{\infty} (k\,!)^{-1}(th^{-1})^k M \leqq M.$$

LEMMA 5.3. *Let h, T, Ω and $S(t)$ be as in Lemma 5.2. For a fixed t, let $k = [th^{-1}]$. Then for any f,*

$$\| S(t)f - T^k f \| \leqq hM^2(\tfrac{1}{2}t \| \Omega^2 f \| + \| \Omega f \|) .$$

Proof. Iteration of (4.5) gives

$$\begin{aligned} S(h)f &= f + \int_0^h S(t)\Omega f\, dt \\ &= f + h\Omega f + \int_0^h\int_0^s S(t)\Omega^2 f\; dtds \\ &= Tf + \int_0^h\int_0^s S(t)\Omega^2 f\; dtds . \end{aligned}$$

Hence $\| S(h)f - Tf \| \leqq \frac{1}{2}h^2 \| \Omega^2 f \|$. By Lemma 4.1

$$\| S(kh)f - T^k f\| \leqq \tfrac{1}{2}M^2 kh^2 \| \Omega^2 f \| \leqq \tfrac{1}{2}M^2 ht \| \Omega^2 f \|$$

when $k=[th^{-1}]$. Also

$$\| S(t)f - S(kh)f \| = \| S(kh)\{S(t - kh) - I\} f \| \leqq hM^2 \| \Omega f \| .$$

THEOREM 5.3. *Let $\{h_n\}$ be a sequence of positive numbers converging to zero, and $\{T_n\}$ a sequence of operators satisfying the stability condition*

$$\| T_n^k \| \leqq Me^{Kkh_n}$$

where M and K are constants independent of n and k. Let $\Omega_n = h_n^{-1}(T_n - I)$ and define $\Omega = \lim_{n\to\infty} \Omega_n$. Suppose that

(i) *Ω is densely defined*

(ii) *for some $\lambda > K$, $R(\lambda - \Omega)$ is dense in X.*

Then the closure of Ω is the infinitesimal generator of a semi-group $T(t)$ and

$$T(t) = \lim_{n\to\infty} T_n[th_n^{-}] .$$

Proof. Define $T'_n = e^{-Kh_n}T_n$, so that for any k, $\| T'^k_n \| \leqq M$. Then

$$\Omega'_n = h_n^{-1}(T'_n - I) = e^{-Kh_n}\Omega_n - h_n^{-1}(1 - e^{-Kh_n})$$

and

$$\Omega' = \lim_{n\to\infty} \Omega'_n = \Omega - K .$$

Hence Ω' is densely defined and $\mathbf{R}(\lambda - K - \Omega')$ is dense. Let $S_n(t)$ be the semi-group generated by Ω'_n. Lemma 5.2 shows that the hypotheses of Theorem 5.2 are satisfied and hence there is a proper semi-group $T'(t) = \lim_{n\to\infty} S_n(t)$, with infinitesimal generator the closure of Ω'. For a fixed t, define $k_n = [th_n^{-1}]$. We shall show that $\lim_{n\to\infty} T'^{k_n}_n = T'(t)$. As in the proof of Theorem 5.1 it suffices to show that $T'(t)f = \lim_{n\to\infty} T'^{k_n}_n P_n f$ for f of the form $f = J^2(\lambda)g$ for some g and λ. Define $f_n = J_n^2(\lambda)P_n g$; then as in the proof of Theorem 5.1 there exists a constant C such that $\| \Omega_n^2 f_n \|$, $\| \Omega_n f_n \| \leqq C$ for all n. Then

$$\begin{aligned} \| S_n(t) - T'^{k_n}_n)P_n f \| \leqq S_n(t)(P_n f - f_n) \| \\ + \| (S_n(t) - T'^{k_n}_n) f_n \| \\ + \| T'^{k_n}_n (f_n - P_n f) \| \\ \leqq 2M \| f_n - P_n f \| + h_n M^2 C(t+1) . \end{aligned}$$

Since $f_n \to f$ and $h_n \to 0$ this shows that

$$\lim_{n\to\infty} T'^{k_n}_n P_n f = \lim_{n\to\infty} S_n(t)P_n f = T'(t) .$$

Hence

$$\lim T_n^{k_n} = \lim e^{Kh_nk_n} T'^{k_n}_n = e^{Kt}T'(t) ,$$

which is a semi-group $T(t)$ with the closure of $\Omega' + K = \Omega$ as infinitesimal generator.

6. Random walks and diffusion processes. Throughout this section, S will denote either the compactified real line $[-\infty, \infty]$, or some closed subinterval of it. Let S' be any Borel subset of S, and suppose that for every $x \in S'$ a positive Borel measure μ_x is given such that

(i) $\mu_x(S) \leqq 1$

(ii) $\mu_x(S - S') = 0$

(iii) $\mu_x(A)$ is a Borel function of x for every Borel set A.

We consider a particle which executes a *random walk* in the following way. Let h be a positive constant. If the particle is at a point $x \in S'$ at time kh (k a non-negative integer) then it remains at x during the interval $[kh, (k+1)h)$ and at time $(k+1)h$ "jumps" so that the probability that it goes to any Borel set A is $\mu_x(A)$. (If $\mu_x(S) < 1$ then the particle disappears with probability $1 - \mu_x(S)$.) The number h is the *basic time-interval* of the random walk, and the measures μ_x are the *one-step transition probability distributions*. The set S' is the *support* of the random walk. If a probability distribution ν on S' for the initial position of the particle is fixed, then the random walk gives rise to what is essentially a discrete parameter Markov process with stationary transition probabilities. For rigorous definition and further details [3, p. 190 ff.] may be consulted.

For any bounded Borel function f on S', the random walk determines a new function Tf defined by setting

$$(Tf)(x) = \int_{S'} f(y)\, d\mu_x(y) \qquad x \in S'.$$

Condition (iii) implies that Tf is again a Borel function. The *one-step transition operator* T is obviously linear and positivity preserving. The space of bounded Borel functions on S' is a Banach space under the norm $\|f\| = \sup_{x \in S'} |f(x)|$, and $\|T\| \leqq 1$ relative to this norm. The adjoint transformation may be considered as acting on the Borel measures on S', and for any such measure ν, νT is given by

$$(\nu T)(A) = \int T\chi_A \, d\nu$$

where A is an arbitrary Borel set and χ_A is its characteristic function. If ν is a probability measure giving the distribution of the initial position of the particle then νT^k is the distribution for its position after k jumps [cf. 3, p. 191]. Thus a random walk is completely characterized by its support, its basic time-interval, and its one-step transition operator, and we may speak of "the random walk $\{S', h, T\}$".

A *diffusion process* (with stationary transition probabilities) on S is characterized by giving for each $x \in S$ and $t > 0$ a measure $\mu_{x,t}$ such that $\mu_{x,t}(S) \leqq 1$, which is to be interpreted as the probability distribution for the position at time $s + t$ of a particle which is at x at time s. To begin with, it is necessary to have $\mu_{x,t}(A)$ a Borel function of x for every t and every Borel set A [3, p. 255]. Then a family of operators $T(t)$ on the space of bounded Borel functions can be defined by setting

$$T(t)f(x) = \int f(y)\, d\mu_{x,t}(y) .$$

As in the case of a random walk, we may consider the adjoint transformation; for ν a Borel measure on S

$$(\nu T(t))(A) = \int T(t)\chi_A\, d\nu$$

for every Borel set A. If ν gives the distribution for the initial position of the particle then $\nu T(t)$ gives the distribution for its position at time t. The further conditions which the measures $\mu_{x,t}$ must satisfy are most easily expressed in terms of the operators $T(t)$, which we call the *transition operators* of the process. In the first place, we impose the condition that $\mu_{x,t}$ be a continuous function of x with respect to the weak topology on the measures; this is equivalent to requiring that $T(t)f$ be continuous when f is continuous (cf. [11]). Secondly, we would like to require that for any continuous f, $T(t)f$ converge uniformly to f as $t \to 0$, but certain processes (those with absorbing barriers) do not satisfy this condition. One or both end-points of S may be absorbing barriers; intuitively speaking, an end-point is an absorbing barrier if a diffusing particle disappears immediately when it reaches that point. In such a case, the range of $T(t)$ contains only functions which vanish at the absorbing barrier(s). Throughout the rest of this section, let X be the Banach space of *continuous functions on S which vanish at those end-points which are absorbing barriers* for the diffusion process under consideration. We shall require that $T(t)f$ converge uniformly to f as $t \to 0$ for all $f \in X$. For the process to have the Markov property, the operators $T(t)$ must have the semi-group property (4.1). It is obvious that $\| T(t) \| \leqq 1$. The conditions imposed so far may be summarized as:

The operators $T(t)$ form a proper semi-group of operators on the space X.

Finally, in order to restrict attention to diffusion processes rather than more general types of Markov process, we suppose that

The infinitesimal generator of the semi-group $T(t)$ is a restriction of an operator of the form

$$\Omega' = \frac{d}{dm}\frac{d}{dx} \tag{6.1}$$

where m is a strictly increasing function of x.

It has been shown by Feller [10, 11] that all one-dimensional diffusion processes satisfying certain regularity conditions have associated semi-groups whose infinitesimal generators can be put into the required form by a suitable choice for the coordinate function x.

Suppose that for each n a random walk $\{S_n, h_n, T_n\}$ is given, and that a diffusion process with transition semi-group $T(t)$ is also given. We shall say that the sequence of random walks *converges* to the diffusion process if

(i) For every probability measure ν on S, there exists a sequence of probability measures ν_n converging weakly to it, with the support of ν_n contained in S_n.

(ii) For every such sequence ν_n, and every $t > 0$, the probability distribution for the position at time t of a particle starting with initial distribution ν_n and executing the nth random walk converges weakly to the probability distribution at time t for the position of a particle starting with initial distribution ν and following the diffusion process.
The weak convergence referred to above is *weak convergence relative to the space* X.

Now let X_n be the Banach space of bounded Borel functions on S_n, and define $P_n : X \to X_n$ by taking $P_n f$ to be the restriction of f to S_n for any $f \in X$. Suppose that the S_n become dense in S, as in Example 2, § 2; then the X_n form a sequence of Banach spaces approximating X. The adjoint space to X consists of the Borel measures on S (with the exception of measures giving non-zero mass to end-points which are absorbing barriers). The Borel measures with support included in S_n may be considered as elements of X_n^*, and for any such measure ν_n, $\nu_n P_n = \nu_n$. It is clear that condition (i) will be satisfied. Condition (ii) may be restated as follows. For any sequence of probability measures ν_n converging weakly to ν, $\nu_n T_n^{[th_n^{-1}]}$ converges weakly to $\nu T(t)$. Thus as a direct consequence of the definition of weak limit in § 3, (ii) will be satisfied if $T^*(t)$ is the weak limit of $T_n^{*[th_n^{-1}]}$. Finally, appealing to Lemma 3.2, we obtain the following *sufficient* condition for convergence.

In order that a sequence of random walks converge to a diffusion process as described above, it is sufficient that the S_n become dense in S and that $T(t) = \lim_{n\to\infty} T_n^{[th_n^{-1}]}$.

Our subsequent discussion will be directed to giving conditions under which this criterion holds.

The simplest example is the convergence of the symmetric random walk to Brownian motion [6, Chap. 14]. Let S be the real line. The standard Brownian motion process is obtained by taking

$$\mu_{x,t}(A) = \int_A (2\pi t)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}t^{-1}(x-y)^2\right\} dy .$$

The nth symmetric random walk has the integral multiples of n^{-1} for support, and basic time-interval $h_n = n^{-2}$. For $x \in S_n$,

$$\mu_x = \frac{1}{2}\delta(x + n^{-1}) + \frac{1}{2}\delta(x - n^{-1}) ,$$

where $\delta(x)$ denotes the measure giving unit weight to the point x.

The convergence in this case follows immediately from a special case of the central limit theorem (namely, the normal approximation to the binomial distribution). From more general versions of the central limit theorem it may be shown that many different random walks converge to Brownian motion. The precise form of the one-step transition probability distributions is largely irrelevant; essentially all that matters is the behaviour of the mean displacements and mean square displacements. *Our goal is to establish similar results for the more general diffusion processes described above.*

Before giving our next set of definitions, we must fix some notational conventions. Assuming that some definite diffusion process is under discussion, we normalize the function m occurring in (6.1) so that it is continuous on the right in the interior of S and continuous (possibly with values $\pm\infty$) at the end-points. The expression $\int_a^b f\,dm$ denotes the integral over the half-open interval $(a, b]$ if $a < b$, and the negative of the integral over $(b, a]$ if $a > b$. The derivatives of functions in $\mathbf{D}(\Omega)$ may have simple discontinuities at the discontinuities of m [cf. **10**], and we use $f'(x)$ to denote always the right-hand derivative of f at x.

The necessary and sufficient conditions for the central limit theorem [5] involve *truncated* means and variances of the distributions concerned, rather than the actual means and variances (which need not exist). Let C be a covering of S by intervals, and for every $x \in S$ define C_x to be the union of those elements of C which contain x. Then for a random walk with transition probabilities μ_x we introduce the following functions, defined for all x in the support of the random walk.
The *residual probability* at x is

$$P^C(x) = \mu_x(S - C_x) . \tag{6.2}$$

The *truncated mass defect* at x is

$$q^C(x) = 1 - \mu_x(C_x) . \tag{6.3}$$

The *truncated mean displacement* at x is

$$(6.4) \qquad k^C(x) = \int_{C_x} (y - x)\, d\mu_x(y) \,.$$

The (*generalized*) *truncated mean square displacement* at x is

$$(6.5) \qquad s^C(x) = \int_{C_x} v_x(y)\, d\mu_x(y) \,.$$

where

$$v_x(y) = \int_x^y \{m(u) - m(x)\}\, du$$

is the solution of $\Omega v = 1$ which satisfies the conditions $v(x) = v_x'(x) = 0$.[6] For functions in $\mathbf{D}(\Omega)$ we obtain the "Taylor expansion"

$$\begin{aligned}(6.6) \qquad f(y) &= f(x) + \int_x^y f'(u)\, du \\ &= f(x) + (y - x)f'(x) + \int_x^y \int_x^u \Omega f(w)\, dm(w)\, du \\ &= f(x) + (y - x)f'(x) + v_x(y)(\Omega f(x) + E(x, y))\end{aligned}$$

where $E(x, y)$ is an error term whose absolute value does not exceed the oscillation of Ωf on the interval (x, y). Putting this estimate together with the definitions (6.2) – (6.5), we obtain

$$\begin{aligned}(6.7) \qquad Tf(x) &= \int_{C_x} f(y)\, d\mu_x(y) + \int_{S - C_x} f(y)\, d\mu_x(y) \\ &= f(x) + s^C(x)\Omega f(x) - q^C(x) f(x) \\ &\quad + k^C(x) f'(x) + s^C(x) E(x) \\ &\quad + O(p^C(x)\, \| f \|)\end{aligned}$$

where $|E(x)|$ does not exceed the oscillation of Ωf on the interval C_x.

A family of coverings will be said to *contain arbitrarily fine members* if it contains a refinement of every finite covering of S by open intervals.

Suppose a diffusion process with transition semi-group $T(t)$ and a sequence of random walks $\{S_n, h_n, T_n\}$ are given and that the space X and the approximating sequence X_n, P_n are defined as described above. Define $\Omega_n = h_n^{-1}(T_n - I)$.

LEMMA 6.1. *Suppose that there exists a family of coverings containing arbitrarily fine members such that for every covering C in the family*

[6] In the case of Brownian motion, $m(x) = 2x$, $v_x(y) = (x - y)^2$ and $s^C(x)$ is the usual truncated mean square displacement. The remark that the definitions given above furnish the appropriate generalization is due to W. Feller.

$p_n^C(x) = o(h_n)$ and $s_n^C(x) = O(h_n)$, uniformly for $x \in S_n$ as $n \to \infty$.[7] *Then for any $f \in \mathbf{D}(\Omega)$,*

$$\lim_{n\to\infty} \sup_{x\in S_n} | \Omega_n P_n f(x) - \Omega f(x) |$$
$$\leqq \lim_{n\to\infty} \sup_{x\in S_n} h_n^{-1} | - q_n^C(x) f(x) + k_n^C(x) f'(x) + (s_n^C(x) - h_n)\Omega f(x) | .$$

Proof. The functions in X are continuous on a compact set; hence Ωf in particular is uniformly continuous. For any $\varepsilon > 0$, there is a covering C in the postulated family which is sufficiently fine that the oscillation of Ωf on any C_x is less then ε. Then for any $x \in S_n$, (6.7) gives

$$(6.8) \quad \Omega_n P_n f(x) - \Omega f(x) = h_n^{-1}\{- q_n^C(x) f(x) + k_n^C(x) f'(x) + (s_n^C(x) - h_n)\, \Omega f(x) + O(\varepsilon s_n^C(x) + p_n^C(x)\, \| f \|)\} .$$

From the assumptions, the last term on the right is $O(\varepsilon)$, uniformly in n, and since ε is arbitrary, the conclusion follows.

In order to apply Theorem 5.3 to show that $\lim_{n\to\infty} T_n^{[th_n^{-1}]} = T(t)$ we must show that $\lim_{n\to\infty} \Omega_n P_n f = \Omega f$ for sufficiently many functions f. According to the definition, this means that we must show that

$$\lim_{n\to\infty} \sup_{x\in S_n} | \Omega_n P_n f(x) - \Omega f(x) | = 0 .$$

We shall in all cases assume the following

Condition A. There exists a family of coverings such that the hypotheses of Lemma 6.1 are satisfied, and such that

(i) $q_n^C(x) = o(h_n)$

(ii) $k_n^C(x) = o(h_n)$

(iii) $s_n^C(x) = h_n + o(h_n)$

as $n \to \infty$, uniformly for x bounded away from the end-points of S.

We remark that since f and Ωf are bounded for any $f \in \mathbf{D}(\Omega)$ while f' is bounded except perhaps near the end-points of S, these conditions imply that

$$(6.9) \qquad -q_n^C(x) f(x) + k_n^C(x) f'(x) + (s_n^C(x) - h_n)\Omega f = o(h_n)$$

for any $f \in \mathbf{D}(\Omega)$, uniformly for x bounded away from the end-points of S.

[7] The first assumption is essentially the Lindeberg condition **[9]**, strengthened by the requirement of uniformity, and implies that $\lim_{n\to\infty} \Omega_n$ is of local character.

The conditions to be imposed on q, k, and s near the boundaries are more complicated and will depend on the boundary conditions used to define Ω as a contraction of Ω'. The classification of types of boundaries and possible boundary conditions has been given in [8], and justification for the assertions made in what follows regarding the behaviour of the solutions of the homogeneous equation $(\lambda - \Omega')u = 0$ is to be found there. In [8] a different canonical form for the operator Ω' is used, so that the statements require some translation to fit our situation; [19] contains a summary of what we need in terms of the present notation. For simplicity, we shall consider only boundary conditions under which Ω is of local character at the end-points of S as well as in the interior. This restriction rules out any interaction between the boundaries of the sort described in [8] and they can be considered separately. We shall discuss in detail only the left-hand boundary. The modifications necessary to deal with the right-hand boundary will be obvious.

Since $T(t)$ is a proper semi-group, the resolvent $(\lambda - \Omega)^{-1}$ is defined on all of X. (From here on we take λ to have some fixed positive value.) Let Y be the linear subset of X consisting of all functions which are constant in some neighbourhood of each end-point. Suppose $g \in \mathrm{Y}$ has the value c in a neighbourhood of an end-point, and let $f = (\lambda - \Omega)^{-1}g$. Then $(\lambda - \Omega)(f - c\lambda^{-1}) = g - c = 0$ on this neighbourhood. Hence in this neighbourhood, f has the form

$$f(x) = bu(x) + \text{constant} \tag{6.10}$$

where b is some constant and u is the solution (unique up to a constant multiple) of $(\lambda - \Omega')u = 0$ which satisfies the boundary condition for Ω at the end-point in question.

LEMMA 6.2. *Let $\tilde{\Omega}$ be the restriction of Ω to those functions in* $\mathbf{D}(\Omega)$ *which are of the form* (6.10) *in some neighbourhood of each boundary. Then $\tilde{\Omega}$ is densely defined, and* $\mathbf{R}(\lambda - \tilde{\Omega})$ *is dense in X.*

Proof. It is clear that $\tilde{\Omega}$ is densely defined. From the remarks preceding the lemma, $\mathbf{R}(\lambda - \tilde{\Omega}) \supset \mathrm{Y}$ which is dense in X.

Let r denote the left-hand end-point of S. We first consider the case where r is a *natural* boundary, so that no additional boundary condition may be imposed. The functions $u(x)$, $u'(x)$, and $\Omega u(x)$ all tend to zero as x approaches r. Hence for $f \in \mathbf{D}(\tilde{\Omega})$, f' and Ωf tend to zero as $x \to r$. Then if

Condition B.1. For every covering of the family postulated in condition A

$$q_n^c(x) = o(h_n)$$

$$k_n^c(x) = O(h_n)$$

$$s_n^c(x) = O(h_n)$$

uniformly in some neighbourhood of r,

is satisfied, it is clear that for every $f \in \mathbf{D}(\tilde{\Omega})$ and every $\varepsilon > 0$, there exists a neighbourhood N of r such that

$$\text{(6.11)} \quad \lim_{n\to\infty} \sup_{x\in N \cap S_n} h_n^{-1} | - q_n^c(x)f(x) + k_n^c(x)f'(x) + (s_n^c(x) - h_n)\Omega f(x) | \leqq \varepsilon .$$

In the case of an *exit* boundary, we shall suppose that the absorbing barrier condition $f(r) = 0$ is imposed. Then the functions $u(x)$ and $\Omega u(x)$ vanish at r, while $u'(x)$ is continuous at r and has a finite non-zero value there. Since $f(r) = 0$, any $f \in \mathbf{D}(\tilde{\Omega})$ is a multiple of u in some neighbourhood of r. At an exit boundary, r is finite and we may integrate to obtain

$$f(x) = f(r) + (x - r)f'(r) + o(x - r) .$$

Making use of this and the continuity of f and Ωf, we see that if

Condition B.2. For every covering of the family postulated in Condition A

$$\begin{aligned} &\text{(i)} & (x - r)q_n^c(x) &= O(h_n) \\ &\text{(ii)} & k_n^c(x) &= O(h_n) \\ &\text{(iii)} & s_n^c(x) &= O(h_n) \\ &\text{(iv)} & -(x - r)q_n^c(x) + k_n^c(x) &= o(h_n) \end{aligned}$$

uniformly in some neighbourhood of r,

is satisfied, then for any $f \in \mathbf{D}(\tilde{\Omega})$ and $\varepsilon > 0$ there exists a neighbourhood of r on which (6.11) holds.

At an *entrance* boundary r is infinite while m is finite. The function u has the property that $u'(r) = 0$. Hence for $f \in \mathbf{D}(\tilde{\Omega})$, $f'(r) = 0$, and integrating with respect to m yields

$$f'(x) = \{m(x) - m(r)\}\Omega f(r) + o(m(x) - m(r)) .$$

Using this it is easy to see that

Condition B.3. For every covering of the family postulated in Condition A

$$\begin{aligned} &\text{(i)} & \{m(x) - m(r)\}k_n^c(x) &= C(h_n) \\ &\text{(ii)} & s^c(x) &= O(h_n) \end{aligned}$$

(iii) $\{m(x) - m(r)\}k_n^c(x) + s_n^c(x) - h_n = o(h_n)$

(iv) $q_n^c(x) = o(h_n)$

uniformly in a neighbourhood of r,

implies (6.11) for some neighbourhood N.

At a *regular* boundary, m and r are both finite, and the boundary condition is of the form

$$aΩf(r) + bf'(r) + cf(r) = 0 \tag{6.12}$$

where a, b, and c are constants, not all zero. For $f \in \mathbf{D}(\Omega)$ we have

$$\Omega f(x) = \Omega f(r) + o(1)$$

and by integrating obtain

$$f'(x) = f'(r) + \{m(x) - m(r)\}\Omega f(r) + o(m(x) - m(r))$$

and

$$f(x) = f(r) + (x - r)f'(r) + v_r(x)\Omega f(r) + o(v_r(x)) .$$

Under these circumstances, the following condition is sufficient to give the conclusion (6.11) for any $\varepsilon > 0$ and any $f \in \mathbf{D}(\Omega)$.

Condition B.4. For every covering of the family postulated in condition A

(i) $v_r(x)q_n^c(x) = O(h_n)$

(ii) $\{m(x) - m(r)\}k_n^c(x) = O(h_n)$

(iii) $s_n^c(x) = O(h_n)$

(iv) $c\{k_n^c(x) - (x - r)q_n^c(x)\} + bq_n^c(x) = o(h_n)$

(v) $a\{k_n^c(x) - (x - r)q_n^c(x)\}$
$-b\{s_n^c(x) - h_n + \{m(x) - m(r)\}k_n^c(x) - v_r(x)q_n^c(x)\} = o(h_n)$

(vi) $c\{s_n^c(x) - h_n + \{m(x) - m(r)\}k_n^c(x) - v_r(x)q_n^c(x)\}$
$+ aq_n^c(x) = o(h_n)$,

omitting (vi) if $a = b = 0$.

The last three conditions are of course not independent, and if neither a, b, nor c is zero any two of them imply the third. Condition (vi) can be omitted if $a = b = 0$ because this is the absorbing barrier condition and hence for $f \in \mathbf{D}(\Omega), f(r) = \Omega f(r) = 0$.

Putting all these results together, we obtain as the final result

THEOREM 6.1. *If a diffusion process and a sequence of random walks are such that*

(i) *The supports S_n become dense in S and*

(ii) *Condition* A *and the appropriate conditions from among* B.1, B.2, B.3, B.4, *and their analogues for the right-hand boundary are satisfied*

then the random walks converge to the diffusion process.

Proof. For any $\varepsilon > 0$ and any $f \in \mathbf{D}(\tilde{\Omega})$ the conditions imposed near the boundaries imply the existence of neighbourhoods of the boundaries on which (6.11) holds. Taken with Condition A and Lemma 6.1, this implies that for $f \in \mathbf{D}(\tilde{\Omega})$, $\lim_{n\to\infty} \Omega_n P_n f = \Omega f$. Hence, from Lemma 6.2, $\lim_{n\to\infty} \Omega_n$ is densely defined and $\lambda - \lim_{n\to\infty} \Omega_n$ has a dense range. Since the operators T_n are all bounded by 1, Theorem 5.3 applies to give the desired conclusion.

For our first example we shall take Brownian motion on the half-line $[0, \infty]$, for which there is a regular boundary at the origin and a natural boundary at infinity. Let $\{h_n\}$ be a sequence of positive numbers converging to zero, and define $d_n = h_n^{1/2}$. Consider a sequence of random walks, with the nth walk having a basic time interval h_n and support S_n consisting of the non-negative integral multiples of d_n. Away from the origin take the ordinary symmetric random walk with

$$\mu_x^n = \frac{1}{2}\delta(x + d_n) + \frac{1}{2}\delta(x - d_n)$$

for $x \in S_n$, $x \neq 0$. The behaviour at the origin will, of course depend on the boundary condition to be imposed on the diffusion process. As the family of coverings required in condition A we shall take the family of all finite coverings of $[0, \infty]$ by open intervals. Since $d_n \to 0$ it is easy to see that for any fixed covering, for sufficiently large n, $p_n^C(x) = q_n^C(x) = k_n^C(x) = 0$ and $s_n^C(x) = h_n$ for all $x \in S_n$, $x \neq 0$. Thus Condition A is satisfied, and so is Condition B.1 at the natural boundary. Furthermore, whatever condition of type B.4 is imposed at the regular boundary, it will automatically be satisfied in any neignbourhood of that boundary except at the origin itself. At the origin the Conditions B.4 reduce to

(i) $s_n^C(0) = O(h_n)$

(ii) $ck_n^C(0) + bq_n^C(0) = o(h_n)$

(iii) $ak_n^C(0) - b\{s_n^C(0) - h_n\} = o(h_n)$

(iv) $c\{s_n^C(0) - h_n\} + aq_n^C(0) = o(h_n)$

where (iv) can be omitted in the absorbing barrier case.

In order to obtain convergence to the *absorbing barrier* process, it is sufficient that no particle which reaches the origin returns to the interior of the interval (i.e., once a particle is at the origin, it either stays there or disappears). In this case, $k_n^C(0) = s_n^C(0) = 0$, and since in the absorbing barrier process $a = b = 0$ and condition (iv) can be omitted, the conditions are satisfied.

It is slightly more complicated to obtain a sequence of random walks converging to a diffusion process with an *elastic barrier* condition $a = 0$, $b = 1$, $c = -\alpha$ (which by taking $\alpha = 0$ specializes to the *reflecting barrier* condition). It can be done by letting a particle reaching the origin be "reflected" to the point d_n with probability $(1 + \alpha d_n)^{-1}$ and disappear with probability $\alpha d_n(1+\alpha d_n)^{-1}$. This gives $q_n^G(0) = \alpha d_n(1+\alpha d_n)^{-1}$, $k_n^G(0) = d_n(2 + \alpha d_n)^{-1}$, and $s_n^G(0) = d_n^2(1 + \alpha d_n)^{-1}$. Remembering that $d_n^2 = h_n$, it easy to verify that the conditions given above are satisfied.

As a further variation of Brownian motion, let us consider the diffusion process on $[-\infty, \infty]$ defined by taking

$$\begin{aligned} m[x] &= 2x - 1 \qquad x < 0 \\ &= 2x + 1 \qquad x \geqq 0 . \end{aligned}$$

This is known to give a Brownian motion process modified by the introduction of a "delay" at the origin so that the set of times for which a particle is at the origin is a nowhere-dense set of positive Lebesgue measure. We shall show that this process is the limit of symmetric random walks, modified so that a particle at the origin has probability $1 - d_n$ of staying there at the next jump, and probability $d_n/2$ of jumping to each of its neighbours. (We use h_n and d_n with the same meaning as in the previous example). As before, we have $q_n^G(x) = k_n^G(x) = 0$ for all $x \in S_n$, while $s_n^G[x] = h_n$ for all $x \neq 0$. Noting that $v_0(y) = y^2$ for $y \geqq 0$, and $v_0(y) = y^2 + 2\,|\,y\,|$ for $y < 0$, we obtain

$$\begin{aligned} s_n^G(0) &= d_n(d_n^2 + d_n) \\ &= h_n + d_n^3 = h_n + o(h_n) . \end{aligned}$$

The boundaries are both natural and Conditions A and B.1 can obviously be satisfied by taking the required family of coverings to consist of all finite coverings of the line.

As a final example, in which it is not so evident *a priori* what the appropriate boundary conditions are, we consider the limiting behaviour of a sequence of random walks encountered in genetic theory. These processes are discussed in [7, p. 232 ff.] and their genetic interpretation is described there. We have made some inessential changes in notation for purposes of convenience.

The processes take place on the interval $0 \leqq y \leqq 1$. (We use y as coordinate to reserve x for the natural scale used to express Ω' in the form (6.1).) The nth random walk has basic time-interval $h_n = n^{-1}$ and support the integer multiples of n^{-1}. The one-step transition probabilities are

$$\mu_y^n = \sum_{k=0}^{n} \binom{n}{k} p_{n,y}^k q_{n,y}^{n-k} (y_k) \tag{6.13}$$

so that

$$T_n f(y) = \sum_{k=0}^{n} \binom{n}{k} p_{n,y}^k q_{n,y}^{n-k} (y_k)$$

where $y_k = kn^{-1}$, $p_{n,y} = 1 - q_{n,y} = y(1 - rn^{-1}) + (1 - y)sn^{-1}$, and r and s are non-negative constants.

It is easy to check that the formal limit of the operators $\Omega_n = n(T_n - I)$ is the differential operator

$$\Omega' = \frac{1}{2} y(1 - y)D^2 + \{s - (r + s)y\}D \tag{6.14}$$

where D denotes differentiation with respect to y.[8] The natural scale x, and the monotone function m needed to express Ω' in the form (6.1) may be taken as [10, formula 4.2]

$$x(y) = \int_{1/2}^{y} z^{-2s}(1 - z)^{-2r}\, dz \tag{6.15}$$

$$m(y) = 2\int_{1/2}^{y} z^{2s-1}(1 - z)^{2r-1}\, dz\ .$$

The nature of the boundary at $y = 0$ depends on the value of s; the nature of the boundary at $y = 1$ depends similarly on the value of r. We shall discuss only the left-hand boundary in detail, since there is obviously a complete symmetry. Checking with the criteria given in [8] or [19], we see that the boundary at $y = 0$ is an exit boundary if $s = 0$, a regular boundary if $0 < 2s < 1$, and an entrance boundary if $2s \geq 1$.

It would presumably be possible to show that the sequence of random walks under consideration satisfies our Condition A and the appropriate conditions at the boundary, but the functions x and m are not elementary, and it would be complicated to obtain satisfactory estimates for the mean displacements, etc. We shall instead make direct use of Lemma 6.2 and Theorem 5.3.

We assert that *if f has a continuous second derivative on the closed interval* $[0, 1]$ *then* $\lim_{n\to\infty} \Omega_n f = \Omega' f$. To prove this, observe first that for the constant function 1, for $g(y) = y$, and for $h(y) = y^2$, simple calculations from the elementary formulas for the mean and variance of a binomial distribution give

[8] Goldberg [12] discusses the solutions of the equation $\partial f/\partial t = \Omega' f$ under various boundary conditions.

$$\Omega_n 1 = 0 = \Omega 1$$

$$\Omega_n g(y) = s - (r + s)y = \Omega g(y)$$

$$\begin{aligned}\Omega_n h(y) &= (1 + 2s)y - (1 + 2r + 2s)y^2 + O(n^{-1}) \\ &= \Omega h(y) + O(n^{-1}) .\end{aligned}$$

It is also easy to verify (for instance, by estimates obtained from the normal approximation to the binomial distribution) that for any $\varepsilon > 0$,

$$p_n^\varepsilon(y) = \int_{|z-y|>\varepsilon} d\mu_y^n(z) = o(n^{-1})$$

uniformly in y. Then using a second-order Taylor expansion, any f with a continuous second derivative can be approximated over the interval $(y - \varepsilon, y + \varepsilon)$ by a linear combination of 1, g, and h so that the error at z is less than $(y - z)^2 E_\varepsilon$ where E_ε is the maximum oscillation of $D^2 f$ on any interval of length 2ε. Estimating f by such a linear combination, we obtain

$$\Omega_n f(y) = \Omega f(y) + O(n^{-1} + E_\varepsilon) + np_n^\varepsilon(y) .$$

Since E_ε can be made arbitrarily small by suitable choice of ε and for each such choice $np_n^\varepsilon(y) = o(1)$, our assertion holds as stated.

We shall complete the proof that the random walks (6.13) converge to a diffusion process associated with the operator (6.14) by showing that (if a suitable boundary condition is imposed in the regular boundary case) all functions in $\mathbf{D}(\tilde{\Omega})$, where $\tilde{\Omega}$ is defined as in Lemma 6.2, have continuous second derivatives on the closed interval. Then we shall have $\lim_{n\to\infty} \Omega_n$ densely defined and $\mathbf{R}(\lambda - \lim_{n\to\infty} \Omega_n)$ dense, so that Theorem 5.3 will apply to give the desired conclusion. Of course, every $f \in \mathbf{D}(\Omega')$ has a continuous second derivative on the open interval, but since the coefficient of D^2 in Ω' vanishes at the end-points, the continuity of $\Omega' f$ on $[0, 1]$ does not imply the continuity of $D^2 f$ there.

The homogeneous equation $(\lambda - \Omega')u = 0$ may be put into the form of the standard hypergeometric equation

$y(1 - y)u''(y) + \{c - (a + b + 1)y\}'(y) - abu = 0$ by determing a, b, and c from the equations

$$\begin{aligned} c &= 2s \\ a + b + 1 &= 2r + 2s \\ ab &= 2\lambda .\end{aligned}$$

The solutions of this equation are given in terms of hypergeometric functions in section 2.3.1 of [2], whose notation we adopt. In the case of a non-regular boundary there is a unique solution which is bounded at 0. This solution is

$$u = yF(a - c + 1, b - c + 1; 2; y)$$

for the exit boundary case, $c = 2s = 0$, and

$$u = F(a, b; c; y)$$

for the entrance boundary case, $c = 2s \geq 1$. Since the hypergeometric function is analytic at the origin, it has a continuous second derivative there.

In the case of a regular boundary, $0 < c = 2s < 1$, there are two independent solutions which are bounded at 0, namely

$$u_1 = F(a, b; c; y)$$

and

$$u_2 = y^{1-c}F(a - c + 1, \; b - c + 1; \; 2 - c; y) \, .$$

We now impose the reflecting barrier condition, $df/dx \to 0$ as $x \to 0$. From (6.15)

$$\frac{df}{dx} = y^{2s}(1 - y)^{2r}D$$

and it is easily verified that $du_1/dx \to 0$ at the boundary, but that du_2/dx does not. Thus solutions of the homogeneous equation which satisfy the boundary condition imposed are multiples of u_1, and have a continuous second derivative at 0.

References

1. S. Banach, *Théorie des opérations linéaires*, Warsaw, 1932.

2. Bateman Manuscript Project (ed. A. Erdélyi), *Higher Transcendental Functions*, vol. 1, New York, 1953.

3. J.L. Doob, *Stohastic processes*, John Wiley & Sons, New York, 1953.

4. J. Douglas, Jr., *On the relation between stability and convergence in the numerical solution of linear parabolic and hyperbolic differential equations*, J. Soc. Indust. Appl. Math. **4** (1956), 20–37.

5. W. Feller, *Uber den zentralen Grenzwertsatz der Wahrscheinlickeitsrechnung*, Math. Z. **40** (1935), 521–559.

6. ———, *An introduction to probability theory and its applications*, vol. 1, John Wiley & Sons, New York, 1950.

7. ———, *Diffusion processes in genetics*, Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1951, 227–246.

8. ———, *The parabolic differential equations and associated semi-groups of transformations*, Ann. of Math., **55** (1952), 468–519.

9. ———, *Diffusion processes in one dimension*, Trans. Amer. Math. Soc., **77** (1954) 1–31.

10. ———, *On differential operators and boundary conditions*, Comm. Pure and Appl. Math., **8** (1955), 203–216.

11. ———, *The general diffusion operator and positivity preserving semi-groups in one dimension*, Ann. of Math. **60** (1954), 417–436.

12. S. Goldberg, *On a singular diffusion equation*, Thesis, Cornell, 1950.

13. D.R. Hartree, *Numerical analysis*, Oxford University Press, 1952.

14. E. Hille, *Functional analysis and semi-groups*, Amer. Math. Soc. Colloquium Publications, vol. 31, New York, 1948.

15. ———, *The abstract Cauchy problem and Cauchy's problem for parabolic differential equations*, Journal d'Analyse Mathématique, **3** (1953–1954), 81–196.

16. L.V. Kantorovitch, *Functional and applied mathematics*, Uspehi Matem. Nauk **3** (1948); Bureau of Standards report 1509, 1952.

17. P.D. Lax and R.D. Richtmyer, *Survey of the stability of linear finite difference equations*, Comm. Pure Appl. Math. **9** (1956), 267–293.

18. W. Leutert, *On the convergence of approximate solutions of the heat equation to the exact solution*, Proc. Amer. Math. Soc. **2** (1951), 433–439.

19. H.P. McKean, Jr., *Elementary solutions for certain parabolic partial differential equations*, Trans. Amer. Math. Soc. **82** (1956), 519–548.

20. G.G. O'Brien, M.A. Hyman, and S. Kaplan, *A study of the numerical solution of partial differential equations*, J. Math. Phys. **29** (1951), 223–251.

21. D.V. Widder, *The Laplace transform*, Princeton University Press, 1941.

22. K. Yosida, *On the differentiability and the representation of one-parameter semi-group of operators*, J. Math. Soc. Japan. **1** (1949).

QUEEN'S UNIVERSITY KINGSTON, CANADA

A FIXED POINT THEOREM FOR MULTI-VALUED FUNCTIONS

L. E. WARD, JR.

1. Introduction. If Y is a topological space then $S(Y)$ denotes the space of closed subsets of Y, endowed with the finite topology [2]. We say that a multi-valued function $F: X \to Y$ is continuous provided $F(x)$ is a closed set for each $x \in X$, and the induced (single-valued) function $f: X \to S(Y)$ is continuous in the usual sense. This definition of continuity for multi-valued functions is equivalent to that of Strother [4]. The space X is said to have the F.p.p. (= fixed point property for continuous multi-valued functions) if and only if for each such function $F: X \to X$, there is an $x \in X$ such that $x \in F(x)$. The space X has the f. p. p. if it has the fixed point property for continuous single-valued functions. It is hardly surprising that the T_1-spaces with the F. p. p. constitute a fairly small subclass of those with f. p. p. Indeed, Plunkett [3] has shown that a Peano continuum has the F. p. p. if and only if it is a dendrite. It is worth noting that Plunkett's argument employs the convex metric of a dendrite in much the same manner as the author [7] has used the order structure of certain acyclic continua to obtain fixed point theorems. A related argument has been used by Capel and Strother [1] to show that a tree has the fixed point property for continuous multi-valued functions for which the image of each point is connected. Their proof depends on being able to produce a continuous selection, in the sense of Michael, on the class of subcontinua of a tree.

In this paper an order-theoretic characterization of a wide class of acyclic spaces is given. This characterization is in the same spirit as the analogous results of [6] and [7] for trees and generalized trees. It is then shown, using their order properties, that such spaces have the F. p. p. To some extent the argument borrows from all of the proofs cited above.

2. Topologically chained continua. If X is a space and $\leqq$ is a partial order on X, we write $L(x) = \{a : a \leqq x\}$ and $M(x) = \{a : x \leqq a\}$. It is natural and convenient to define

$$[x, y] = M(x) \cap L(y)$$

and, if $A \subset X$, we write $M(A)$ for the union of all $M(x)$ for which

Received August 7, 1957, and in revised forms November 4, 1957, and February 6, 1958. Presented to the American Mathematical Society, August 29, 1957,

$x \in A$. An *antichain* of X is a subset in which no two distinct elements are related under the partial order. A *zero* of a subset A of X is a member a_0 of A such that $A \subset M(a_0)$.

In what follows it will be convenient to use two theorems from [5]. These are stated below as Theorems A and B. The partial-order $\leqq$ is said to be *upper* (*lower*) *semicontinuous* provided $M(x)$ ($L(x)$) is closed, for each $x \in X$. It is *semicontinuous* if it is both upper and lower semicontinuous.

THEOREM A. *If the compact space X is endowed with an upper (lower) semicontinuous partial order then X admits a maximal (minimal) element.*

THEOREM B. *If the compact space X is endowed with an order-dense semicontinuous partial order and if the set of maximal elements or the set of minimal elements is connected, then X is connected.*

Recall that a continuum (= compact connected Hausdorff space) is *unicoherent* if it cannot be represented as the union of two subcontinua whose intersection is not connected. A continuum is *hereditarily unicoherent* if each of its subcontinua is unicoherent.

A continuum is *topologically chained* if each pair of points is contained in a subcontinuum which has exactly two non-cutpoints, and such a subcontinuum is called a topological chain. This concept is a natural generalization of the notion of an arc. Note that, if x and y are distinct elements of a continuum which is topologically chained and hereditarily unicoherent, there is a unique subcontinuum $C(x, y)$ which is irreducible about x and y. Moreover, $C(x, y)$ is a topological chain with x and y for endpoints.

Consider the following five properties enjoyed by some spaces X admitting a partial order, $\leqq$.

I. *$[x, y]$ is closed and simply ordered for each x and y in X.*

II. *$\leqq$ is order-dense.*

III. *There exists $e \in X$ such that $M(e) = X$.*

IV. *If x and y are points of the subcontinuum Y and $x \leqq y$, then $[x, y] \subset Y$.*

V. *If A is an antichain of X and P is a continuum contained in $M(A)$, then $P \subset M(x)$ for some $x \in A$.*

LEMMA 1. *If X is a compact Hausdorff space with a partial order satisfying* I *and* II, *and if $x < y$ in X, then $[x, y]$ is a topological chain,* *Proof.* For each $t \in [x, y]$ the sets

$$L(t) \cap [x, y] = [x, t],$$
$$M(t) \cap [x, y] = [t, y]$$

are closed so that $\leqq$ is semicontinuous on $[x, y]$. Since $[x, y]$ has only one minimal element it is connected by Theorem B. If $x < t < y$ then the continua $[x, t]$ and $[t, y]$ have only the point t in their intersection, so that x and y are the only non-cutpoints of $[x, y]$; that is, $[x, y]$ is a topological chain.

LEMMA 2. *If X is a compact Hausdorff space with a partial order satisfying* I, III, *and* V, *then each continuum contained in X has a zero.*

Proof. By I and III the set $L(x) = [e, x]$ is closed for each $x \in X$ and consequently, by Theorem A, each compact subset of X contains a minimal element. If Y is a subcontinuum of X then the set A of minimal elements of Y forms a nonempty antichain. It follows from V that A contains only one element, that is, Y has a zero.

THEOREM 1. *A necessary and sufficient condition that X be a topologically chained, hereditarily unicoherent continuum is that X be a compact Hausdorff space which admits a partial order satisfying* I-V.

Proof. Necessity. Let X be a topologically chained, hereditarily unicoherent continuum. Fix $e \in X$ and let $x \leqq y$ mean that $x \in C(e, y)$. Then $L(x) = C(e, x)$ and I, II and III are easily verified. If $x \leq y$ and x and y are elements of the subcontinuum Y then, by hereditary unicoherence, $[x, y] \cap Y$ is connected. Since x and y are elements of Y it follows that $[x, y] \subset Y$. This establishes IV. To see that V is satisfied let A be an antichain of X and suppose that P is a continuum contained in $M(A)$. If P meets $M(x)$ and $M(y)$ where x and y are distinct points of A, we may select $p \in M(x) \cap P$ and $q \in M(y) \cap P$. In view of I the points p and q are not comparable. Therefore, the continua $L(p) \cup P$ and $L(q) \cup P$ meet in the non-connected set $P \cup (L(x) \cap L(y))$, and this contradicts the hereditary unicoherence of X.

Sufficiency. Let X be a compact Hausdorff space which admits a partial order satisfying I-V. By Lemma 1 each set $L(x) = [e, x]$ is a topological chain and therefore X is a topologically chained continuum. Now suppose A and B are subcontinua of X and that x and y are distinct elements of $A \cap B$. In order to show that $A \cap B$ is connected and thus that X is hereditarily unicoherent it is sufficient to prove that

$$Z = [z, x] \cup [z, y] \subset A \cap B,$$

where z is the supremum of $L(x) \cap L(y)$. The existence of z is assured by Theorem A. By Lemma 2, A has a zero, a_0, so that a_0 is a predecessor of x and y and hence also of z. By IV,

$$[z, x] \subset [a_0, x] \subset A .$$

It is clear that, by a sequence of analogous arguments, $Z \subset A \cap B$.

3. The fixed point theorem. The proof of Theorem 2 which follows is patterned after that of Theorem 10 in [7] where it was assumed that the partial order had a closed graph. In view of our weaker hypotheses it has been necessary to revise that argument extensively. For the remainder of this paper the term *sequence* is used in its generalized sense, that is, a function defined on the predecessors of some ordinal number. If x is a sequence then a *subsequence* of x is the restriction of x to some cofinal subset of its domain.

A useful continuity property of a partial order satisfying I-V is given by the following lemma.

LEMMA 3. *Let X be a compact Hausdorff space which contains no indecomposable continuum and which admits a partial order satisfying* I-V. *If x is an increasing sequence in X then* lim x *exists and* $x_\alpha \leqq$ lim x *for each α in the domain of x.*

Proof. Let x_0 be a limit point of x and suppose $x_0 \in X - M(x_\beta)$ for some β. Without loss of generality we may take $\beta = 1$. Let C be the union of the topological chains $[x_\alpha, x_{\alpha'}]$, $\alpha' > \alpha$. If σ is the zero of $\overline{C}$ then $x_0 \in \overline{C}$ implies $\sigma < x_1$. If K is a subcontinuum of $\overline{C}$ which contains the values of some subsequence of x, then by IV, $x_\alpha \in K$ implies $[x_\alpha, x_\alpha'] \subset K$ for each $\alpha' > \alpha$ and hence K contains x_0 and σ. Again by IV, $K = \overline{C}$. It follows that $\overline{C}$ is the union of no two proper subcontinua and this contradicts our hypothesis that X contains no indecomposable continuum. Hence $x_\alpha \leqq x_0$ for each α in the domain of x. If $y \in X$ is not a predecessor of x_0 then y is a member of the open set $U = X - L(x_0)$. Since $x_\alpha \in X - U$ for each α, y is not a limit point of x. Therefore x_0 is the unique limit point of x, that is, $x_0 = \lim x$.

LEMMA 4. *Let K be a connected topological space and let F be a continuous multi-valued function defined on K. Suppose that $F(x)$ is a compact set for each $x \in K$. If Q is a quasi-component of $F(K)$ then $F(x) \cap Q$ is nonempty for each $x \in K$.*

Proof. By definition $Q = \cap \{V_\alpha\}$ where $\{V_\alpha\}$ is the family of open and closed sets containing Q. If $F(x)$ meets each V_α then, by the compactness of $F(x)$, it meets Q and therefore $F^{-1}(Q) = \cap \{F^{-1}(V_\alpha)\}$. By an elementary continuity argument it can be verified that each set $F^{-1}(V_\alpha)$ is both open and closed and hence is equal to K. Therefore, $K = F^{-1}(Q)$ and the lemma follows.

For the remainder of this section, X denotes a topologically chained, hereditarily unicoherent continuum which contains no indecomposable subcontinuum, and $F: X \to X$ is a continuous multi-valued function. By Theorem 1 the space X admits a partial order satisfying I-V. Let J be the set of all elements x of X such that (i) there exists a minimal element t_x in the set $F(x) \cap M(x)$, and (ii) if $x < p \leqq t_x$ then $F(p)$ and $[p, t_x]$ are disjoint.

LEMMA 5. *If F is fixed point free then J is not empty. Indeed if $a \in X$ such that $F(a) \cap M(a)$ is not empty then $M(a) \cap J$ is not empty.*

Proof. By Lemma 2, X has a zero and hence there exists $a \in X$ such that $F(a) \cap M(a)$ is not empty. If $b \in F(a) \cap M(a)$ then by Theorem A there is a minimal element t_a in the compact set $F(a) \cap [a, b]$. Clearly t_a is minimal in $F(a) \cap M(a)$. Let K denote the set of all $p \in [a, t_a]$ such that $F(p) \cap [p, t_a]$ is not empty. Then $a \in K$ and $t_a \in X - K$. If $x_0 = \sup K$ it is easy to construct an increasing sequence y in K such that $\lim y = x_0$. For each α in the domain of y let $z_\alpha \in F(y_\alpha) \cap [y_\alpha, t_a]$. Then the sequence z has a limit point z_0 and by continuity $z_0 \in F(x_0)$. Further, since $y_\alpha < z_\alpha \leqq t_a$ for each α, it follows that $z_0 \in [x_0, t_a]$ and hence $x_0 \in K$. Now let t_0 be minimal in $F(x_0) \cap [x_0, z_0]$. If $x_0 < p \leqq t_0$ and $F(p) \cap [p, t_0]$ is not empty then $t_0 \leqq t_a$ implies that $F(p) \cap [p, t_a]$ is not empty. But this is a contradiction since x_0 is the supremum of K. Therefore $F(p)$ and $[p, t_0]$ are disjoint, that is, $x_0 \in J$ and the lemma is proved.

Consider the set S of all sequences x such that J contains the range of x and

$$x_1 < x_2 < \cdots < x_\alpha < \cdots,$$

$$t_\alpha \leqq x_{\alpha+1} \text{ for each } \alpha + 1 \text{ in the domain of } x,$$

where $t_\alpha = t_{x_\alpha}$. If J is not empty then S contains at least one nonempty sequence. We partially order the elements of S in the following manner: x precedes y provided x is an initial segment of y. Clearly, if N is a simply ordered subset of S then $\cup N$ is a member of S. By Zorn's lemma S contains a maximal element.

LEMMA 6. *If F is fixed point free and x is an element of S then the domain of x is not a limit ordinal.*

Proof. If the domain of x is a limit ordinal, let $x_0 = \lim x$, $t_0 = \lim t$. By Lemma 3, $x_\alpha \leqq x_0$ and $t_\alpha \leqq t_0$ for each α. From the definition of J we have $x_\alpha < t_\alpha$ for each α and hence $x_0 \leqq t_0$. But by the definition of S, $t_{\alpha+1} \leqq x_\alpha$ for each α and hence $t_0 \leqq x_0$. This implies that $x_0 = t_0$ is a fixed point of F, which is a contradiction.

LEMMA 7. *If F is fixed point free and x is an element of S then the domain of x is finite.*

Proof. Suppose the domain of x is not finite. We denote by ω the first infinite ordinal. Then by Lemma 6, ω is a member of the domain of x. From the definition of S it is clear that x, restricted to ω, is an element of S, contrary to Lemma 6.

THEOREM 2. *Each topologically chained, hereditarily unicoherent continuum which contains no indecomposable continuum has the* F. p. p.

Proof. If $F: X \to X$ is fixed point free then, by Lemma 5, J is not empty and hence S contains a maximal element, x. By Lemma 7 the domain of x is a set $\{1, 2, \cdots, N\}$ of integers. We write $t_N = t_{x_N}$. By Lemma 5 and the maximality of x it follows that $M(t_N) \cap F(t_N)$ is empty. Now from Corollary 9.6 of [2] the set $F([x_N, t_N])$ is compact and hence its components and quasi-components are identical. Let D be that component of $F([x_N, t_N])$ such that $t_N \in D$. By Lemma 4, $F(t_N) \cap D$ is not empty and hence $D - M(t_N)$ is not empty. Now D is a continuum and hence has a zero z which precedes t_N. By IV, $[z, t_N] \subset D$ but since F is fixed point free there exists $q \in X$ with $x_N < q < t_N$ such that $[q, t_N] \subset D - F(t_N)$. Let b be an infinite increasing sequence in $[q, t_N]$ such that $\lim b = t_N$. For each b_α there exists a_α, $x_N < a_x < t_N$, such that $b_x \in F(a_\alpha)$. Since $x_N \in J$ we infer that $b_\alpha < a_\alpha$. Clearly $\lim a = t_N$ from which it follows that $t_N \in F(t_N)$, a contradiction. Therefore F cannot be fixed point free.

4. **Remarks.** The following question is of some interest. Can the assumption of hereditary decomposability be omitted from the hypotheses of Theorem 2? This assumption is essential to Lemma 3 so that an affirmative answer to our question would require a substantially different proof. Even more difficult is the problem of characterizing the F. p. p. among the topologically chained continua. In view of Plunkett's characterization [3] of the F. p. p. for Peano continua, there would seem to be some hope of discovering a succinct characterization.

REFERENCES

1. C. E. Capel and W. L. Strother, *A new proof of a fixed point theorem of Wallace*, Bull. Amer. Math. Soc. Abstract 62-1-111.

2. E. Michael, *Topologies on spaces of subsets*, Trans. Amer. Math. Soc., **71** (1951), 152-182.

3. R. L. Plunkett, *A fixed point theorem for continuous multi-valued transformations*, Proc. Amer. Math. Soc., **7** (1956) 160-163.

4. W. L. Strother, *On an open question concerning fixed points*, Proc. Amer. Math. Soc., **4** (1953), 988-993.

5. L. E. Ward, Jr., *Partially ordered topological spaces*, Proc. Amer. Math. Soc., **5** (1954), 144-161.

6. ———, *A note on dendrites and trees*, Proc. Amer. Math. Soc., **5** (1954), 992-994.

7. ———, *Mobs, trees, and fixed points*, Proc. Amer. Math. Soc., **8** (1957), 798-804.

U. S. NAVAL ORDNANCE TEST STATION
CHINA LAKE, CALIFORNIA

ON THE NUMBER OF LATTICE POINTS IN $x^t + y^t = n^{t/2}$

R. E. WILD.

Introduction. Suppose that t is independent of $n, n > 1$; $t = (2M)/(2N+1)$; $M = 1, 2, 3, \cdots$; $N = 0, 1, 2, \cdots$; $M \geqq N+1$, so that $t > 1$. Let $L_t(n^{t/2})$ be the number of lattice points, (x, y), satisfying $x^t + y^t \leqq n^{t/2}$. Our main objective is the proof of the relation

$$(1.1) \quad S(n) = t/2 \; n^{1-t/2} \int_0^n L_t(w^{t/2}) w^{t/2-1} dw$$

$$= c_1 n^2 - c_2/\pi n^{(2t-1)/(2t)} \sum_{\alpha=1}^{\infty} \alpha^{-2-1/t} \cos(2\pi\sqrt{n}\,\alpha - \pi/(2t))$$

$$- \frac{2t}{\pi^2\sqrt{t-1}} n^{3/4} \sum_{\alpha=1}^{\infty} \sum_{\beta=1}^{\infty} \frac{\cos(2\pi H\sqrt{n} - \pi/4)}{(\alpha\beta)^{(t-2)/(2t-2)} H^{(3t-1)/(2t-2)}} + O(\sqrt{n})$$

with $t > 1$, $c_1 = \dfrac{2\Gamma^2(1/t)}{(t+2)\Gamma(2/t)}$, $\quad c_2 = \dfrac{2^{(2t-1)/t} t^{1/t} \Gamma(1/t)}{\pi^{(t+1)/t}}$,

$H = (\alpha^{t/(t-1)} + \beta^{t/(t-1)})^{(t-1)/t}$. The case $t = 2$ is known in connection with the classical problem of the lattice points in a circle [4, pp. 221, 235].

By choosing t as specified above the analysis is less bulky than it would be if we considered the slightly more general problem of $L_T(n^{T/2})$ corresponding to the curve $|x|^T + |y|^T = n^{T/2}$ with real $T > 0$. Expressions and estimates for $L_T(n^{T/2})$ have been obtained by Bachmann [1, pp. 447–450], Cauer [2], and van der Corput [3]. In particular van der Corput [3] found that

$$(1.2) \quad L_T(n^{T/2}) = c_1' n - 8T^{(1-T)/T} n^{(T-1)/(2T)} \int_0^{\infty} g_1(\sqrt{n} - x) x^{(1-T)/T} dx$$

$$+ O(n^{1/3}),\ T > 3\,;$$

$$= c_1' n - 8 \sum_{j=1}^{\infty} (-1)^{j+1} \binom{1/T}{j} \zeta(-jT) n^{(1-jT)/2}$$

$$+ O(n^{1/3}),\ 0 < T \leqq 3,\ T \neq 1\,;$$

where

$$c_1' = \frac{2\Gamma^2(1/T)}{T\Gamma(2/T)},$$

$g_1(x) = x - [x] - 1/2$, $[x]$ is the integral part of x, $\zeta(s)$ is the Riemann zeta function and $\binom{a}{b}$ is the binomial coefficient. From (1.2) it follows that

Received June 3, 1958. This paper is condensed from the author's Ph. D. thesis written under the guidance of Professor E. G. Straus at U. C. L. A.

(1.3) $L_T(n^{T/2}) = c_1'n + O(n^{(T-1)/(2T)}),\ L_T(n^{T/2}) = c_1'n + \Omega(n^{(T-1)(2T)}, \quad T > 3\ .$

These results in (1.3) and analogous results can be obtained from (1.1) also. Our methods fail to establish the analogue of (1.1) for $0 < t < 1$.

2. First auxiliary result. We first obtain the result

$$(2.1)\quad S(n) = n^2 \sum_{\alpha=-\infty}^{\infty} \sum_{\beta=-\infty}^{\infty} \iint_{x^t+y^t \leqq 1} (1 - x^t - y^t)\cos 2\pi\sqrt{n}\,(\alpha x + \beta y)dxdy\ , \qquad t > 1\ .$$

In § 4 we prove that the double series is absolutely convergent.

We have [4, p. 205]

$$(2.2)\quad \int_0^W L_t(w)dw = \int_0^W \sum_{j^t+k^t \leqq w}\sum |\,dw = \sum_{j^t+k^t \leqq W}\sum \int_{j^t+k^t}^W dw$$
$$= \sum_{j^t+k^t \leqq W}\sum (W - j^t - k^t) = \sum_{-W^{1/t} \leqq j \leqq W^{1/t}} \ \sum_{-(W-j^t)^{1/t} \leqq k \leqq (W-j^t)^{1/t}} (W - j^t - k^t)\ .$$

To this we apply the Poisson summation formula [4, p. 204] to obtain

$$(2.3)\quad \int_0^W L_t(w)dw = \sum_{\alpha=-\infty}^{\infty} \int_{-W^{1/t}}^{W^{1/t}} \cos 2\pi\alpha x \sum_{-(W-x^t)^{1/t} \leqq k \leqq (W-x^t)^{1/t}} (W - x^t - k^t)dx$$
$$= \sum_{\alpha=-\infty}^{\infty} \int_{-W^{1/t}}^{W^{1/t}} \cos 2\pi\alpha x \sum_{\beta=-\infty}^{\infty} \int_{-(W-x^t)^{1/t}}^{(W-x^t)^{1/t}} \cos 2\pi\beta y \cdot (W - x^t - y^t)dydx.$$

Integrating by parts and applying the second mean value theorem for integrals, we have, for the inner integral,

$$\frac{t}{\pi\beta}\int_0^{(W-x^t)^{1/t}} \sin 2\pi\beta y \cdot y^{t-1}dy = \frac{t(W-x^t)^{(t-1)/t}}{\pi\beta}\int_\xi^{(W-x^t)^{1/t}} \sin 2\pi\beta y dy\ ,$$

where $0 \leqq \xi < (W - x^t)^{1/t}$, so that the sum over β is uniformly convergent in x. Hence we can interchange the order of operations in $\int dx \sum_\beta$ in (2.3) to obtain

$$(2.4)\quad \int_0^W L_t(w)dw = \sum_{\alpha=-\infty}^{\infty} \sum_{\beta=-\infty}^{\infty} \iint_{x^t+y^t \leqq W} \cos 2\pi\alpha x \cos 2\pi\beta y \cdot (W - x^t - y^t)dxdy\ .$$

By symmetry we can replace $\cos 2\pi\alpha x \cos 2\pi\beta y$ by $\cos 2\pi(\alpha x + \beta y)$. If also we set $w = z^{t/2}$, $x = W^{1/t}r$, $y = W^{1/t}s$, $W = n^{t/2}$, we reduce (2.4) to

$$(2.5)\quad t/2\int_0^n L_t(z^{t/2})z^{t/2-1}dz = n^{t/2+1} \sum_{\alpha=-\infty}^{\infty} \sum_{\beta=-\infty}^{\infty} \iint_{r^t+s^t \leqq 1} (1-r^t-s^t) \cos 2\pi\sqrt{n}(\alpha r \beta s)drds$$

and then (2.1) follows upon multiplication of each side by $n^{(2-t)/t}$.

3. Second auxiliary result. For $t > 1$, we shall obtain from (2.1) the identity

$$(3.1)\quad S(n) = T_1 + T_2 + T_3 + T_4 + T_5$$

where

$$T_1 = c_1 n^2,\ c_1 = \frac{2\Gamma^2(1/t)}{(t+2)\Gamma(2/t)}\ ;$$

$$T_2 = c_2 n^{5/4-1/(2t)} \sum_{\alpha=1}^{\infty} \alpha^{-3/2-1/t} J_{3/2+1/t}(2\pi\sqrt{n}\,\alpha)\ ,$$

where

$$c_2 = \frac{2^{(2t-1)/t} t^{1/t} \Gamma(1/t)}{\pi^{(t+1)/t}}\ ,$$

and $J_r(x)$ is the ordinary Bessel function of order r;

$$T_3 = c_3 n^2 \sum_{\alpha=1}^{\infty} \int_0^1 f(x, t) \cos 2\pi\sqrt{n}\,\alpha x dx\ ,\quad c_3 = \frac{16t}{t+1}\ ,$$

and $f(x, t) = (1 - x^t)^{(t+1)/t} - (t/2)^{(t+1)/t}(1 - x^2)^{(t+1)/t}$;

$$T_4 = -\frac{2t}{\pi\sqrt{t-1}} n \sum_{\alpha=1}^{\infty}\sum_{\beta=1}^{\infty} \frac{J_0(2\pi H\sqrt{n})}{(\alpha\beta)^{(t-2)/(2t-2)} H^{t/(t-1)}}\ ,\quad H = (\alpha^{t/(t-1)} + \beta^{t/(t-1)})^{(t-1)/t}\ ;$$

$$T_5 = \frac{2t}{\pi^2} n \sum_{\alpha=1}^{\infty}\sum_{\beta=1}^{\infty} \frac{1}{\alpha\beta} \int_{-\infty}^{\infty} G(u, \alpha, \beta) \cos 2\pi H\sqrt{n}\, v(u, \alpha, \beta) \cdot v'(u, \alpha, \beta) du\ ,$$

where

$$v(u, \alpha, \beta) = H^{-1} A_0^{-1/t}(u)\ ,\quad A_i(u) = (-1)^i \alpha^{-t}(P\alpha - u)^{t-i} + \beta^{-t}(Q\beta + u)^{t-i}\ ,$$

$$P = \frac{\alpha^{1/(t-1)}}{\alpha^{t/(t-1)} + \beta^{t/(t-1)}}\ ,\quad Q = \frac{\beta^{1/(t-1)}}{\alpha^{t/(t-1)} + \beta^{t/(t-1)}}\ ,$$

$$G(u, \alpha, \beta) = \frac{A_{-1}(u)A_1(u) - A_0^2(u)}{v'(u, \alpha, \beta)A_0^2(u)} - a_{-1}(\alpha, \beta)\operatorname{sgn} u[1 - v^2(u, \alpha, \beta)]^{-1/2}\ ,$$

$$a_{-1}(\alpha, \beta) = \frac{(\alpha\beta)^{t/(2t-2)}}{\sqrt{t-1}(\alpha^{t/(t-1)} + \beta^{t/(t-1)})}\ .$$

In the proof of (3.1) we make use of the following result on Bessel functions [5, p. 366],

$$(3.2)\quad \int_0^1 (1 - x^2)^{m-1/2} \cos Kx dx = \sqrt{\pi} 2^{m-1} K^{-m} \Gamma(m + 1/2) J_m(K) \quad m > -1/2\ .$$

First, it is convenient to break up the double sum in (2.1) as follows,

$$S(n) = \sum_{\alpha=0}\sum_{\beta=0} + \sum_{\substack{\alpha=-\infty\\ \alpha\neq 0}}^{\infty}\sum_{\beta=0} + \sum_{\alpha=0}\sum_{\substack{\beta=-\infty\\ \beta\neq 0}}^{\infty} + \sum_{\alpha=1}^{\infty}\sum_{\beta=1}^{\infty} + \sum_{\alpha=-\infty}^{-1}\sum_{\beta=-\infty}^{-1} + \sum_{\alpha=-\infty}^{-1}\sum_{\beta=1}^{\infty} + \sum_{\alpha=1}^{\infty}\sum_{\beta=-\infty}^{-1} . \tag{3.3}$$

By symmetry this can be written as

$$\begin{aligned} S(n) &= n^2 \iint_{x^t+y^t\leq 1} (1 - x^t - y^t)dxdy \\ &+ 4n^2 \sum_{\alpha=1}^{\infty} \iint_{x^t+y^t\leq 1} (1 - x^t - y^t)\cos 2\pi\sqrt{n}\,\alpha x dxdy \\ &+ 4n^2 \sum_{\alpha=1}^{\infty}\sum_{\beta=1}^{\infty} \iint_{x^t+y^t\leq 1} (1 - x^t - y^t)\cos 2\pi\sqrt{n}(\alpha x + \beta y)dxdy \\ &= S_1 + S_2 + S_3 . \end{aligned} \tag{3.4}$$

S_1 can be evaluated in terms of gamma functions to obtain

$$S_1 = \frac{2\Gamma^2(1/t)}{(t+2)\Gamma(2/t)} n^2 = c_1 n^2 . \tag{3.5}$$

Let I_2 denote the integral in S_2. Then

$$\begin{aligned} I_2 &= 4\int_0^1 \cos 2\pi\sqrt{n}\,\alpha x dx \int_0^{(1-x^t)^{1/t}} (1 - x^t - y^t)dy \\ &= \frac{4t}{t+1}\int_0^1 (1 - x^t)^{(t+1)/t} \cos 2\pi\sqrt{n}\,\alpha x dx \\ &= \frac{4t}{t+1}\left(\frac{t}{2}\right)^{(t+1)/t}\int_0^1 (1 - x^2)^{(t+1)/t} \cos 2\pi\sqrt{n}\,\alpha x dx \\ &+ \frac{c_3}{4}\int_0^1 f(x, t)\cos 2\pi\sqrt{n}\,\alpha x dx \end{aligned} \tag{3.6}$$

by the definition of $f(x, t)$ in (3.1). Applying (3.2) to (3.6) we have

$$S_2 = 4n^2 \sum_{\alpha=1}^{\infty} I_2 = T_2 + T_3 . \tag{3.7}$$

Let I_3 denote the integral in S_3. Then by symmetry

$$I_3 = 2 \iint_{\substack{x^t+y^t\leq 1\\ \alpha x+\beta y\geq 0}} (1 - x^t - y^t)\cos 2\pi\sqrt{n}(\alpha x - \beta y)dxdy . \tag{3.8}$$

The transformation

$$x = Hv(P - u/\alpha) , \quad y = Hv(Q + u/\beta) \tag{3.9}$$

transforms $x^t + y^t = 1$ into

$$v = H^{-1}A_0^{-1/t}(u) \tag{3.10}$$

where H, P, Q, and $A_i(u)$ are defined in (3.1). The transformation (3.9) is one to one for $\alpha x + \beta y \geqq 0$ and the absolute value of the Jacobian is

$$J\left(\frac{x,\, y}{v,\, u}\right) = \frac{H^2 v}{\alpha\beta}\ . \tag{3.11}$$

The graph of (3.10) resembles that of $v = 1/(1 + u^2)$ except that the curve is not symmetric to the v axis unless $t = 2$. The curve has a relative maximum at (0, 1).

Applying (3.9) to (3.8) we transform $x^t + y^t \leqq 1$ and $\alpha x + \beta y \geqq 0$ into $v \leqq H^{-1}A_0^{-1/t}(u)$ and $v \geqq 0$ respectively, so that (3.8) becomes

$$I_3 = \frac{2H^2}{\alpha\beta}\int_{-\infty}^{\infty} du \int_0^{v(u)} [1 - H^t v^t A_0(u)] v \cos 2\pi H\sqrt{n}\, v dv\ . \tag{3.12}$$

Upon integration by parts with respect to v, the integrated terms vanish and we obtain

$$\begin{aligned} I_3 &= -\frac{H}{\pi\sqrt{n}\,\alpha\beta}\int_{-\infty}^{\infty} du \int_0^{v(u)} [1 - (t + 1)H^t v^t A_0(u)] \sin 2\pi H\sqrt{n}\, v dv \\ &= -\frac{H}{\pi\sqrt{n}\,\alpha\beta}\int_0^1 \sin 2\pi H\sqrt{n}\, v dv \int_{u_-(v)}^{u_+(v)} [1 - (t + 1)H^t v^t A_0(u)] du \end{aligned} \tag{3.13}$$

where $u_+(v)$ and $u_-(v)$ refer to the first and second quadrant branches of (3.10) respectively. Since

$$A_i(u) = (-1)^i \alpha^{-t}(P\alpha - u)^{t-i} + \beta^{-t}(Q\beta + u)^{t-i}, \qquad \frac{d}{du}A_i(u) = (t - i)A_{i+1}(u)\ , \tag{3.14}$$

we can write (3.13) as

$$\begin{aligned} I_3 = &-\frac{H}{\pi\sqrt{n}\,\alpha\beta}\int_0^1 [u_+(v) - H^t v^t A_{-1}(u_+(v))] \sin 2\pi H\sqrt{n}\, v dv \\ &-\frac{H}{\pi\sqrt{n}\,\alpha\beta}\int_0^1 [-u_-(v) + H^t v^t A_{-1}(u_-(v))] \sin 2\pi H\sqrt{n}\, v dv\ . \end{aligned} \tag{3.15}$$

By the change of variable (3.10) this can be written as

$$I_3 = \frac{H}{\pi\sqrt{n}\,\alpha\beta}\int_{-\infty}^{\infty}\left[u - \frac{A_{-1}(u)}{A_0(u)}\right] \sin 2\pi H\sqrt{n}\, v(u) \cdot v'(u) du\ . \tag{3.16}$$

From (3.14) we obtain

$$u - \frac{A_{-1}(u)}{A_0(u)} = \frac{P\alpha^{1-t} - Q\beta^{1-t}}{\alpha^{-t} + \beta^{-t}} + O\left(\frac{1}{u}\right) \tag{3.17}$$

for large u, so that upon integrating by parts again we obtain

$$I_3 = \frac{t}{2\pi^2 n\alpha\beta}\int_{-\infty}^{\infty} F(u)\cos 2\pi H\sqrt{n}\,v(u)\cdot v'(u)du \tag{3.18}$$

where

$$F(u) = F(u, \alpha, \beta) = \frac{A_{-1}(u)A_1(u) - A_0^2(u)}{v'(u)A_0^2(u)}\,. \tag{3.19}$$

The function $a_{-1}\operatorname{sgn} u[1 - v^2(u)]^{-1/2}$ is an asymptotic equivalent of $F(u)$ in the neighborhood of $(0, 1)$, even though $v(0) = 1$ and $v'(0) = 0$, if $a_{-1} = a_{-1}(\alpha, \beta)$ is determined from

$$\begin{aligned} a_{-1} &= \lim_{u\to 0+} F(u)\sqrt{1 - v^2(u)} = \lim_{u\to 0+}\frac{\sqrt{1-v^2}}{-v'} \\ &= \lim_{u\to 0+}\frac{vv'(1-v^2)^{-1/2}}{v''} = \frac{1}{v''(0)}\lim_{u\to 0+}\frac{v'}{\sqrt{1-v^2}} \\ &= \frac{-1}{v''(0)a_{-1}} = \frac{1}{\sqrt{|v''(0)|}}\,. \end{aligned} \tag{3.20}$$

From (3.10) and (3.14) we obtain

$$v''(u) = -H^{-1}A_0^{-(1+2t)/t}(u)[-(t+1)A_1^2(u) + (t-1)A_0(u)A_2(u)] \tag{3.21}$$

from which a_{-1}, as given in (3.1), can be determined.

We now write (3.18) as

$$\begin{aligned} I_3 &= \frac{ta_{-1}}{2\pi^2 n\alpha\beta}\int_{-\infty}^{\infty}\operatorname{sgn} u[1 - v^2(u)]^{-1/2}\cos 2\pi H\sqrt{n}\,v(u)\cdot v'(u)du \\ &\quad + \frac{t}{2\pi^2 n\alpha\beta}\int_{-\infty}^{\infty}[F(u) - a_{-1}\operatorname{sgn} u[1-v^2(u)]^{-1/2}]\cos 2\pi H\sqrt{n}v(u)\cdot v'(u)du \\ &= -\frac{ta_{-1}}{\pi^2 n\alpha\beta}\int_0^1 (1 - v^2)^{-1/2}\cos 2\pi H\sqrt{n}\,v dv \\ &\quad + \frac{t}{2\pi^2 n\alpha\beta}\int_{-\infty}^{\infty} G(u)\cos 2\pi H\sqrt{n}\,v(u)\cdot v'(u)du \end{aligned} \tag{3.22}$$

where $G(u) = G(u, \alpha, \beta)$ is defined in (3.1). Applying (3.2) to (3.22) we obtain

$$S_3 = 4n^2\sum_{\alpha=1}^{\infty}\sum_{\beta=1}^{\infty} I_3 = T_4 + T_5\,. \tag{3.23}$$

Collecting the results of (3.4), (3.5), (3.7), and (3.23), we have (3.1).

4. Convergence investigations. We next prove that the double series in (2.1) is absolutely convergent. We write (3.18) as

$$I_3 = \frac{t}{2\pi^2 n\alpha\beta}\left(\int_{-\infty}^{0} + \int_{0}^{\sigma} + \int_{\sigma}^{P\alpha} + \int_{P\alpha}^{\infty}\right) \tag{4.1}$$

$$= \frac{t}{2\pi^2 n\alpha\beta}(I_4 + I_5 + I_6 + I_7)$$

where $0 < \sigma < P\alpha$.

First we consider

$$I_7 = \int_{P\alpha}^{\infty} F(u)\cos 2\pi H\sqrt{n}\,v(u)\cdot v'(u)du\,. \tag{4.2}$$

By (3.14) and (3.19) we have

$$F(u) = \frac{H(P\alpha-u)^{t-1}(Q\beta+u)^{t-1}}{(\alpha\beta)^t A_5^{1-1/t}(u)A_1(u)} = \frac{-H[\alpha^{-t}(u-P\alpha)^t+\beta^{-t}(u+Q\beta)^t]^{(1-t)/t}}{\alpha^t(u-P\alpha)^{1-t}+\beta^t(u+Q\beta)^{1-t}} \tag{4.3}$$

From (4.3) we find that

$$\frac{dF(u)}{du} = \frac{(1-t)H}{(\alpha\beta)^t}\left(\frac{(u-P\alpha)^t}{\alpha^t} + \frac{(u+Q\beta)^t}{\beta^t}\right)^{(1-2t)/t} \tag{4.4}$$

$$\times\left(\frac{-\beta^{2t}(u-P\alpha)^{2t-1}+\alpha^{2t}(u+Q\beta)^{2t-1}}{(u-P\alpha)^t(u+Q\beta)^t[\alpha^t(u-P\alpha)^{1-t}+\beta^t(u+Q\beta)^{1-t}]^2}\right).$$

From (4.3) and (4.4) we derive certain information about the graph of F(u), namely,

$$\begin{aligned} &F(u) > 0,\ F'(u) < 0,\ 0 < u < P\alpha\,; \\ &F'(P\alpha) = \infty,\ 1 < t < 2\,;\ F'(P\alpha) = 0,\ 2 < t\,; \\ &F(u) < 0,\ P\alpha < u < \infty\,; \\ &F'(u) = 0,\ u = u_1,\ P\alpha < u_1 < \infty,\ \beta > \alpha\,; \\ &F'(u) < 0,\ P\alpha < u < \infty,\ \beta \leqq \alpha\,. \end{aligned} \tag{4.5}$$

The point (u_1, v_1) is a relative minimum and from (4.3) and (4.4) we find that

$$u_1 = \frac{Q\beta\alpha^{(2t)/(2t-1)} + P\alpha\beta^{(2t)/(2t-1)}}{\beta^{(2t)/(2t-1)} - \alpha^{(2t)/(2t-1)}}, \tag{4.6}$$

$$v_1 = F(u_1) = -H(\alpha^{t/(2t-1)} + \beta^{t/(2t-1)})^{-(2t-1)/t}\,.$$

Thus by (4.5) and the second mean value theorem for integrals we have, for $\beta > \alpha$, and $P\alpha \leqq \xi_1 < u_1 < \xi_2 \leqq \infty$,

$$I_7 = \int_{P\alpha}^{u_1} + \int_{u_1}^{\infty} = F(u_1)\int_{\xi_1}^{u_1} + F(u_1)\int_{u_1}^{\xi_2} \tag{4.7}$$

$$= F(u_1)\int_{\xi_1}^{\xi_2} \cos 2\pi H\sqrt{n}\,v(u)\cdot v'(u)du$$
$$= O\{F(u_1)H^{-1}n^{-1/2}\} = O\{(\alpha^{t/(2t-1)} + \beta^{t/(2t-1)})^{-(2t-1)/t}n^{-1/2}\}$$
$$= O\{(n\alpha\beta)^{-1/2}\}$$

by the inequality $x^2 + y^2 \geqq 2xy$, $x > 0$, $y > 0$. Similarly, for $\beta \leqq \alpha$, and $P\alpha \leqq \xi_3 < \infty$, we have

$$(4.8)\qquad I_7 = \int_{P\alpha}^{\infty} = F(\infty)\int_{\xi_3}^{\infty} \cos 2\pi H\sqrt{n}\,v(u)\cdot v'(u)du$$
$$= O\{F(\infty)H^{-1}n^{-1/2}\} = \{(\alpha^{-t} + \beta^{-t})^{(1-t)/t}(\alpha^t + \beta^t)^{-1}n^{-1/2}\}$$
$$= O\{(\alpha\beta)^{t-1}(\alpha^t + \beta^t)^{-(2t-1)/t}n^{-1/2}\} = O\{(n\alpha\beta)^{-1/2}\}\ .$$

We next consider I_5 in (4.1). By (4.3) we can write

$$(4.9)\qquad I_5 = \int_0^{\sigma} F_1(u)\cos 2\pi H\sqrt{n}\,v(u)du$$

where

$$(4.10)\qquad -F_1(u) = \frac{(pq)^{t-1}}{(\alpha\beta)^t A_0^2(u)}\,,\ p = P\alpha - u,\ q = Q\beta + u\ ,$$
$$= \frac{A_0^{-2/t}(u)}{\alpha\beta}\cdot\frac{[(pq)/(\alpha\beta)]^{t-1}}{[(p/\alpha)^t + (q/\beta)^t]^{2(t-1)/t}} < \frac{A_0^{-2/t}(u)}{\alpha\beta}\cdot\frac{1}{2^{2(t-1)/t}}$$
$$< \frac{A_0^{-2/t}(0)}{\alpha\beta 2^{2(t-1)/t}} = O\left(\frac{H^2}{\alpha\beta}\right)\ .$$

Therefore

$$(4.11)\qquad I_5 = O\left(\frac{H^2}{\alpha\beta}\int_0^{\sigma} du\right) = O\left(\frac{H^2\sigma}{\alpha\beta}\right)\ .$$

Turning next to I_6 in (4.1), we note that by the first line of (4.5) we can use the second mean value theorem to write, for some ξ_4 satisfying $\sigma < \xi_4 \leqq P\alpha$,

$$(4.12)\qquad I_6 = F(\sigma)\int_{\sigma}^{\xi_4} \cos 2\pi H\sqrt{n}\,v(u)\cdot v'(u)du = O\left(\frac{F(\sigma)}{H\sqrt{n}}\right)\ .$$

To examine the question of the order of $F(u)$ in $0 < u < P\alpha$ we use (4.3) with, $p = P\alpha - u$, $q = Q\beta + u$, and write

$$(4.13)\qquad F(u) = \frac{H}{(\alpha\beta)^{1/2}}\cdot\frac{[(pq)/(\alpha\beta)]^{(t-1)/2}}{[(p/\alpha)^t + (q/\beta)^t]^{(t-1)t}}\cdot$$
$$\cdot\frac{1}{-(\beta/\alpha)^{t/2}(p/q)^{(t-1)/2} + (\alpha/\beta)^{t/2}(q/p)^{(t-1)/2}}$$
$$\leqq \frac{H}{(\alpha\beta)^{1/2}}\cdot\frac{1}{2^{(t-1)/2}}\cdot\frac{1}{F_2(u)}\ .$$

Since $F_2(0) = 0$ and

$$(4.14)\quad F_3(u) = \frac{dF_2(u)}{du} = \frac{t-1}{2}\left[\left(\frac{\beta}{\alpha}\right)^{t/2}\left(\frac{p}{q}\right)^{(t-3)/2}\frac{1}{q^2} + \left(\frac{\alpha}{\beta}\right)^{t/2}\left(\frac{q}{p}\right)^{(t-3)/2}\frac{1}{p^2}\right],$$

we have, by the mean value theorem,

$$(4.15)\quad F(u) \leqq \frac{H\lambda}{(\alpha\beta)^{1/2}} \cdot \frac{(t-1)/2}{F_3(u_3)u},\ \lambda = \frac{2^{(5-t)/2}}{t-1},\ 0 < u < P\alpha,\ 0 < u_3 < P\alpha\ .$$

Setting $p_3 = P\alpha - u_3$, $q_3 = Q\beta + u_3$, we obtain

$$\begin{aligned}(4.16)\quad F(u) &\leqq \frac{H\lambda}{(\alpha\beta)^{1/2}} \cdot \frac{p_3q_3}{[(\beta/\alpha)^{t/2}(p_3/q_3)^{(t-1)/2} + (\alpha/\beta)^{t/2}(q_3/p_3)^{(t-1)/2}]u} \\ &\leqq \frac{H\lambda}{(\alpha\beta)^{1/2}} \cdot \frac{p_3q_3}{[(\beta/\alpha)^{t/2} + (\alpha/\beta)^{t/2}]u} \leqq \frac{H\lambda}{(\alpha\beta)^{1/2}} \cdot \frac{4}{[(\beta/\alpha)^{t/2} + (\alpha/\beta)^{t/2}]u} \\ &= O\left\{\frac{H(\alpha\beta)^{(t-1)/2}}{(\alpha^t + \beta^t)u}\right\}.\end{aligned}$$

Hence combining (4.11), (4.12), and (4.16), we obtain

$$(4.17)\quad I_5 + I_6 = O\left(\frac{H^2\sigma}{\alpha\beta}\right) + O\left(\frac{(\alpha\beta)^{(t-1)/2}}{(\alpha^t + \beta^t)\sigma n^{1/2}}\right) = O\left(\frac{H(\alpha\beta)^{(t-3)/4}n^{-1/4}}{(\alpha^t + \beta^t)^{1/2}}\right).$$

In the further analysis of $I_5 + I_6$ we use the inequalities,

$$(4.18)\quad 1 + x^m < (1 + x)^m,\ 0 < x < 1,\ m > 1\ ,$$

$$(4.19)\quad (x + 1)^m < 2^{m-1}(x^m + 1),\ x > 1,\ m > 1\ .$$

In (4.17) suppose $1<t\leqq 2$. Since $H=(\alpha^{t/(t-1)}+\beta^{t/(t-1)})^{(t-1)/t}$ and $t/(t-1)>t$, we have by (4.18), $H < (\alpha^t + \beta^t)^{1/t}$, and therefore, for $1 < t \leqq 2$, we have

$$(4.20)\quad \frac{H(\alpha\beta)^{(t-3)/4}}{(\alpha^t + \beta^t)^{1/2}} < \frac{(\alpha^t+\beta^t)^{(2-t)/(2t)}}{(\alpha\beta)^{(3-t)/4}} < \frac{(\alpha + \beta)^{(2-t)/2}}{(\alpha\beta)^{(t-1)/4}(\alpha\beta)^{(2-t)/2}} < \frac{2^{(2-t)/2}}{(\alpha\beta)^{(t-1)/4}}\ .$$

Hence from (4.17) and (4.20) we have, for $1 < t \leqq 2$,

$$(4.21)\quad I_5 + I_6 = O\{(\alpha\beta)^{-(t-1)/4}n^{-1/4}\}\ .$$

If $t > 2$ is (4.17), then $t > t/(t-1)$ and so by (4.19) we have

$$\begin{aligned}(4.22)\quad \frac{H(\alpha\beta)^{(t-3)/4}}{(\alpha^t + \beta^t)^{1/2}} &= \frac{(\alpha^{t/(t-1)}+\beta^{t/(t-1)})^{(t-1)/t}}{(\alpha^t + \beta^t)^{1/t}} \cdot \frac{(\alpha\beta)^{(t-3)/4}}{(\alpha^t+\beta^t)^{(t-2)/(2t)}} \\ &< 2^{(t-2)/t} \cdot \frac{(\alpha\beta)^{(t-3)/4}}{(\alpha^t + \beta^t)^{(t-2)/(2t)}} \\ &< \frac{2^{(t-2)/t}(\alpha\beta)^{(t-3)/4}}{2^{(t-2)/(2t)}(\alpha\beta)^{(t-2)/4}} = \frac{2^{(t-2)/(2t)}}{(\alpha\beta)^{1/4}}\ .\end{aligned}$$

Hence from (4.17) and (4.22) we have, for $t > 2$,

$$I_5 + I_6 = O\{(n\alpha\beta)^{-1/4}\} \ . \tag{4.23}$$

By (3.10) $v(-u, \alpha, \beta) = v(u, \alpha, \beta)$ so that an estimate for $I_5 + I_6 + I_7$ holds also for I_4 in (4.1). By this fact, and the results of (4.7), (4.8), (4.21), and (4.23), it now follows that for S_3, defined by (3.4), (3.23), and (4.1), we have,

$$\begin{aligned} S_3 &= \frac{2tn}{\pi^2} \sum_{\alpha=1}^{\infty} \sum_{\beta=1}^{\infty} \frac{1}{\alpha\beta} \int_{-\infty}^{\infty} F(u) \cos 2\pi H\sqrt{n}\, v(u) \cdot v'(u) du \\ &= O(n^{3/4}),\ t > 1 \ , \end{aligned} \tag{4.24}$$

the double series being absolutely convergent.

Integrating by parts and applying the second mean value theorem, we have, from (3.6), for $x_1 = [(t-1)/t]^{1/t}$,

$$\begin{aligned} S_2 &= \frac{16t}{t+1} n^2 \sum_{\alpha=1}^{\infty} \int_0^1 (1 - x^t)^{(t+1)/t} \cos 2\pi\sqrt{n}\, \alpha x dx \\ &= \frac{8t}{\pi} n^{3/2} \sum_{\alpha=1}^{\infty} \frac{1}{\alpha} \int_0^1 (1 - x^t)^{1/t} x^{t-1} \sin 2\pi\sqrt{n}\, \alpha x dx \\ &= \frac{8t}{\pi} n^{3/2} \sum_{\alpha=1}^{\infty} \frac{1}{\alpha} \left\{ \int_0^{x_1} + \int_{x_1}^1 \right\} \\ &= \frac{8t}{\pi} n^{3/2} \sum_{\alpha=1}^{\infty} \frac{1}{\alpha} \left\{ (1 - x_1^t)^{1/t} x_1^{t-1} \int_{\xi_5}^{x_1} \sin 2\pi\sqrt{n}\, \alpha x dx \right. \\ &\qquad \left. + (1 - x_1^t)^{1/t} x_1^{t-1} \int_{x_1}^{\xi_6} \sin 2\pi\sqrt{n}\, \alpha x dx \right\} \\ &= \frac{8t}{\pi} n^{3/2} \sum_{\alpha=1}^{\infty} O\left(\frac{1}{\sqrt{n}\, \alpha^2}\right) = O(n),\ t > 1 \ , \end{aligned} \tag{4.25}$$

the series being absolutely convergent. The absolute convergence of the double series in (2.1) now follows from the results leading to (4.24) and (4.25).

5. Proof of (1.1). Finally we deduce (1.1) from (3.1). We make use of the asymptotic expansion for the general Bessel function, namely [5, p. 368],

$$J_m(K) = \sqrt{\frac{2}{\pi K}} \cos\left(K - \frac{m\pi}{2} - \frac{\pi}{4}\right) + O(K^{-3/2}) \ , \tag{5.1}$$

for large K and m independent of K.

By (5.1) and the absolute convergence of the sum we have

$$(5.2)\quad T_2 = c_2 n^{5/4-1/(2t)} \sum_{\alpha=1}^{\infty} \alpha^{-3/2-1/t} \left\{ \frac{\cos[2\pi\sqrt{n}\,\alpha - \pi(1+1/(2t))]}{(\pi^2\sqrt{n}\,\alpha)^{1/2}} + O(n^{-3/4}\alpha^{-3/2}) \right\}$$

$$= -\frac{c_2}{\pi} n^{1-1/(2t)} \sum_{\alpha=1}^{\infty} \alpha^{-2-1/t} \cos(2\pi\sqrt{n}\,\alpha - \pi/(2t)) + O(n^{1/2-1/(2t)}) .$$

In $T_3 f'(0, t) = 0$ and $f^{(k)}(1, t) = 0$, $k = 0, 1, 2$. Hence if we integrate by parts twice the integrated terms vanish and we have left

$$(5.3)\qquad T_3 = -\frac{4t}{\pi^2(t+1)} n \sum_{\alpha=1}^{\infty} \frac{1}{\alpha^2} \int_0^1 f''(x, t) \cos 2\pi\sqrt{n}\,\alpha x dx .$$

$f''(x, t)$ is continuous in $0 \leqq x \leqq 1$ and independent of n and α and so it has a finite number, independent of n and α, of relative and absolute extrema whose values are also independent of n and α. Hence dividing the interval of integration into pieces in which $f''(x, t)$ is monotonic, we obtain by the second mean value theorem, for appropriate ξ_j, ξ'_j, ξ'_{j+1} in the interval from 0 to 1, the ξ's depending on n and α, the result,

$$(5.4)\quad T_3 = -\frac{4t}{\pi^2(t+1)} n^{3/4} \sum_{\alpha=1}^{\infty} \frac{1}{\alpha^2} \sum_j f''(\xi_j, t) \int_{\xi'_j}^{\xi'_{j+1}} \cos 2\pi\sqrt{n}\,\alpha x dx = O(\sqrt{n}) .$$

Applying (5.1) to T_4 we obtain

$$T_4 = -\frac{2t}{\pi^2\sqrt{t-1}} n^{3/4} \sum_{\alpha=1}^{\infty} \sum_{\beta=1}^{\infty} \frac{\cos(2\pi H\sqrt{n} - \pi/4)}{(\alpha\beta)^{(t-2)/(2t-2)} H^{(3t-1)/(2t-2)}}$$
$$- \frac{2tn}{\pi\sqrt{t-1}} \sum_{\alpha=1}^{\infty} \sum_{\beta=1}^{\infty} \frac{O\{(\alpha^{t/(t-1)} + \beta^{t/(t-1)})^{-(3t-3)/(2t)} n^{-3/4}\}}{(\alpha\beta)^{(t-1)/(2t-2)} (\alpha^{t/(t-1)} + \beta^{t/(t-1)})} .$$

Since

$$(\alpha^{t/(t-1)} + \beta^{t/(t-1)})^{-(5t-3)/(2t)} \leqq 2^{-(5t-3)/(2t)} (\alpha\beta)^{-(5t-3)/(4t-4)} ,$$

the double series are absolutely convergent so that

$$(5.5)\quad T_4 = -\frac{2t}{\pi^2\sqrt{t-1}} n^{3/4} \sum_{\alpha=1}^{\infty} \sum_{\beta=1}^{\infty} \frac{\cos(2\pi H\sqrt{n} - \pi/4)}{(\alpha\beta)^{(t-2)/(2t-2)} H^{(3t-1)/(2t-2)}} + O(n^{1/4}) .$$

Next we consider T_5. We have shown that $-T_4$ and S_3 are absolutely convergent double series for $t > 1$ and hence so is their term by term sum which is identical with T_5. We break up the interval of integration in T_5 into a finite number, independent of n, α, β, of subintervals in which $G(u, \alpha, \beta)$ is monotonic and write

$$(5.6)\quad T_5 = \frac{2tn}{\pi^2} \sum_{\alpha=1}^{\infty} \sum_{\beta=1}^{\infty} \frac{1}{\alpha\beta} \sum_j \int_{\xi_j}^{\xi_{j+1}} G(u, \alpha, \beta) \cos 2\pi H\sqrt{n}\, v(u) \cdot v'(u) du .$$

Now $G(u, \alpha, \beta)$ is continuous in each $\xi_j \leqq u \leqq \xi_{j+1}$. The only doubt arises, at $u = 0$ where $v'(u) = (1 - v^2)^{1/2} = 0$, and at $u = \infty$ where $v'(u) = 0$. But, using the definitions in (3.1) and evaluating an indeterminate form, we obtain

$$(5.7) \qquad G(0+, \alpha, \beta) = \frac{-HA_{-1}(0)}{A_0^{1-1/t}(0)} - \lim_{u \to 0+}\left(\frac{1}{v'(u)} + \frac{a_{-1}}{\sqrt{1 - v^2(u)}}\right)$$
$$= \frac{-HA_{-1}(0)}{A_0^{1-1/t}(0)} + O\left(\frac{\alpha^{t/(t-1)} - \beta^{t/(t-1)}}{\alpha^{t/(t-1)} + \beta^{t/(t-1)}}\right)$$

which is bounded. On the other hand, by (4.3),

$$(5.8) \qquad G(\infty, \alpha, \beta) = -H(\alpha\beta)^{t-1}(\alpha^t + \beta^t)^{(1-2t)/t} - a_{-1},$$

which is also bounded.

Applying the second mean value theorem to (5.6) we obtain

$$(5.9) \quad T_5 = \frac{2tn}{\pi^2} \sum_{\alpha=1}^{\infty} \sum_{\beta=1}^{\infty} \frac{1}{\alpha\beta} \sum_j G(\zeta_j', \alpha, \beta) \int_{\zeta_j}^{\zeta_{j+1}} \cos 2\pi H\sqrt{n}\, v(u) \cdot v'(u) du$$

for appropriate ζ_j', ζ_j, ζ_{j+1} in the interval from ξ_j to ξ_{j+1}. Further we have

$$(5.10) \qquad T_5 = \frac{2tn}{\pi^2} \sum_{\alpha=1}^{\infty} \sum_{\beta=1}^{\infty} \frac{1}{\alpha\beta} \sum_j G(\zeta_j', \alpha, \beta) \frac{O(1)}{H\sqrt{n}}$$
$$= \frac{2t\sqrt{n}}{\pi^2} \sum_{\alpha=1}^{\infty} \sum_{\beta=1}^{\infty} \frac{1}{\alpha\beta H} \sum_j G(\zeta_j', \alpha, \beta) O(1)$$
$$= O(\sqrt{n})$$

by the absolute convergence of the double series.

The relation (1.1) now follows from (3.1), (3.5), (5.2), (5.4), (5.5), and (5.10).

References

1. P. Bachmann, *Zahlentheorie*, **2**, Teubner, (1894).
2. D. Cauer, *Neue Anwendungen der Pfeifferschen Methode zur Abschätzung zahlentheoretischer Funktionen*, Diss. Göttingen, **55** s. 8° (1914).
3. J. G. van der Corput, *Over roosterpunten in het platte vlak*, Diss. Leiden, Groningen: Noordhoff, **128** s. 8° (1919).
4. E. Landau, *Vorlesungen über Zahlentheorie*, **2**, Chelsea, (1947).
5. Whittaker and Watson, *Modern Analysis*, Macmillan, (1946).

University of California at Los Angeles

Pacific Journal of Mathematics

Vol. 8, No. 4 June, 1958