Pacific Journal of Mathematics

IN THIS ISSUE—

Maynard G. Arsove, *The Paley-Wiener theorem in metric linear spaces* 365
Robert (Yisrael) John Aumann, *Acceptable points in games of perfect information* 381
A. V. Balakrishnan, *Fractional powers of closed operators and the semigroups generated by them* 419
Dallas O. Banks, *Bounds for the eigenvalues of some vibrating systems* 439
Billy Joe Boyer, *On the summability of derived Fourier series* 475
Robert Breusch, *An elementary proof of the prime number theorem with remainder term* 487
Edward David Callender, Jr., *Hölder continuity of n-dimensional quasi-conformal mappings* 499
L. Carlitz, *Note on Alder's polynomials* 517
P. H. Doyle, III, *Unions of cell pairs in* E^3 521
James Eells, Jr., *A class of smooth bundles over a manifold* 525
Shaul Foguel, *Computations of the multiplicity function* 539
James G. Glimm and Richard Vincent Kadison, *Unitary operators in* C^**-algebras* 547
Hugh Gordon, *Measure defined by abstract* L_p *spaces* 557
Robert Clarke James, *Separable conjugate spaces* 563
William Elliott Jenner, *On non-associative algebras associated with bilinear forms* 573
Harold H. Johnson, *Terminating prolongation procedures* 577
John W. Milnor and Edwin Spanier, *Two remarks on fiber homotopy type* 585
Donald Alan Norton, *A note on associativity* 591
Ronald John Nunke, *On the extensions of a torsion module* 597
Joseph J. Rotman, *Mixed modules over valuations rings* 607
A. Sade, *Théorie des systèmes demosiens de groupoï des* 625
Wolfgang M. Schmidt, *On normal numbers* 661
Berthold Schweizer, Abe Sklar and Edward Oakley Thorp, *The metrization of statistical metric spaces* 673
John P. Shanahan, *On uniqueness questions for hyperbolic differential equations* 677
A. H. Stone, *Sequences of coverings* 689
Edward Oakley Thorp, *Projections onto the subspace of compact operators* 693
L. Bruce Treybig, *Concerning certain locally peripherally separable spaces* 697
Milo Wesley Weaver, *On the commutativity of a correspondence and a permutation* 705
David Van Vranken Wend, *On the zeros of solutions of some linear complex differential equations* 713
Fred Boyer Wright, Jr., *Polarity and duality* 723

Vol. 10, No. 2 **October, 1960**

PACIFIC JOURNAL OF MATHEMATICS

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

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

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

Subscriptions, orders for back numbers, and changes of address should be sent to Pacific Journal of Mathematics, 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 PALEY-WIENER THEOREM IN METRIC LINEAR SPACES

MAYNARD G. ARSOVE

1. Introduction. By a *basis* in a topological linear space $\mathscr{T}$ we mean a sequence $\{x_n\}$ of points of $\mathscr{T}$ such that to every x in $\mathscr{T}$ there corresponds a unique sequence $\{a_n\}$ of scalars for which

$$x = \sum_{n=1}^{\infty} a_n x_n \ .$$

Denoting the coefficient functionals here by φ_n, we can rewrite this as

$$x = \sum_{n=1}^{\infty} \varphi_n(x) x_n \ . \tag{1.1}$$

If it happens that all φ_n are continuous on $\mathscr{T}$, the basis will be referred to as a *Schauder basis*. Every basis in a Fréchet space [**14**, pp. 59, 110] is known to be a Schauder basis (see Newns [**21**], pp. 431–432), and it will be shown here that the same holds for bases in an arbitrary complete metric linear space over the real or complex field.

The classical Paley-Wiener theorem asserts that for $\mathscr{T}$ a Banach space, all sequences which sufficiently closely approximate bases must themselves be bases. A more precise statement of the theorem is obtained by replacing $\mathscr{M}$ in Theorem 1 by a Banach space $\mathscr{B}$.

The bibliography at the end of the present paper includes a chronological listing of articles on the Paley-Wiener theorem, and we give now a brief résumé of its history. As originally presented in 1934 by Paley and Wiener [**1**, p. 100], the theorem was derived specifically for the Hilbert space L^2. Then, in applying the theorem to the Pincherle basis problem [**2**, p. 469], Boas observed in 1940 that the proof of Paley and Wiener remains valid for Banach spaces. Boas also succeeded in simplifying a portion of the proof. However, the first really elementary proof of the theorem was published in 1949 by Schafke [**8**], to whom conclusion (3) is due. The remaining articles on the Paley-Wiener theorem deal mainly with various generalizations of condition (2.1) for Hilbert spaces.

From the viewpoint of modern functional analysis, the key to theorems of Paley-Wiener type lies in the inversion of an operator $I+T$ by means of a geometric series in T. This crucial observation was made by Buck [**15**, p. 410] in 1953.[1]

Received July 28, 1958, and in revised form June 22, 1959. The research reported upon here was supported in part by the National Science Foundation.

[1] The same technique was used also in [**9**], the author having been unaware of the earlier remarks of Buck. A further application (to generalized bases) appears in [**12**].

Our purpose in the present note is to utilize the operator technique in deriving a number of variants of the Paley-Wiener theorem. For reference, we begin by sketching a proof of the theorem itself for complete metric linear spaces. The ensuing variants then have in common the hypothesis that $\{x_n\}$ be a Schauder basis and $\{y_n\}$ a sequence triangular with respect to $\{x_n\}$. This is evidently motivated by the case of Pincherle bases in spaces of analytic functions (see, for example, [9]), and we conclude, in fact, with a generalization to arbitrary Fréchet spaces of the theorem of Boas [2, p. 447, Theorem 4.1] on Pincherle bases.

The author is indebted to Professor Robert C. James for reading the manuscript and suggesting a number of important simplifications. In particular, Theorem 2 replaces a weaker theorem of the original manuscript.

2. The proof for metric linear spaces. In what follows, we shall denote by $\mathscr{M}$ a complete metric linear space over the real or complex field and employ the notation of Banach:

$$\|x\| = \rho(x, 0) \qquad (x \in \mathscr{M}),$$

where ρ is the metric on $\mathscr{M}$. It will be assumed further that ρ is translation invariant.[2]

With these conventions the Paley-Wiener theorem can be formulated as

THEOREM 1. *Let $\{x_n\}$ and $\{y_n\}$ be sequences in $\mathscr{M}$, and let λ be a real number $(0 < \lambda < 1)$ such that*

$$\left\|\sum_{n=1}^{m} a_n(y_n - x_n)\right\| \leq \lambda \left\|\sum_{n=1}^{m} a_n x_n\right\| \tag{2.1}$$

holds for all finite sequences $a_1, a_2, \cdots, a_m$ of scalars. Then

(1) *if $\{x_n\}$ is total in $\mathscr{M}$, so is $\{y_n\}$;*

(2) *if $\{x_n\}$ is a basis in $\mathscr{M}$, so is $\{y_n\}$, and the coefficients in any expansion $\sum b_n y_n$ satisfy*

$$\left\|\sum_{n=1}^{\infty} b_n x_n\right\| \leq \frac{1}{1-\lambda} \left\|\sum_{n=1}^{\infty} b_n y_n\right\| ; \tag{2.2}$$

(3) *if $\{x_n\}$ is a basis in $\mathscr{M}$, there exists an automorphism[3] A on $\mathscr{M}$ such that $y_n = Ax_n$ $(n = 1, 2, \cdots)$.*

---

[2] A translation-invariant metric yielding the original topology always exists (see, for example, [**19**, p. 34]).

[3] The term *automorphism* is used to designate any linear homeomorphic mapping of the space onto itself. By the open mapping theorem [**13**, p. 41, Theorem 5] every one-to-one continuous linear mapping of $\mathscr{M}$ onto itself is an automorphism on $\mathscr{M}$.

Proof. For convenience we consider first the case in which $\{x_n\}$ is a basis. Condition (2.1) then allows us to define a continuous linear operator T on $\mathscr{M}$ as

$$Tx = \sum_{n=1}^{\infty} \varphi_n(x) \cdot (y_n - x_n) \tag{2.3}$$

and yields the inequality

$$\| T^n x \| \leq \lambda^n \| x \| \qquad (n = 0, 1, \cdots).$$

By comparison with the corresponding geometric series in λ we infer convergence of the operator series

$$U = \sum_{n=0}^{\infty} (-T)^n \tag{2.4}$$

and obtain the inequality

$$\| Ux \| \leq (1 - \lambda)^{-1} \| x \| . \tag{2.5}$$

Hence, the linear operator U is continuous on $\mathscr{M}$.

For any x in $\mathscr{M}$ the element y of $\mathscr{M}$ defined by $y = Ux$ has the evident property that $x = (I + T)y$, where I is the identity operator. From $y = \sum b_n x_n$ it therefore follows that $x = \sum b_n y_n$, and this proves that $\{y_n\}$ spans $\mathscr{M}$ in the infinite-series sense. That $\{y_n\}$ is linearly independent in the infinite-series sense can then be seen by rewriting (2.5) in the form (2.2). Assertions (2) and (3) are thereby established, the latter with A taken as $I + T (= U^{-1})$.

No essential change in the above argument is required to prove (1). We can clearly presume the x_n to be finitely linearly independent and replace the infinite series in (2.3) by corresponding finite sums. Thus defined on a dense subset of $\mathscr{M}$, T is then extended to all of $\mathscr{M}$ in the usual fashion.

It should perhaps be mentioned that the automorphism A in (3) is uniquely determined by the way it correlates the basis elements x_n and y_n. In fact,

$$Ax = \sum_{n=1}^{\infty} \varphi_n(x) y_n . \tag{2.6}$$

3. Coefficient functionals and coordinate subspaces. We recall that a Fréchet space is defined [14, pp. 59, 110] as a metrizable, complete, locally convex topological linear space over the real or complex field. Generalizing a theorem of Banach, Newns has shown [21, pp. 431–432] that for bases in Fréchet spaces the coefficient functionals φ_n are always continuous. This can, however, be carried one step farther by discarding the hypothesis of local convexity.

Such is the content of

THEOREM 2. *Every basis in $\mathscr{M}$ is a Schauder basis.*

Proof. As observed in footnote 2, there is no loss of generality in taking the metric ρ on $\mathscr{M}$ to be translation invariant. Having done this, we can conveniently make use of the functional $\|x\| = \rho(x, 0)$.

Let $\{x_n\}$ be a basis in $\mathscr{M}$, so that for each x in $\mathscr{M}$ we have the expansion (1.1), or equivalently

$$\lim_{m\to\infty}\left\| x - \sum_{n=1}^{m} \varphi_n(x)x_n \right\| = 0 .$$

Since this yields boundedness in m (for fixed x) of

$$\left\| \sum_{n=1}^{m} \varphi_n(x)x_n \right\| ,$$

the quantity

$$\|x\|' = \sup_{m\geq 1}\left\| \sum_{n=1}^{m} \varphi_n(x)x_n \right\| \tag{3.1}$$

is always finite. Thus, $\rho'(x, y) = \|x - y\|'$ defines a translation-invariant metric ρ' on $\mathscr{M}$ with the property that $\rho(x, y) \leq \rho'(x, y$ for all x, y in $\mathscr{M}$.

It is immediate from (3.1) that

$$\|\varphi_n(x)x_n\| \leq 2\|x\|' \qquad (n = 1, 2, \cdots), \tag{3.2}$$

and the corollary to Proposition 2, pp. 25–26, of Bourbaki [14] then ensures that each φ_n is continuous in the metric ρ'. The proof will be completed by showing that ρ and ρ' define the same topology on $\mathscr{M}$.

We establish, first of all, that $\mathscr{M}$ is complete in the metric ρ'. To this end, let $\{z_k\}$ be a Cauchy sequence in the metric ρ'. From (3.2) and the result of Bourbaki just cited it follows that, for each n, $\{\varphi_n(z_k)\}_{k=1}^{\infty}$ is a Cauchy sequence of scalars and therefore converges to some scalar c_n. Now, given $\varepsilon > 0$, there exists a positive integer N such that $\|z_j - z_k\|' \leq \varepsilon$ for $j, k > N$. For arbitrary positive integers m and $m' \leq m$ we thus have

$$\left\| \sum_{n=m'}^{m} [\varphi_n(z_j) - \varphi_n(z_k)]x_n \right\| \leq 2\varepsilon \qquad (j, k > N), \tag{3.3}$$

which yields in the limit as $j \to \infty$

$$\left\| \sum_{n=m'}^{m} c_n x_n \right\| \leq 2\varepsilon + \left\| \sum_{n=m'}^{m} \varphi_n(z_k)x_n \right\| \qquad (k > N).$$

The ρ-convergence of $\sum \varphi_n(z_k)x_n$ (for fixed k) gives rise to a Cauchy condition on its partial sums and thereby on the partial sums of $\sum c_n x_n$. Hence, $\sum c_n x_n$ converges (ρ) to some point z of $\mathscr{M}$. Taking $m' = 1$ in (3.3) and passing to the limit on j, we arrive at

$$|| z - z_k ||' = \sup_{m \geq 1} \left\| \sum_{n=1}^{m} [c_n - \varphi_n(z_k)] x_n \right\| \leq 2\varepsilon \; (k > N) .$$

That is, $\{z_k\}$ converges to z in the metric ρ'.

The remainder of the proof involves simply a routine application of a corollary of the open mapping theorem [**13**, p. 41, Theorem 6] to conclude that ρ and ρ' define the same topology on $\mathscr{M}$.

Relative to a given basis $\{x_n\}$, a *coordinate subspace* of $\mathscr{M}$ is defined as a subspace of the form $\{x: \varphi_n(x) = 0 \text{ for } n \in K\}$, where K is some set of positive integers. The coordinate subspaces arising when K consists of the first $k - 1$ positive integers are of special interest in the sequel, and we denote them by $\mathscr{M}_k$. That is, for each positive integer k, $\mathscr{M}_k$ is the set of all elements of $\mathscr{M}$ expressible as infinite linear combinations of the basis elements $x_k, x_{k+1}, \cdots$. For convenience, $\mathscr{M}_k$ will be referred to as a *terminal coordinate subspace* (or, more precisely, as *the* kth *terminal coordinate subspace*) of $\mathscr{M}$ relative to $\{x_n\}$.

Since coordinate subspaces relative to Schauder bases are necessarily closed, we have

COROLLARY 2.1. *All coordinate subspaces of $\mathscr{M}$ are closed.*

4. Some variants of the Paley-Wiener theorem. A sequence $\{y_n\}$ in $\mathscr{M}$ will be called *triangular* with respect to a basis $\{x_n\}$ provided that each y_n has the representation

$$y_n = x_n + \sum_{i=n+1}^{\infty} \varphi_i(y_n) x_i . \tag{4.1}$$

In the present section we shall be concerned with the problem of determining conditions under which $\{y_n\}$ will itself be a basis in $\mathscr{M}$. This arises as a natural analogue of the Pincherle basis problem, and our methods here have much in common with those of [**9**].

We take advantage of the following special properties of triangular sequences.

LEMMA 1. *Let $\{x_n\}$ be a basis in $\mathscr{M}$, and let $\mathscr{M}_k$ be a terminal coordinate subspace of $\mathscr{M}$ relative to $\{x_n\}$. If $\{y_n\}$ is a sequence in $\mathscr{M}$ triangular with respect to $\{x_n\}$, then*

(1) *$\{y_n\}$ is linearly independent in the infinite-series sense, and*

(2) *for $\{y_n\}_{n=k}^{\infty}$ to be a basis in $\mathscr{M}_k$ it is necessary and sufficient that $\{y_n\}_{n=1}^{\infty}$ be a basis in $\mathscr{M}$.*

Proof. To show that $\{y_n\}$ is linearly independent in the infinite-series sense, we suppose that

$$\sum_{n=1}^{\infty} b_n y_n = 0 .$$

Then, from (4.1) and the fact that $\mathscr{M}_2$ is closed, it is immediate that $b_1 x_1 + z_2 = 0$, where z_2 is some point in $\mathscr{M}_2$. Hence $b_1 = 0$, and an obvious inductive argument establishes $b_n = 0$ $(n = 1, 2, \cdots)$.

The second assertion is dealt with similarly. Let $\{y_n\}_{n=k}^{\infty}$ be a basis in $\mathscr{M}_k$, and let y be any element of $\mathscr{M}$. It is evident that, for a suitably chosen scalar b_1, the point $y - b_1 y_1$ will lie in $\mathscr{M}_2$. Proceeding inductively, we then see that there exist scalars $b_1, b_2, \cdots, b_{k-1}$ yielding

$$y - \sum_{n=1}^{k-1} b_n y_n \in \mathscr{M}_k .$$

Consequently, $\{y_n\}_{n=1}^{\infty}$ spans $\mathscr{M}$ in the infinite-series sense and is therefore a basis in $\mathscr{M}$. The converse in (2) is trivial.

This leads to our first variant of Theorem 1.

THEOREM 3. *Let $\{x_n\}$ be a basis in $\mathscr{M}$ and $\{y_n\}$ a sequence triangular with respect to $\{x_n\}$. If there exist a positive number $\lambda < 1$ and a positive integer k such that*

$$\left\| \sum_{n=k}^{m} a_n (y_n - x_n) \right\| \leq \lambda \left\| \sum_{n=k}^{m} a_n x_n \right\| \tag{4.2}$$

holds for all finite sequences $a_k, a_{k+1}, \cdots, a_m$ of scalars, then

(1) *$\{y_n\}$ is a basis in $\mathscr{M}$, and*

(2) *there exists an automorphism A on $\mathscr{M}$ such that $y_n = Ax_n$ $(n = 1, 2, \cdots)$.*

Proof. For conclusion (1) we apply Theorem 1 to infer that $\{y_n\}_{n=k}^{\infty}$ is a basis in $\mathscr{M}_k$, and then invoke (2) of Lemma 1. Theorem 1 shows also that the mapping

$$A_k x = \sum_{n=k}^{\infty} \varphi_n(x) y_n$$

is an automorphism on $\mathscr{M}_k$. We can obviously extend A_k to a mapping A of $\mathscr{M}$ into itself by setting

$$Ax = \sum_{n=1}^{\infty} \varphi_n(x) y_n ,$$

and from the fact that $\{y_n\}$ is a basis in $\mathscr{M}$ it is then clear that A maps $\mathscr{M}$ onto itself in one-to-one fashion. There remains simply to observe that the continuity of A_k implies continuity of A, so that A is an automorphism on $\mathscr{M}$.

A further variant of the main theorem is at hand when $\{x_n\}$ is an *absolutely ρ-convergent basis* in $\mathscr{M}$, that is, when $\{x_n\}$ is a basis for which all $x \in \mathscr{M}$ satisfy

$$\sum_{n=1}^{\infty} \| \varphi_n(x)x_n \| < +\infty .$$[4]

THEOREM 4. *Let $\{x_n\}$ be an absolutely ρ-convergent basis in $\mathscr{M}$ and $\{y_n\}$ a sequence triangular with respect to $\{x_n\}$. If*

$$\limsup_{n\to\infty} \left\{ \sup_{a\neq 0} \frac{\sum_{i=n+1}^{\infty} \| a\varphi_i(y_n)x_i \|}{\| ax_n \|} \right\} < 1 \qquad (a, \text{ scalar}), \tag{4.3}$$

then

(1) *$\{y_n\}$ is an absolutely ρ-convergent basis in $\mathscr{M}$, and*

(2) *there exists an automorphism A on $\mathscr{M}$ such that $y_n = Ax_n$ $(n = 1, 2, \cdots)$.*

Proof. We first remetrize $\mathscr{M}$ by setting $\rho'(x, y) = \| x - y \|'$, where

$$\| x \|' = \sum_{n=1}^{\infty} \| \varphi_n(x)x_n \| \tag{4.4}$$

for all x in $\mathscr{M}$. Obviously ρ' is translation invariant, and $\| x \| \le \| x \|'$.

Condition (4.3) can now be restated as follows: there exist a positive number $\lambda < 1$ and a positive integer k such that

$$\| a(y_n - x_n) \|' \le \lambda \| ax_n \|' \tag{4.5}$$

holds for $n \ge k$ and all scalars a. This, together with (4.4), yields the inequality (4.2) in the metric ρ'. Hence, $\{y_n\}$ is a basis in $\mathscr{M}$, and there exists an automorphism A on $\mathscr{M}$ such that $y_n = Ax_n$ $(n = 1, 2, \cdots)$. It follows that, for arbitrary scalar sequences $\{b_n\}$, convergence of the series $\sum b_n y_n$ implies convergence (and thereby absolute ρ-convergence) of the series $\sum b_n x_n$. Since (4.5) results in

$$\| b_n y_n \| \le (1 + \lambda) \| b_n x_n \|$$

for $n \ge k$, we see that $\{y_n\}$ is, in fact, an absolutely ρ-convergent basis in $\mathscr{M}$. This completes the proof.

As noted in the derivation, there is no real loss of generality in requiring that

$$\left\| \sum_{n=1}^{\infty} \varphi_n(x)x_n \right\| = \sum_{n=1}^{\infty} \| \varphi_n(x)x_n \|$$

[4] In metric linear spaces the notion of absolute ρ-convergence coincides with that of absolute convergence as defined by Day [**16**, pp. 11, 59] in terms of the Minkowski functional. Here, absolutely convergent bases are defined only for Fréchet spaces (see § 5).

for all x in $\mathscr{M}$. Whenever the metric ρ and the basis $\{x_n\}$ are interrelated in this fashion, condition (4.3) assumes the simpler form

$$\limsup_{n\to\infty}\left\{\sup_{a\neq 0}\frac{\|a(y_n-x_n)\|}{\|ax_n\|}\right\}<1 \qquad (a,\ \text{scalar}). \tag{4.3'}$$

5. The case of Fréchet spaces. Proposition 6, p. 97, of [**14**] ensures that the topology on a Fréchet space $\mathscr{F}$ can be described by a sequence $\{\|\ \ \|_q\}$ of continuous semi-norms, and with no loss of generality this sequence will be taken as monotone increasing (a condition automatically fulfilled in spaces of analytic functions). Thus, $\|x\|_p \leq \|x\|_q$ for $q > p$ and all $x \in \mathscr{F}$; and convergence in $\mathscr{F}$ is equivalent to convergence with respect to each of the semi-norms $\|\ \ \|_q$. The topology on $\mathscr{F}$ is then that of the translation-invariant metric

$$\rho(x, y) = \sum_{q=1}^{\infty}\frac{1}{2^q}\frac{\|x-y\|_q}{1+\|x-y\|_q}\ . \tag{5.1}$$

As we proceed to show, the Paley-Wiener theorem and its variants can be generalized in the case of Fréchet spaces by replacing the inequalities on the metric by corresponding inequalities on the semi-norms.

THEOREM 5. *Let $\{x_n\}$ and $\{y_n\}$ be sequences in a Fréchet space $\mathscr{F}$, and let $\{\lambda_q\}$ be a sequence of real numbers $(0 < \lambda_q < 1)$ such that*

$$\left\|\sum_{n=1}^{m}a_n(y_n-x_n)\right\|_q \leq \lambda_q\left\|\sum_{n=1}^{m}a_nx_n\right\|_q \qquad (q=1,2,\cdots)$$

holds for all finite sequences $a_1, a_2, \cdots, a_m$ of scalars. Then

(1) *if $\{x_n\}$ is total in $\mathscr{F}$, so is $\{y_n\}$;*

(2) *if $\{x_n\}$ is a basis in $\mathscr{F}$, so is $\{y_n\}$, and the coefficients in any expansion $\sum b_ny_n$ satisfy*

$$\left\|\sum_{n=1}^{\infty}b_nx_n\right\|_q \leq \frac{1}{1-\lambda_q}\left\|\sum_{n=1}^{\infty}b_ny_n\right\|_q \qquad (q=1,2,\cdots);$$

(3) *if $\{x_n\}$ is a basis in $\mathscr{F}$, there exists an automorphism A on $\mathscr{F}$ such that $y_n = Ax_n$ $(n = 1, 2, \cdots)$.*

THEOREM 6. *Let $\{x_n\}$ be a basis in a Fréchet space $\mathscr{F}$ and $\{y_n\}$ a sequence triangular with respect to $\{x_n\}$. If there exist positive integers k_q and positive numbers $\lambda_q < 1$ such that*

$$\left\|\sum_{n=k_q}^{m}a_n(y_n-x_n)\right\|_q \leq \lambda_q\left\|\sum_{n=k_q}^{m}a_nx_n\right\|_q \qquad (q=1,2,\cdots) \tag{5.2}$$

holds for all finite sequences $a_{k_q}, a_{k_q+1}, \cdots, a_m$ of scalars, then

(1) *$\{y_n\}$ is a basis in $\mathscr{F}$, and*

(2) *there exists an automorphism* A *on* $\mathscr{F}$ *such that* $y_n = Ax_n$ $(n = 1, 2, \cdots)$.

The proof of Theorem 5 duplicates that of Theorem 1. The proof of Theorem 6 would likewise duplicate that of Theorem 3 if we knew that $\{k_q\}$ were bounded (so that in effect we could replace it by a single number k). Failing this, we use the following argument, based directly on the properties of the transformation T of (2.3).

Convergence of the series

$$Tx = \sum_{n=1}^{\infty} \varphi_n(x)(y_n - x_n)$$

is ensured by condition (5.2). In fact, if x lies in the k_qth terminal coordinate subspace $\mathscr{F}_{k_q}$, we have

$$\| Tx \|_q \leq \lambda_q \| x \|_q \qquad (q = 1, 2, \cdots). \tag{5.3}$$

Since the complementary subspace corresponding to each $\mathscr{F}_{k_q}$ is finite dimensional, it follows that T is continuous on $\mathscr{F}$.[5]

Now, taking account of the fact that $\mathscr{F}_k$ is closed, we verify at once that x in $\mathscr{F}_{k-1}$ implies Tx in $\mathscr{F}_k$. Hence, for arbitrary x in $\mathscr{F}$, the point $T^{k-1}x$ must lie in $\mathscr{F}_k$ $(k = 1, 2, \cdots)$. This result, combined with (5.3), leads to the inequality

$$\| T^n x \|_q \leq (\lambda_q)^{n-k_q} \| T^{k_q} x \|_q \qquad (q = 1, 2, \cdots)$$

for $n \geq k_q$ and all x in $\mathscr{F}$. As in the proof of Theorem 1, it follows that the operator series

$$U = \sum_{n=0}^{\infty} (-T)^n$$

converges and that $U = (I + T)^{-1}$. From this we conclude that $A = I + T$ is an automorphism on $\mathscr{F}$ carrying x_n into y_n $(n = 1, 2, \cdots)$, and that $\{y_n\}$ is a basis in $\mathscr{F}$.

To frame an analogue of Theorem 4, we first define an *absolutely convergent basis* in the Fréchet space $\mathscr{F}$ as a basis $\{x_n\}$ such that

$$\sum_{n=1}^{\infty} \| \varphi_n(x) x_n \|_q < +\infty \qquad (q = 1, 2, \cdots)$$

for all x in $\mathscr{F}$.[6]

---

[5] Any x in $\mathscr{F}$ can be expressed as $x = x' + x''$, where x' is the projection of x on the complementary subspace to $\mathscr{F}_{k_q}$ and x'' is the projection of x on $\mathscr{F}_{k_q}$. By continuity of the coefficient functionals, $x \to 0$ implies $x' \to 0$ and thereby $x'' \to 0$. Then $Tx' \to 0$ and $Tx'' \to 0$, so that $Tx \to 0$.

[6] It is evident from [**14**, p. 101] that this definition is independent of the choice of semi-norm sequence from among those defining the topology on $\mathscr{F}$.

THEOREM 7. *Let $\mathscr{F}$ be a Fréchet space, $\{x_n\}$ an absolutely convergent basis in $\mathscr{F}$, and $\{y_n\}$ a sequence triangular with respect to $\{x_n\}$. If*

$$\limsup_{n\to\infty} \frac{\sum_{i=n+1}^{\infty} \| \varphi_i(y_n)x_i \|_q}{\| x_n \|_q} < 1 \qquad (q = 1, 2, \cdots), \tag{5.4}$$

then

(1) *$\{y_n\}$ is an absolutely convergent basis in $\mathscr{F}$, and*

(2) *there exists an automorphism A on $\mathscr{F}$ such that $y_n = Ax_n$ $(n = 1, 2, \cdots)$.*

Proof. Setting

$$\| x \|'_q = \sum_{n=1}^{\infty} \| \varphi_n(x)x_n \|_q \qquad (q = 1, 2, \cdots) \tag{5.5}$$

for all x in $\mathscr{F}$, we observe that $\{\| \;\; \|'_q\}$ is an increasing sequence of semi-norms on $\mathscr{F}$. It thus defines a metric ρ' on $\mathscr{F}$ according to (5.1), and there is no difficulty in showing that ρ' yields the same topology as ρ.[7] Then, to each index q there correspond a positive number $\lambda_q < 1$ and a positive integer k_q such that

$$\| y_n - x_n \|'_q < \lambda_q \| x_n \|'_q$$

holds for $n > k_q$. The additivity property (5.5) assures us also that (5.2) holds for the primed semi-norms, and the proof is completed just as in the case of Theorem 4.

Again we note that the semi-norms on $\mathscr{F}$ can be required to have the additivity property

$$\left\| \sum_{n=1}^{\infty} \varphi_n(x)x_n \right\|_q = \sum_{n=1}^{\infty} \| \varphi_n(x)x_n \|_q \qquad (q = 1, 2, \cdots)$$

relative to the absolutely convergent basis $\{x_n\}$. In terms of a "natural" sequence of semi-norms of this sort, condition (5.4) reduces to

$$\limsup_{n\to\infty} \frac{\| y_n - x_n \|_q}{\| x_n \|_q} < 1 \qquad (q = 1, 2, \cdots). \tag{5.4'}$$

It is readily seen that the coefficients for an element in a given basis are finite linear combinations of the coefficients in a basis triangular with respect to the given one. We have, in fact,

LEMMA 2. *Let $\{x_n\}$ be a basis in $\mathscr{M}$ and $\{y_n\}$ a basis triangular with respect to $\{x_n\}$. If x is an element of $\mathscr{M}$ having expansions in*

[7] This argument appears also in the proof of Lemma 4 of [**11**].

the two bases as

$$x = \sum_{n=1}^{\infty} a_n x_n \qquad \textit{and} \qquad x = \sum_{n=1}^{\infty} b_n y_n \,,$$

then

$$a_1 = b_1 \qquad \textit{and} \qquad a_n = b_n + \sum_{k=1}^{n-1} b_{n-k}\varphi_n(y_{n-k}) \qquad (n \geq 2).$$

Proof. The expansion of x in the basis $\{y_n\}$ appears as

$$\begin{aligned} x = b_1[x_1 + \varphi_2(y_1)x_2 + \varphi_3(y_1)x_3 + \cdots] \\ + b_2[x_2 + \varphi_3(y_2)x_3 + \cdots] \\ + b_3[x_3 + \cdots] \\ + \cdots . \end{aligned}$$

Since $\mathscr{M}_2$ is closed, it follows from the linear independence of $\{x_n\}$ that $a_1 = b_1$. The fact that $\mathscr{M}_3$ is closed then results in $a_2 = b_2 + \varphi_2(y_1)$, and the general formula is obtained by induction. (Note that the proof in no way depends on rearrangement of the series.)

Using this lemma, we show how certain inequalities on the coefficients a_n give rise to corresponding inequalities on the coefficients b_n. The underlying space will be taken as a Fréchet space $\mathscr{F}$, and $\{y_n\}$ will again be assumed to be a basis triangular with respect to the basis $\{x_n\}$.

Thus, let x be an element of $\mathscr{F}$ having the expansion

$$x = \sum_{n=1}^{\infty} a_n x_n \,,$$

and for each q let M_q be a constant such that

$$|\, a_n \,| \leq \frac{M_q}{||\, x_n \,||_q} \qquad (n = 1, 2, \cdots).$$

Constants of this sort always exist if the basis $\{x_n\}$ is absolutely convergent, since we can, for example, put $M_q = \sum ||\, a_n x_n \,||_q$. (In spaces of analytic functions, where we have access to the Cauchy inequalities, the maximum modulus of course yields a better choice for M_q.) In similar fashion $H_q(y_n)$ will be taken as any constant for which

$$(5.6) \qquad |\, \varphi_i(y_n) \,| \leq H_q(y_n) \frac{||\, x_n \,||_q}{||\, x_i \,||_q} \qquad (i \geq n + 1).$$

Absolute convergence of $\{x_n\}$ again suffices to ensure the existence of such a constant, for example

$$H_q(y_n) = \frac{\sum_{i=n+1}^{\infty} \| \varphi_i(y_n) x_i \|_q}{\| x_n \|_q},$$

and our remark on the case of analytic function spaces carries over.

Combined with the identities on the coefficients given in Lemma 2, the above inequalities provide the estimates

$$| b_1 | \cdot \| x_1 \|_q \leq M_q ,$$
$$| b_n | \cdot \| x_n \|_q \leq M_q + \sum_{k=1}^{n-1} H_q(y_{n-k}) \, | b_{n-k} | \cdot \| x_{n-k} \|_q \qquad (n \geq 2).$$

We apply now a procedure based on the techniques (due to Narumi [20]) used in proving Theorem 5 of [10]. With $\{B_n\}$ defined inductively according to the equations

$$B_1 \| x_1 \|_q = M_q ,$$
$$B_n \| x_n \|_q = M_q + \sum_{k=1}^{n-1} B_{n-k} H_q(y_{n-k}) \| x_{n-k} \|_q \qquad (n \geq 2)$$

it is readily verified that

$$B_n \| x_n \|_q - B_{n-1} \| x_{n-1} \|_q = B_{n-1} H_q(y_{n-1}) \| x_{n-1} \|_q .$$

Thus, for $n \geq 2$

$$B_n \| x_n \|_q = [1 + H_q(y_{n-1})] B_{n-1} \| x_{n-1} \|_q ,$$

so that

$$B_1 \| x_1 \|_q = M_q ,$$
$$B_n \| x_n \|_q = M_q \prod_{k=1}^{n-1} [1 + H_q(y_k)] \qquad (n \geq 2).$$

There follows

THEOREM 8. *Let $\mathscr{F}$ be a Fréchet space, x an element of $\mathscr{F}$, $\{x_n\}$ a basis in $\mathscr{F}$, and $\{y_n\}$ a basis triangular with respect to $\{x_n\}$. Let us suppose further that there exist constants M_q such that the coefficients in the expansion*

$$x = \sum_{n=1}^{\infty} a_n x_n$$

satisfy

$$| a_n | \leq \frac{M_q}{\| x_n \|_q} \qquad (n = 1, 2, \cdots)$$

for each index q, and that there exist constants $H_q(y_n)$ for which (5.6) *holds. Then the coefficients in the expansion*

$$x = \sum_{n=1}^{\infty} b_n y_n$$

satisfy the inequalities

$$|b_1| \leq \frac{M_q}{\|x_1\|_q}, \qquad |b_n| \leq \frac{M_q}{\|x_n\|_q} \prod_{k=1}^{n-1} [1 + H_q(y_k)] \qquad (n \geq 2).$$

If in addition there exist constants J_q *such that*

$$\limsup_{n\to\infty} H_q(y_n) < J_q \qquad (q = 1, 2, \cdots), \tag{5.7}$$

then there also exist constants K_q *such that*

$$|b_n| < (1 + J_q)^n \frac{K_q}{\|x_n\|_q} \qquad (n = 1, 2, \cdots)$$

for all q, *and the constants* K_q *are independent of* q *whenever the same is true of* M_q, $H_q(y_n)$, *and* J_q. *In particular, condition* (5.4) *implies*

$$|b_n| < 2^n \frac{K_q}{\|x_n\|_q}.$$

6. Concluding remarks. We begin by making explicit the specialization of Theorem 7 to spaces of analytic functions.

Thus, let Ω be a non-empty plane region, and fix $\{\Omega_q\}$ as any sequence of non-empty subregions of Ω such that $\overline{\Omega}_q$ is a compact subset of Ω_{q+1} $(q = 1, 2, \cdots)$ and $\Omega = \bigcup \Omega_q$. The family of all functions f analytic on Ω, topologized by the sequence of semi-norms

$$M_q(f) = \max_{\overline{\Omega}q} |f|,$$

is a Fréchet space which we shall denote by $\mathscr{A}(\Omega)$.

Applied to $\mathscr{A}(\Omega)$, Theorem 7 yields the following variant of Theorem 2, p. 117, of Evgrafov [17].[8]

THEOREM 9. *Let* $\{\alpha_n\}$ *be an absolutely convergent basis in* $\mathscr{A}(\Omega)$, *and let* $\{\beta_n\}$ *be the triangular sequence defined by*

$$\beta_n(z) = \alpha_n(z) + \sum_{k=1}^{\infty} A_{nk}\alpha_{n+k}(z),$$

where the A_{nk} *are any complex numbers for which the indicated series converge. If*

[8] Evgrafov's theorem, stated in terms of total systems, is given only for Ω simply connected and all α_n bounded on Ω. Also, our condition of absolute convergence is replaced in the hypotheses of Evgrafov by the existence of a rather special sort of biorthogonal system.

$$\limsup_{n\to\infty} \sum_{k=1}^{\infty} |A_{nk}| \frac{M_q(\alpha_{n+k})}{M_q(\alpha_n)} < 1 \qquad (q = 1, 2, \cdots), \tag{6.1}$$

then $\{\beta_n\}$ is an absolutely convergent basis in $\mathscr{A}(\Omega)$, and there exists an automorphism on $\mathscr{A}(\Omega)$ carrying α_n into β_n for each n.

A further specialization results in the theorem of Boas (cited in § 1) on Pincherle bases in spaces of functions analytic on the discs $N_R(0) = \{z\colon |z| < R\}$ $(0 < R \leq +\infty)$. It is convenient here to let the index n start with 0 and to put $\delta_n(z) = z^n$ $(n = 0, 1, \cdots)$.

COROLLARY 9.1. (*Boas*). *Let*

$$\alpha_n(z) = z^n\left(1 + \sum_{k=1}^{\infty} A_{nk} z^k\right) \qquad (n = 0, 1, \cdots),$$

where the A_{nk} are complex numbers, define a sequence in $\mathscr{A}(N_R(0))$. If

$$\limsup_{n\to\infty} \sum_{k=1}^{\infty} |A_{nk}| r^k < 1 \tag{6.2}$$

for each $r < R$, then $\{\alpha_n\}$ is an absolutely convergent basis in $\mathscr{A}(N_R(0))$, and there exists an automorphism A on $\mathscr{A}(N_R(0))$ such that $\alpha_n = A\delta_n$ $(n = 0, 1, \cdots)$.[9]

Returning to the case of general Fréchet spaces, we observe that the results of § 5 remain valid if we assume only that the conditions on the semi-norms are satisfied for infinitely many indices q. In fact, the topology on $\mathscr{F}$ obviously is not affected if we replace the initial sequence of semi-norms by any subsequence of it.

Finally, we remark that when the underlying space is a Banach space, Theorems 4 and 7 coalesce. The common theorem is, however, somewhat restricted in scope, since every Banach space admitting an absolutely convergent basis is isomorphic to the space l^1 of absolutely summable sequences (see Karlin [**18**, p. 974]).

BIBLIOGRAPHY

A. Chronological listing of articles on the Paley-Wiener theorem

1. R. E. A. C. Paley and N. Wiener, *Fourier transforms in the complex domain*, New York, 1934.
2. R. P. Boas, Jr., *Expansions of analytic functions*, Trans. Amer. Math. Soc., **48** (1940), 467–487.
3. Y. Y. Tseng, *On generalized biorthogonal expansions in metric and unitary spaces*, Proc. Nat. Acad. Sci. U.S.A., **28** (1942), 35–43.

[9] The theorem of Boas is enunciated in essentially this form by Evgrafov [**17**, p. 122] Theorem 2 of [**9**] states the result in a slightly restricted way by taking the limit as $r \to R$ of the left-hand member of (6.2). The proof in [**9**] is, however, valid in the present case.

4. H. Pollard, *Completeness theorems of Paley-Wiener type*, Ann. of Math. (2), **45** (1944), 738–739.

5. B. Sz. Nagy, *Expansion theorems of Paley-Wiener type*, Duke Math. J., **14** (1947), 975–978.

6. S. H. Hilding, *Note on completeness theorems of Paley-Wiener type*, Ann. of Math. (2), **49** (1948), 953–955.

7. F. W. Schafke, *Über einige unendliche lineare Gleichungssysteme*, Math. Nachr., **3** (1949), 40–58.

8. ———, *Das Kriterium von Paley und Wiener im Banachschen Raum*, Math. Nachr., **3** (1949), 59–61.

B. Other references

9. M. G. Arsove, *The Pincherle basis problem and a theorem of Boas*, Math. Scand., **5** (1957), 271–275.

10. ———, *Proper Pincherle bases in the space of entire functions*, Quart. J. Math. (Oxford) (2), **9** (1958), 40–54.

11. ———, *Similar bases and isomorphisms in Fréchet spaces*, Math. Ann., **135** (1958), 283–293.

12. M. G. Arsove and R. E. Edwards, *Generalized bases in topological linear spaces*, Studia Math. (to appear).

13. S. Banach, *Théorie des opérations linéaires*, Varsovie, 1932.

14. N. Bourbaki, *Espaces vectoriels topologiques* (Ch. I, II), Paris, 1953.

15. R. C. Buck, *Expansion theorems for analytic functions*, Conference on functions of a complex variable, University of Michigan (1953), 409–419.

16. M. M. Day, *Normed linear spaces*, Berlin, 1958.

17. M. A. Evgrafov, *The method of approximating systems in the space of analytic functions and its applications to interpolation*, Trudy Moskovskovo Mat. Obshchestva, **5** (1956), 89–201. (Russian).

18. S. Karlin, *Bases in a Banach space*, Duke Math. J., **15** (1948), 971–985.

19. D. Montgomery and L. Zippin, *Topological transformation groups*, New York, 1955.

20. S. Narumi, *A theorem on the expansion of analytic functions by infinite series*, Tôhoku Math. J., **30** (1928), 441–444.

21. W. F. Newns, *On the representation of analytic functions by infinite series*. Phil. Trans. Royal Soc. London (A), **245** (1953), 429–468.

INSTITUT HENRI POINCARÉ
AND
UNIVERSITY OF WASHINGTON

ACCEPTABLE POINTS IN GAMES OF PERFECT INFORMATION

ROBERT J. AUMANN

Summary. This is the second of a series of papers on the theory of acceptable points in n-person games. The first was [1]; in it the notion of acceptable points was defined for cooperative games, and a fundamental theorem was proved relating the acceptable expected payoffs for a single play of a game to probable average payoffs for "strong equilibrium points" in its supergame.[1]

The chief result of the current paper, Theorem 5.4, is a generalization of von Neumann's classical Theorem on two-person zero-sum games of perfect information (see [11]). Roughly, it states that strong equilibrium points in the supergame of a stable game of perfect information can be achieved in pure supergame strategies. An example shows that not all games possess this property; and in fact, it is conjectured that the property is characteristic of game structures of perfect information.

The theorem stated above holds whether G is interpreted as a cooperative or as a non-cooperative game. To lend meaning to this statement, we will have to extend the theory introduced in [1] to non-cooperative games. We plan to do this in full in a subsequent paper. Here just enough definitions and theorems will be used to enable us to state and prove the chief result for non-cooperative games of perfect information.

The paper is divided into two parts, the first centering around the proof of the chief result for cooperative games, the second dealing with the extension to non-cooperative games. Section 1, the introduction, serves mainly to supply background from [1] and from the literature. In § 2, we show that the naive approach to generalizing von Neumann's theorem on games of perfect information fails; that is, we bring an example of a stable game of perfect information that has no acceptable point in pure strategies. It is then shown intuitively that an appropriate generalization of the von Neumann Theorem should involve the supergame. Sections 3 and 4 are devoted to the proof of preliminary theorems, dealing with supergame pure strategies and supergames of perfect information, respectively. In § 5 we establish the chief result. Section 6, which completes the first part of the paper, is devoted to the example and conjecture mentioned in connection with the chief result.

Received September 4, 1958.

[1] Readers not familiar with [1] should read the introduction (section 1) before continuing with this summary.

The second part begins with § 7, a summary of the additional notation needed for the non-cooperative case. In § 8 the concept of acceptability is defined for non-cooperative games. In § 9, we show that in a game G of perfect information, it makes no difference, insofar as the theory of acceptable points is concerned, whether G is to be considered as a non-cooperative or as a cooperative game. More precisely, it is shown that the set of acceptable payoffs in the non-cooperative sense, coincides with the set of acceptable payoffs in the cooperative sense. This is a consequence of the lemma, interesting in its own right, that in a game of perfect information, the set of payoff vectors to correlated strategy vectors coincides with the set of payoff vectors to mixed strategy vectors. Again, this lemma seems to be characteristic of game structures of perfect information. In § 10, we define supergame strategies for non-cooperative games and prove some preliminary results. Section 11 is devoted to the statement and proof of the chief result for non-cooperative games of perfect information.

As in [1], the games under consideration contain no chance moves.

We will make unrestricted use of the notations, ideas, definitions, theorems and proofs of [1]. We will not in general repeat explanations and proofs that are similar to those given there. Especially heavy use will be made of § 6 of [1].

1. Introduction and background. Up to the present, the starting point for all work on games of perfect information has been the theorem of von Neumann that every two-person zero-sum game of perfect information with finitely many moves has a solution in pure strategies. Subsequent work has dealt with extensions to n-person games and the concomitant generalizations of the solution notion, with various converses to the von Neumann theorem, with extensions to games containing infinitely many moves (i. e., positions), and with various combinations of these. We mention also the notion of stochastic games of perfect information with infinitely many moves.

In the first of these areas, Kuhn [9] showed that the von Neumann theorem could be extended to n-person games if the "equilibrium point" notion of Nash [12] was substituted for the classical solution notion. Dalkey [4] proved a converse of this theorem, which reduces to a converse of the von Neumann theorem in the two-person, zero-sum case. Gale and Stewart [6] were the first to treat games of perfect information with infinitely many moves; they showed that certain such (two-person zero-sum) games possess no pure strategy solutions, and derived sufficient conditions for the existence of a pure strategy solution. Wolfe [14] extended their results. By adopting a definition of payoff that is somewhat more restricted than that of Gale and Stewart, Berge [2] was able to extend von Neumann's theorem to some games with infinitely many moves. He was also able to show [2, 3] that under very general conditions on the structure of the game, Kuhn's theorem on the existence of a pure strategy equilibrium point in a game of perfect information holds true. The work of Shapley [13] and Gillette [7] on Stochastic Games of perfect information will be discussed in detail below.

The current paper deals with an extension of the von Neumann theorem to n-person games. The solution notion that we use is that of "acceptable" points, introduced in [1]. The notion of acceptability is a generalization of the "core" introduced by Gillies [8] for the cooperative game with side payments. More precisely, an n-tuple x of strategies is called *acceptable* if the players of any given coalition can be prevented by the players not in that coalition from each obtaining a higher payoff than when x is played (Definition 4.1 of [1]). Intuitively, it would seem that in a long sequence of plays of a game, a "steady state" would have to represent an acceptable point, because the players would certainly tend to move away from any point that is *not* acceptable.

In order to obtain a precise statement and proof of this intuitive idea, we introduced (in § 6 of [1]) the formal notion of the "supergame" of a given game G. The *supergame* of G is a game each play of which consists of an infinite sequence of plays of G. The payoff to a superplay (i.e., a play of the supergame) is given by the average (i.e., first césaro limit, if it exists) of the payoffs to the individual plays of G that constitute the superplay. Many of the notions that apply to ordinary games can also be applied to supergames. In particular, it is possible to define the notion of strategy in the supergame, and also the notion of a strategy equilibrium point in the sense of Nash. A much stronger form of the Nash equilibrium notion may be defined as follows: An n-tuple x of strategies is called a "strong equilibrium point" if for no coalition B can *all* the members of B increase their payoff by adopting strategies different from those at x while the remaining players (those in $N - B$) play as they did at x. The notion of strong equilibrium applied to the supergame provides a formalization of the "steady state" idea (§ 7 of [1]).

The basic result of [1] (§ 10) may be stated as follows: The payoffs for the acceptable points in a game G are the same as the payoffs for the strong equilibrium points in its supergame. Since the notion of acceptability depends only on the payoff, this means that the acceptable points in G correspond precisely to the steady state points in the supergame of G. For two-person zero-sum games, a point is acceptable if and only if its payoff is the game value, whereas it is a strong equilibrium point if and only if it is a solution (§ 5 of [1]).

The object of this paper is to apply the theory of acceptable points to games of perfect information, with a view to obtaining an appropriate

n-person generalization of the von Neumann theorem. In other words, we want to accomplish for acceptable points in games of perfect information what Kuhn did in [9] for equilibrium points in games of perfect information. The first conjecture in this direction might be that every game of perfect information has an acceptable point in pure strategies. This is unreasonable, because according to an example given in [1] (§ 11), not every game of perfect information need have an acceptable point at all, let alone one in pure strategies. However, it turns out that not even all *stable* games (games that do have acceptable points) of perfect information have pure strategy acceptable points. The reasons for this are discussed in § 2, and it is also shown there that a more appropriate place to look for a generalization of the von Neumann theorem is in the supergame. We would like to show that if G is a game of perfect information, then each player can restrict himself to pure strategies in each play of an infinite sequence of plays of G. In fact, we prove (Theorem 5.4) that every acceptable point (and hence every strong equilibrium point) in a game of perfect information can be "achieved" in pure supergame strategies, in the sense that there is a pure strategy strong equilibrium point with the same payoff. In particular, if the supergame of a game of perfect information has a strong equilibrium point at all, then it already has one in pure strategies.

Formally, the supergame defined in [1] bears some resemblance to the stochastic games treated by Gillette in [7]. The two concepts are similar in that both involve games consisting of an infinite sequence of plays of finite games, and the payoffs in both cases are given by a form of the average of the payoffs to the individual plays. The main differences are that Gillette considers a set of M games, any one of which may be the game played at a given stage, whereas we are concerned with repeated plays of one game only; and that Gillette considers two-person zero-sum games, while we deal with n-person games. The "intersection" of the two theories is an infinite sequence of plays of the same two-person zero-sum game of perfect information, a trivial situation once von Neumann's theorem is known (obviously both players play their optimal pure strategies on each play). The two theories provide totally "disjoint" generalizations of the von Neumann theorem.

All of Gillette's positive results involve "stationary" strategies, i.e., supergame strategies that are obtained by repeating the same strategy on each play of the infinite sequence of plays that constitutes a superplay. In a somewhat similar situation, Everett [5] gives a formal definition of some strategies that are not stationary, and obtains positive results with them; but the strategies he defines are still "almost" stationary in the sense that the choice of a player at a given game of the supergame can depend only on which game he is at, not on the choices of the other players on previous occasions.

It is of interest to ask whether these restricted notions of strategy are sufficient for our theory. The answer is no. The existence of a strong equilibrium point in stationary pure strategies would imply the existence of an acceptable point in pure strategies; and the example in § 2 shows that even in stable games of perfect information such an acceptable point in pure strategies need not always exist. The same example shows that there is no strong equilibrium point in "almost" stationary pure strategies.

Finally, we mention that in the supergames of games of perfect information (even unstable ones), there is always a Nash equilibrium point (as opposed to a strong equilibrium point) in stationary pure strategies. This is a consequence of Kuhn's theorem.

2. Failure of the naive approach. We saw in [1] (§ 5) that the concept of acceptability constitutes a generalization of the concept of solution in two-person, zero-sum games. As a generalization of Von Neumann's Theorem on two-person zero-sum games of perfect information, we might hope that *every* game of perfect information that has any acceptable points also has acceptable points in pure strategies. An example shows that this is false.

The game G is a two-person, non-zero-sum game of perfect information. In the game tree, given in Figure 1, the moves are labelled with the names of the players and the terminals with the payoff vectors.

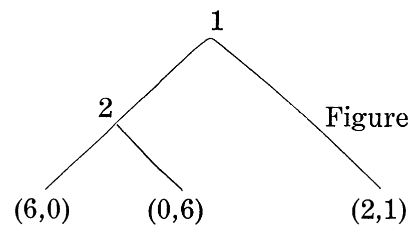

Figure 1

Each player has two strategies, the *left* and the *right* strategies. Notation in the following payoff matrix is obvious.

	L^1	R^1
L^2	(6,0)	(2,1)
R^2	(0,6)	(2,1)

Player 1 cannot be prevented from obtaining at least 2 (he can play R^1); player 2 cannot be prevented from obtaining at least 1 (he can play R^2). This shows that (L^1, L^2) and (L^1, R^2) are not acceptable. The other two pure strategy pairs are not acceptable because the coalition (1,2) cannot be prevented from obtaining (3,3)—by playing $(L^1, 1/2L^2 + 1/2R^2)$—and (3,3) is strictly larger than the payoff vector at both (R^1, L^2) and (R^1, R^2).

Hence G has no acceptable point in pure strategies. (Note that (3,3) is an acceptable payoff vector, so that G does have some acceptable points.)

The intuitive feeling that a game of perfect information should have a "good" point in pure strategies can be traced to the traditional purpose of mixed strategies—namely, to hide one's intensions from one's opponent by the use of a random device. In a game of perfect information, we somehow feel that it is unnessary to hide one's intensions, that it is in the nature of the game that everything may just as well be open and above-board. The conclusion is that mixed strategies are unnecessary in such a game, and that therefore we may just as well confine ourselves to the consideration of pure strategies.

The counter example points up the fallacy in this intuitive argument. It is quite true that the hiding of one's intentions, and the concommitant use of a random device should be unnecessary in a game of perfect information. This does not mean, though, that one can achieve one's desires by means of pure strategies. Indeed, if there were some means of mixing one's strategies other than by the use of a random device, this would be perfectly satisfactory in Example 3. For example, the pair $(L^1, 1/2L^2 + 1/2R^2)$ happens to be acceptable. If, instead of tossing a coin before each play of a sequence of plays, 2 were to announce beforehand that he will alternate L^2 and R^2, this would in no way affect the actions of 1. Contrary to the situation in, say, penny matching, the purpose of playing $1/2L^1 + 1/2R^1$ here is simply to achieve a payoff not provided in the matrix, not to avoid "discovery" by the opponent.

This discussion shows that though we cannot expect pure-strategy acceptable points in a game G of perfect information, we should be able to expect that the players may, without loss, restrict themselves to pure strategies in each of the plays that constitute a superplay of G. This is in fact the case, as we shall see in the sequel.

3. **Supergame pure strategies.** A supergame pure strategy vector (or p-strategy vector) is a c-strategy vector in which there are no coalitions and the players choose pure strategies on each play. Here the second condition is the essential one; the first condition is adopted only for convenience. If the first condition were eliminated, the resulting supergame strategy vectors would be essentially equivalent to those obtained under our definition.

The formal definition is as follows:

DEFINITION 3.1. *A supergame c-strategy f^i is said to be "pure" if*

$$f^i_k(y) \in P^i$$

for each $k \geq 0$ and $y \in J^i_1 \times \cdots \times J^i_k$.

We also say that f^i is a supergame p-strategy.

The following are lemmas that will be needed later.

LEMMA 3.2. *If f is a supergame p-strategy vector and B is a (possibly empty) subset of N, then for each $k \geq 0$ and $y \in J_1 \times \cdots \times J_k$, we have*

$$f_k^{N-B}(y) \mid R^{N-B} = d_e^{N-B} .$$

Furthermore, for each $k \geq 1$, we have

$$z_k(f) = (x_1(f), \cdots, x_k(f)) .$$

Proof. The first statement follows at once from 3.1. The second statement follows by induction from 6.3, 6.4, and 6.5 of [1], and from 3.1.

LEMMA 3.3. *Let f be a supergame p-strategy vector, and let g be a supergame c-strategy vector for which*

$$g^{N-B} = f^{N-B} .$$

Let $v = (v_1, \cdots, v_k, \cdots) \in J_1 \times \cdots \times J_k \times \cdots$ occur with positive probability when g is played (see definition 10.22 of [1]). *Then for all $k \geq 1$, we have*

$$v_k \mid R^{N-B} = d_e^{N-B}$$

and for all $k \geq 0$, we have

$$f_k^{N-B}(v_1, \cdots, v_k) = f_k^{N-B}((v_1 \mid U_1, d_e), \cdots, (v_k \mid U_k, d_e)) .$$

Proof. The first statement is an immediate consequence of the previous lemma. As for the second statement, it follows from the first statement that

$$\begin{aligned} f_k^{N-B}(v_1, \cdots, v_k) &= f_k^{N-B}((v_1^B, v_1^{N-B}), \cdots, (v_k^B, v_k^{N-B})) \\ &= f_k^{N-B}((v_1^B, (v_1^{N-B} \mid U_1^{N-B}, d_e^{N-B})), \cdots, (v_k^B, (v_k^{N-B} \mid U_k^{N-B}, d_e^{N-B}))) . \end{aligned}$$

But by Definition 6.1 of [1], f_k^{N-B} is independent of $(v_1^B, \cdots, v_k^B)$. The result follows at once.

For a supergame c-strategy vector f, define $S_k(f)$ $(= S_k)$ by

$$S_k(f) = \frac{1}{k} \sum_{j=1}^{k} H_j(f) . \tag{3.4}$$

Parallel to the definition of strong equilibrium c-point, we may make the following definition:

DEFINITION 3.5. *A strong equilibrium p-point f is a summable supergame p-strategy vector for which there is no $B \subset N$ for which there is a supergame p-strategy vector g satisfying*

$$g^{N-B} = f^{N-B} \tag{3.6}$$

and

$$\limsup_{k\to\infty} \min_{i\in B} (S_k^i(g) - H^i(f)) > 0 \ . \tag{3.7}$$

The set of all strong equilibrium p-points is denoted by S_p. The Condition 3.7 may also be replaced by the following condition:

$$\liminf_{k\to\infty} (S_k^B(g) - H^B(f)) > 0 \ . \tag{3.8}$$

We denote by $\tilde{S}_p$ the set of supergame p-strategy vectors that satisfy a condition that differs from 3.5 only in that 3.7 is replaced by 3.8.

The essential difference between a strong equilibrium p-point and a pure strong equilibrium c-point is that in the former, $N - B$ need only be prepared to defend against all supergame *pure* strategy B-vectors, whereas in the latter, $N - B$ must be prepared to defend against *all* supergame correlated strategy B-vectors. We will show in 3.11 that the two conditions are nevertheless equivalent. As for 3.7 and 3.8, they are merely translations of 7.2 and 7.3 of [1] to the case of pure strategies, where the consideration of probabilities becomes superfluous.

THEOREM 3.9. *If f is a supergame p-strategy vector, then $z_k(f)$, $x_k(f)$, and $E_k(f)$ are "pure" for each $k \geq 0$; that is, they are discrete probability distributions in which one of the events occurs with probability* 1, *all others with probability* 0.

Proof. This is a trivial consequence of (6.2), (6.3), (6.4), (6.5) and (6.6) of [1], and of 3.1.

Theorem 3.9 enables us to replace probability statements involving the random variable $S_k(v)$ by statements involving the constants $S_k(f)$ only. More precisely, we have

COROLLARY 3.10. *Let $F(x_1, x_2, \cdots)$ be a predicate depending on a sequence of B-vectors $x_1, x_2, \cdots$. Let Λ be the proposition a function that assigns the number* 1 *to true propositions and the number* 0 *to false propositions. Suppose f is a supergame p-strategy vector for the game G. Then*

$$\text{Prob}_f\, F(S_1^B(v), S_2^B(v), \cdots) = \Lambda(F(S_1^B(f), S_2^B(f) \cdots))$$

Similar results hold for $z_k(f)$, $x_k(f)$ and $E_k(f)$.

THEOREM 3.11. *Every strong equilibrium p-point is a strong equilibrium c-point. Conversely, every pure strong equilibrium c-point is a strong equilibrium p-point. In symbols*

$$F_p \cap S_c = S_p ,$$

where F_p is the set of supergame p-strategy vectors.

Proof. We consider first the converse, the easier of the two statements. Let f be a pure strong equilibrium c-point. It is sufficient to prove that there is no pure g satisfying 3.6 and 3.7. Suppose there is such a g. Then g must satisfy 7.1 of [1], which is identical with 3.6. Furthermore, from 3.7 we deduce the existence of an $\varepsilon > 0$ such that for infinitely many k, we have

$$\min_{i \in B} (S_k^i(g) - H^i(f)) > \varepsilon .$$

It follows that for infinitely many k, we have

$$S_k^B(g) > H^B(f) + \varepsilon^B ,$$

where ε^B is a B-vector defined by

$$\varepsilon^i = \varepsilon$$

for all $i \in B$. Hence it follows that for all k, we have

$$S_r^B(g) > H^B(f) + \varepsilon^B \text{ for some } r \geq k .$$

Applying 3.10, we obtain

$$\text{Prob}_g (S_r^B(v) \geq H^B(f) + \varepsilon^B \text{ for some } r \geq k) = 1$$

for all k. Hence it follows that

$$\lim_{k \to \infty} \text{prob}_g (S_r^B(v) \geq H^B(f) + \varepsilon^B \text{ for some } r \geq k) = 1 > 0 .$$

But this is exactly Condition 7.2 of [1]. We have established that g satisfies 7.1 and 7.2 of [1], whence f cannot be a strong equilibrium c-point. This contradicts the hypothesis, and we must conclude that g satisfies 3.6 and 3.7. This completes the proof of the converse.

Now assume that f is a strong equilibrium p-point, but not a strong equilibrium c-point. Then there is a supergame c-strategy vector g satisfying 7.1 and 7.2 of [1]. From 7.1 of [1] we obtain

$$g^{N-B} = f^{N-B} . \tag{1}$$

From 7.2 of [1], we obtain that there is a B-vector $\varepsilon^B > 0$ for which

$$\lim_{k \to \infty} \text{Prob}_g (S_r^B(v) \geq H^B(f) + \varepsilon^B \text{ for some } r \geq k) > 0 . \tag{2}$$

Now the expression inside the limit on the left side of (2) is monotone decreasing with k; hence (2) implies the existence of a

$$(3) \qquad \delta > 0$$

such that

$$(4) \quad \mathrm{Prob}_g\,(S_r^B(v) \geq H^B(f) + \varepsilon^B \text{ for some } r \geq k) > \delta, \text{ for all } k \geq 1 .$$

From (4) we obtain

$$(5) \quad \mathrm{Prob}_g\,(\text{For all } k \geq 1,\, S_r^B(v) \geq H^B(f) + \varepsilon^B \text{ for some } r \geq k) \geq \delta ,$$

which is the same as

$$(6) \qquad \mathrm{Prob}_g\,(S_r^B(v) \geq H^B(f) + \varepsilon^B \text{ for infinitely many } r) \geq \delta .$$

That (5) follows from (4) is an immediate consequence of the fact that the measure of the intersection of a monotone decreasing sequence of measurable sets is the limit (or g.l.b.) of the measures of the sets.

From (3) and (6) it follows that there is a sequence

$$v = (v_1, \cdots, v_k, \cdots) \in J_1 \times \cdots \times J_k \times \cdots ,$$

occurring with positive probability when g is played, for which

$$(7) \qquad S_r^B(v) \geq H^B(f) + \varepsilon^B \text{ for infinitely many } r .$$

Since v occurs with positive probability we deduce from 6.4 and 6.2 of [1], and from the definitions in § 2 of [1] that for each k,

$$\begin{aligned} 0 < s(g_{k-1}(v_1, \cdots, v_{k-1}))(v_k) &\leq u(c(g_{k-1}(v_1, \cdots, v_{k-1})))(v_k \mid U_k) \\ &= \sum_{p_k \in u^{-1}(v_k \mid U_k)} c(g_{k-1}(v_1, \cdots, v_{k-1}))(p_k) . \end{aligned}$$

It follows that for each k, there is a p_k satisfying

$$(8) \qquad v_k \mid U_k = u(p_k)$$

and

$$(9) \qquad c(g_{k-1}(v_1, \cdots, v_{k-1}))(p_k) > 0 .$$

Now as a consequence of 2.7 of [1], Lemmas 3.2 and 3.3, and (1), we have that for each $k \geq 0$,

$$(10) \qquad c^{N-B}(g_k(v_1, \cdots, v_k)) = g_k^{N-B}(v_1, \cdots, v_k) .$$

From (9) it follows that

$$c^{N-B}(g_k(v_1, \cdots, v_k))(p_{k+1}^{N-B}) > 0 .$$

Applying (10), we deduce that

$$g_k^{N-B}(v_1, \cdots, v_k)(p_{k+1}^{N-B}) > 0 ,$$

and it then follows from (1) that

$$f_k^{N-B}(v_1, \cdots, v_k)(p_{k+1}^{N-B}) > 0 . \tag{11}$$

Since f_k^{N-B} must be a pure strategy $(N-B)$-vector, it follows from (11) that

$$f_k^{N-B}(v_1, \cdots, v_k) = p_{k+1}^{N-B} . \tag{12}$$

We now define a supergame p-strategy vector q by

$$q^{N-B} = f^{N-B} \tag{13}$$

$$q_{k-1}^i = p_k^i , \quad i \in B, k \geq 1 . \tag{14}$$

Next, we prove that for $k \geq 1$,

$$z_k(q) = ((v_1 \mid U_1, d_e), \cdots, (v_k \mid U_k, d_e)) . \tag{15}$$

That

$$z_k(q) \mid R_1 \times \cdots \times R_k = (d_e, \cdots, d_e) \tag{16}$$

follows at once from (13), (14) and the fact that f is a supergame p-strategy vector. The remainder of (15) is proved by induction on k. For $k = 1$, we have by 6.2 and 6.3 of [1],

$$\begin{aligned} z_1(q) \mid U_1 &= u(c(q_0)) \\ &= u(c(p_1^B, f_0^{N-B})) \text{ (by (13) and (14))} \\ &= u(c(p_1)) \text{ (by (12))} \\ &= u(p_1) \text{ (by 2.7 of [1])} \\ &= v_1 \mid U_1 . \end{aligned} \tag{17}$$

Now let us assume that we have established

$$z_k(q) \mid U_1 \times \cdots \times U_k = (v_1 \mid U_1, \cdots, v_k \mid U_k) . \tag{18}$$

Then by 6.2 and 6.4 of [1],

$$z_{k+1}(q) \mid U_1 \times \cdots U_{k+1} = \sum_{y \in J_1 \times \cdots \times J_k}^{*} z_k(q)(y)(y \mid U_1 \times \cdots \times U_k, u(c(q_k(y)))) .$$

By (16) and (18), all the coefficients $z_k(q)(y)$ in this sum vanish, unless

$$y = ((v_1 \mid U_1, d_e), \cdots, (v_k \mid U_k, d_e)) .$$

Hence

$$z_{k+1}(q) \mid U_1 \times \cdots \times U_{k+1} = (v_1 \mid U_1, \cdots, v_k \mid U_k, u(c(q_k(z_k(q))))) . \tag{19}$$

Now by (14),

$$q_k^B(z_k(q)) = p_{k+1}^B \tag{20}$$

and by (10),

$$\begin{aligned} q_k^{N-B}(z_k(q)) &= f_k^{N-B}(z_k(q)) \\ &= f_k^{N-B}((v_1 \mid U_1, d_e), \cdots, (v_k \mid U_k, d_e)) \text{ (by (16) and (18))} \\ &= f_k^{N-B}(v_1, \cdots, v_k) \text{ by (Lemma 3.3)} \\ &= p_{k+1}^{N-B} \text{ (by (12))} . \end{aligned} \tag{21}$$

Combining (20) and (21), we obtain $q_k(z_k(q)) = p_{k+1}$.
Hence

$$c(q_k(z_k(q))) = p_{k+1} ,$$

and it follows from this and (8) that

$$u(c(q_k(z_k(q)))) = u(p_{k+1}) = v_{k+1} \mid U_{k+1} . \tag{22}$$

Combining (19) and (22), we obtain

$$z_{k+1}(q) \mid U_1 \times \cdots \times U_{k+1} = (v_1 \mid U_1, \cdots, v_{k+1} \mid U_{k+1}) ,$$

which completes the inductive step and the proof of (15). Hence (22) holds for all k, and therefore

$$\begin{aligned} H_{k+1}(q) &= H(c(q_k(z_k(q)))) \text{ (by 6.6 and 6.7 of [1])} \\ &= (\psi \circ u)(c(q_k(z_k(q)))) \text{ (by § 6 of [1])} \\ &= \psi(u(c(q_k(z_k(q))))) \\ &= \psi(v_{k+1} \mid U_{k+1}) \text{ (by (22))} \\ &= H_{k+1}(v) \text{ (by 6.10 of [1])} . \end{aligned}$$

It then follows from 6.11 of [1] and from 3.4 that

$$S_k(q) = S_k(v) .$$

Applying (7), we obtain that

$$S_k^B(q) \geq H^B(f) + \varepsilon^B \text{ for infinitely many } k.$$

In particular,

$$\min_{i \in B} (S_k^i(q) - H^i(f)) \geq \min_{i \in B} \varepsilon^i$$

for infinitely many k, and it follows that

$$\limsup_{k \to \infty} \min_{i \in B} (S_k^i(q) - H^i(f)) \geq \min_{i \in B} \varepsilon^i > 0 . \tag{23}$$

Now by (13) and (14), q is a supergame p-strategy vector. By (1) it satisfies 3.6 and by (23) it satisfies 3.7. Hence by 3.5, f cannot be a strong equilibrium p-point, a contradiction. This completes the proof of 3.11.

THEOREM 3.12. $F_p \cap \tilde{S}_c = \tilde{S}_p$.

Proof. The proof is similar to that of 3.11. It will be omitted.

The following formulae follow easily from the indicated definitions and theorems.

(3.13) $$H(S_p) \subset H(S_c) \text{ (by 3.11)} .$$

(3.14) $$H(\tilde{S}_p) \subset H(\tilde{S}_c) \text{ (by 3.12)} .$$

(3.15) $$S_p \subset \tilde{S}_p \text{ (by 3.5)} .$$

(3.16) $$H(S_p) \subset H(\tilde{S}_p) \text{ (by 3.15)} .$$

Finally, we mention the following theorem, which will be needed in the sequel.

THEOREM 3.17. *A supergame p-strategy vector f is summable if and only if it is summable in the mean.*

Proof. The necessity follows at once from 6.9 of [**1**]. For sufficiency, we must show that if f is summable in the mean, then a sequence of random variables distributed according to $E_k(f)$ obeys the strong law of large numbers. But this follows at once from 3.9.

4. Supergame pure strategies in games of perfect information. In a game G of perfect information, the information that a player i has about the outcome of each previous play[2] may be described as follows,

(4.1) He knows which terminal was reached.

(4.2) He knows which pure strategy he himself played.

Formally, let W be the set of terminals in G, and let

$$\lambda: P \rightarrow W$$

be the function that associates with each pure strategy vector p the terminal $\lambda(p)$ that results when p is played (in the notation of [**9**], if $\pi \in P$, $\lambda(\pi)$ is the unique $w \in W$ for which $p_\pi(w) = 1$). Then for each $i \in B$ and $p \in P$,

(4.3) $$u^i(p) = (\lambda(p), p^i) .$$

If he wishes, the reader may regard 4.3 as the definition of u^i for games of perfect information.

Actually, each player may with impunity discard the additional information obtained from 4.2 as long as he restricts himself to the use of supergame p-strategies. Formally, we may say that in a game of perfect information, each supergame p-strategy f^i is equivalent to one

[2] We are discussing that information that is characterized by the information function u^i.

in which f^i_k depends only on the $\lambda(p)$, not on the p^i. To lend meaning to this statement, we must give a suitable definition of equivalence.

DEFINITION 4.4. *Two supergame p-strategies f^i and g^i are said to be equivalent ($f^i \simeq g^i$) if for each supergame p-strategy $(N-i)$-vector θ^{N-i}, we have*

$$H_k(f^i, \theta^{N-i}) = H_k(g^i, \theta^{N-i}) , \qquad k \geq 1 .$$

COROLLARY 4.5. *Let $B \subset N$. If two supergame p-strategy vectors f and g are equivalent, then for each supergame p-strategy $(N-B)$-vector θ^{N-B}, we have*

$$H_k(f^B, \theta^{N-B}) = H_k(g^B, \theta^{N-B}) , \qquad k \geq 1 .$$

Proof. Let

$$B = \{i_1, \cdots, i_b\} .$$

Then since $f^i \simeq g^i$ for each i it follows that for each $k \geq 1$,

$$\begin{aligned} H_k(f^B, \theta^{N-B}) &= H_k(g^{i_1}, (f^{B-i_1}, \theta^{N-B})) \\ &= H_k(g^{i_1}, g^{i_2}, (f^{B-i_1-i_2}, \theta^{N-B})) \\ &= H_k((g^{i_1}, \cdots, g^{i_b}), \theta^{N-B}) \\ &= H_k(g^B, \theta^{N-B}) . \end{aligned}$$

This completes the proof.

DEFINITION 4.6. *Let G be a game of perfect information. A supergame p^*-strategy f^i is a supergame p-strategy for which for each $k \geq 1$, and pair $(p_1, \cdots, p_k)$ and $(q_1, \cdots, q_k)$ of sequences of pure strategy vectors, we have*

$$\begin{aligned} &\lambda(p_j) = \lambda(q_j) , \qquad 1 \leq j \leq k , \\ \Rightarrow \quad &f^i_k(((\lambda(p_1), p^i_1), d_e), \cdots, ((\lambda(p_k), p^i_k), d_e)) \\ &= f^i_k(((\lambda(q_1), q^i_1), d_e), \cdots, ((\lambda(q_k), q^i_k), d_e)) . \end{aligned}$$

For convenience, we will sometimes make use of the following conventions:

CONVENTION 4.7. When f^i is a supergame p^*-strategy, write

$$f^i_k(\lambda(p_1), \cdots, \lambda(p_k))$$

instead of

$$f^i_k(((\lambda(p_1), p^i_1), d_e), \cdots, ((\lambda(p_k), p^i_k), d_e)) .$$

CONVENTION 4.8. When f^i is a supergame p-strategy, write

$$f^i_k(u^i(p_1), \cdots, u^i(p_k))$$

instead of

$$f^i_k((u^i(p_1), d_e), \cdots, (u^i(p_k), d_e)) .$$

The use of these conventions is justified by Definition 4.6 and Lemma 3.3 respectively

THEOREM 4.9. *In a game G of perfect information, every supergame p-strategy is equivalent to a supergame p^*-strategy.*

Proof. Let f^i be a supergame p-strategy in G. For any sequence of terminals $(\alpha_1, \cdots, \alpha_k)$ we may define $g^i_k(\alpha_1, \cdots, \alpha_k)$ by means of the following recursion:

$$(1) \qquad g^i_0 = f^i_0$$

$$(2) \quad g^i_j(\alpha_1, \cdots, \alpha_j) = f^i_j((\alpha_1, g^i_0), (\alpha_2, g^i_1(\alpha_1)), \cdots, (\alpha_j, g^i_{j-1}(\alpha_1, \cdots, \alpha_{j-1}))), \quad j \leq k .$$

Let θ^{N-i} be an arbitrary supergame p-strategy $(N - i)$-vector. We prove by induction on k that

$$(3) \qquad z_k(f^i, \theta^{N-i}) = z_k(g^i, \theta^{N-i}) , \qquad k \geq 1 .$$

For $k = 1$, (3) follows at once from (1) and 6.3 of [**1**]. Suppose (3) has been proved for $k \leq j$. Set

$$(4) \qquad \xi = (f^i, \theta^{N-i}) , \qquad \zeta = (g^i, \theta^{N-i}) .$$

Then by the induction hypothesis,

$$(5) \qquad z_k(\xi) = z_k(\zeta) , \qquad k \leq j .$$

Hence

$$\begin{aligned}
(6) \quad g^i_j(z_j(\xi)) &= g^i_j(x_1(\xi), x_2(\xi), \cdots, x_j(\xi)) \qquad \text{(by 3.2)} \\
&= g^i_j(x_1(\xi) \mid W, x_2(\xi) \mid W, \cdots, x_j(\xi) \mid W) \qquad \text{(by 4.7)} \\
&= f^i_j((x_1(\xi) \mid W, g^i_0), (x_2(\xi) \mid W, g^i_1(z_1(\xi))), \cdots, (x_j(\xi) \mid W, g^i_{j-1}(z_{j-1}(\xi)))) \ \text{(by (2))} \\
&= f^i_j((x_1(\xi) \mid W, g^i_0), (x_2(\xi) \mid W, g^i_1(z_1(\zeta))), \cdots, (x_j(\xi) \mid W, g^i_{j-1}(z_{j-1}(\zeta)))) \ \text{(by (5))}.
\end{aligned}$$

Now by 6.2 and 6.5 of [**1**], we have

$$\begin{aligned}
(7) \qquad x_k(\zeta) &= s(\zeta_{k-1}(z_{k-1}(\zeta))) \\
&= (u(c(\zeta_{k-1}(z_{k-1}(\zeta)))), d(\zeta_{k-1}(z_{k-1}(\zeta)))) .
\end{aligned}$$

But by 2.7 of [**1**] and by 3.2,

$$c(\zeta_{k-1}(z_{k-1}(\zeta))) = \zeta_{k-1}(z_{k-1}(\zeta))$$

Applying this to (7), we obtain

$$\begin{aligned}
x^i_k(\zeta) \mid U^i_k &= u^i(\zeta_{k-1}(z_{k-1}(\zeta))) \\
&= (\lambda(\zeta_{k-1}(z_{k-1}(\zeta))), \zeta^i_{k-1}(z_{k-1}(\zeta))) \qquad \text{(by 4.3)} \\
&= (\lambda(\zeta_{k-1}(z_{k-1}(\zeta))), g^i_{k-1}(z_{k-1}(\zeta))) \qquad \text{(by (4))} .
\end{aligned}$$

Hence

$$x_k^i(\zeta) \mid P^i = g_{k-1}^i(z_{k-1}(\zeta)) , \qquad k \geq 1 .$$

Applying this to (6), we obtain

$$\begin{aligned} g_j^i(z_j(\xi)) &= f_j^i((x_1(\xi) \mid W, x_j^i(\zeta) \mid P^i), \cdots, (x_j(\xi) \mid W, x_1^i(\zeta) \mid P^i)) \\ &= f_j^i((x_1(\xi) \mid W, x_j^i(\xi) \mid P^i), \cdots, (x_j(\xi) \mid W, x_1^i(\xi) \mid P^i)) \qquad \text{(by (5))} \\ &= f_j^i(x_1, (\xi) \mid U_1^i, \cdots, x_j(\xi) \mid U_k^i) \qquad \text{(by 4.3)} \\ &= f_j^i(x_1(\xi), \cdots x_j(\xi)) \qquad \text{(by 4.8)} \\ &= f_j^i(z_j(\xi)) . \end{aligned}$$

Applying (4), we obtain

$$\xi_j(z_j(\xi)) = \zeta_j(z_j(\xi)) . \tag{8}$$

Hence

$$\begin{aligned} & x_{j+1}(\zeta) \mid U_{j+1} \\ &= u(\zeta_j(z_j(\zeta))) \qquad \text{(by 6.5 of [1])} \\ &= u(\zeta_j(z_j(\xi))) \qquad \text{(by (5))} \\ &= u(\xi_j(z_j(\xi))) \qquad \text{(by (8))} \\ &= x_{j+1}(\xi) \mid U_{j+1} . \end{aligned}$$

Hence by 3.2,

$$z_{j+1}(\zeta) \mid U_1 \times \cdots \times U_{j+1} = z_{j+1}(\xi) \mid U_1 \times \cdots \times U_{j+1} , \tag{9}$$

and since

$$z_{j+1}(\zeta) \mid R_1 \times \cdots \times R_{j+1} = d_e \times \cdots \times d_e = z_{j+1}(\xi) \mid R_1 \times \cdots \times R_{j+1} ,$$

we conclude from (9) that

$$z_{j+1}(\zeta) = z_{j+1}(\xi) .$$

This completes the induction and the proof of (3). Applying 6.6 and 6.7 of [1] to (3), we obtain

$$H_k(f^i, \theta^{N-i}) = H_k(g^i, \theta^{N-i}) , \qquad k \geq 1 .$$

Hence by 4.4,

$$f^i \simeq g^i .$$

But g^i is by its definition a supergame p^*-strategy, and thus our proof is completed.

Parallel to Definition 3.5, we may make the following definition:

DEFINITION 4.10. *A strong equilibrium p^*-point f is a summable supergame p^*-strategy vector for which there is no $B \subset N$ and supergame p^*-strategy vector g satisfying* 3.6 *and* 3.7.

The set of all strong equilibrium p^*-points is denoted by S_{p^*}. If 3.7 is replaced by 3.8, the resulting set of points is denoted by $\tilde{S}_{p^*}$.

If we can succeed in restricting our considerations to supergame p^*-strategies then we will have considerably simplified our problem, because then the information available about previous plays is the same for all players (so that the information function may be regarded as 1-dimensional rather than n-dimensional). That we may without loss of generality restrict ourselves in this way is the content of the next theorem.

THEOREM 4.11. *In a game C of perfect information, a summable supergame p-strategy vector f is a strong equilibrium p-point if there is a strong equilibrium p^*-point f_* equivalent to f.*

Proof. Suppose

$$f \notin S_p .$$

Then there is a $B \subset N$ and a supergame p-strategy vector g satisfying 3.6 and 3.7. In accordance with 4.9, there is a supergame p^*-strategy B-vector g_*^B for which

(1) $$g_*^B \simeq g^B .$$

Define

(2) $$g_*^{N-B} = f_*^{N-B} .$$

By hypothesis we have

(3) $$f_*^{N-B} \simeq f^{N-B} .$$

Combining (1), (2), (3), and 3.6, we obtain

(4) $$g_* \simeq g .$$

From (4), 3.4 and 4.5 it follows that

(5) $$S_k(g) = S_k(g_*)$$

for each k. By hypothesis,

(6) $$f \simeq f_* .$$

Applying 6.8 of [1] and 4.5 to (6), we obtain

(7) $$H(f) = H(f_*) .$$

From 3.7, (5), and (7) it follows that

(8) $$\limsup_{k\to\infty} \min_{i \in B} (S_k^i(g_*) - H^i(f_*)) > 0 .$$

From (2), (8), and 4.10 it follows that $f_* \notin S_{p^*}$, which contradicts the hypothesis. This completes the proof.

COROLLARY 4.12. $H(S_{p^*}) \subset H(S_p)$.

Proof. Follows from 6.8 of [1], 4.5, and 4.11.

The following theorems (4.13 through 4.16) will not be used in the sequel; they are included for the sake of completeness. The proofs use the same ideas as those already, given, and will be omitted.

THEOREM 4.13. *Conversely to* 4.11, *a summable supergame p-strategy vector f is a strong equilibrium p-point only if there is a strong equilibrium p^*-point f_* equivalent to f.*

COROLLARY 4.14. $H(S_p) = H(S_{p^*})$.

THEOREM 4.15. $S_{p^*} \subset \tilde{S}_{p^*}$.

THEOREM 4.16. $H(\tilde{S}_p) = H(\tilde{S}_{p^*})$.

Theorems analogous to 4.11 and 4.13 for S_{p^*} may also be proved.

For supergame p^*-strategy vectors f, formulas 6.3 through 6.7 of [1] may be rewritten as follows:

$$z_1 = \lambda(f_0) \tag{4.17}$$

$$z_k = (z_{k-1}, x_k) \tag{4.18}$$

$$x_k = \lambda(f_{k-1}(z_{k-1})) \tag{4.19}$$

$$E_k = E(f_{k-1}(z_{k-1})) \tag{4.20}$$

$$H_k = H(E_k)\ . \tag{4.21}$$

Here we are making use of the notation introduced in convention 4.7.

5. The main theorem. We make use of two lemmas. The first tells us that at an acceptable point in a game of perfect information, $N - B$ can always retaliate for a defection by B by means of a single pure strategy. The second tells us that any payoff that can be obtained by a c-strategy vector in G can also be obtained by a supergame p-strategy vector (or even by a supergame p^*-strategy vector).

LEMMA 5.1. *Let G be a game of perfect information. Let $B \subset N$, and let h be a vector. If there is a $c^{N-B} \in C^{N-B}$ such that for all $c^B \in C^B$, there is an $i \in B$ for which*

$$H^i(c^B, c^{N-B}) \leq h^i\ , \tag{1}$$

then there is a $p^{N-B} \in P^{N-B}$ such that for all $c^B \in C^B$, there is an $i \in B$ for which

(2) $$H^i(c^B, p^{N-B}) \leq h^i .$$

Proof. $H^B(C^B, c^{N-B})$ is easily seen to be a convex subset of the euclidean B-space R^B. By the hypothesis of the lemma, $H^B(C^B, c^{N-B})$ cannot intersect the open "corner" or "octant" in R^B given by the inequalities

(3) $$x^i > h^i , \qquad i \in B .$$

This "corner" is also convex. Applying the separation theorem for convex sets[3], we obtain a hyperplane

(4) $$\sum_{i \in B} a^i x^i = k$$

which passes through h^B, and which separates $H^B(C^B, c^{N-B})$ from the "corner" given by (3). In other words, we have

(5) $$\sum_{i \in B} a^i h^i = k ,$$

and we may assume without loss of generality that

(6) $$\sum_{i \in B} a^i H^i(c^B, c^{N-B}) \leq k \quad \text{for } c^B \in C^B$$

and

(7) $$\sum_{i \in B} a^i x^i > k \text{ for } x^B \text{ satisfying (3) .}$$

(if the inequalities (6) and (7) are reversed, then we may obtain them in the given form by multiplying both sides of (4) by -1). From (3) and (7) it follows that

(8) $$a^B \geq 0 .$$

Since (4) defines a hyperplane, there must also be an $i \in B$ for which

(9) $$a^i \neq 0 .$$

Define a two-person, zero-sum game G_* as follows: There are two players, 1 and 2. The game tree of G_* is the same as that of G, and G_* is also a game of perfect information. Player 1 has all the moves that members of B have in C, and player 2 has all the moves that members of $N - B$ have in G. Thus the mixed strategy space of player 1 is C^B, and the mixed strategy space of player 2 is C^{N-B} (we will also use the notation M^1 and M^2 for these mixed strategy spaces). The payoff in G_* will be denoted by H_*; it is defined by

(10) $$H^1_*(p^B, p^{N-B}) = \sum_{i \in B} a^i H^i(p^B, p^{N-B})$$

[3] See for instance [**10**], pp. 29 and 81.

(11) $$H_*^2 = -H_*^1 .$$

From (10) it follows that

(12) $$H_*^1(c^B, c^{N-B}) = \sum_{i \in B} a^i H^i(c^B, c^{N-B})$$

for all $c^B \in C^B$. Combining (12) with (6) and the hypothesis of the lemma, we obtain the existence of a $c^{N-B} \in C^{N-B}$ such that for all $c^B \in C^B$, we have

$$H_*^1(c^B, c^{N-B}) \leq k .$$

Restated in terms of mixed strategies in C_*, we have the existence of a mixed strategy $m_*^2 \in M_*^2$ (namely c^{N-B}), such that for all $m_*^1 \in M_*^1$, we have

(13) $$H_*^1(m_*^1, m_*^2) \leq k .$$

By (11), G_* is zero-sum as well as two person. (13) merely tells us that

(14) $$v(G_*) \leq k ,$$

where $v(G_*)$ denotes the value of G_*. By the theorem of von Neumann on two-person zero-sum games of perfect information, we have the existence of an optimal *pure* strategy for player 2 in G_*. Hence there is a $p_*^2 \in p_*^2$ (i.e. a $p^{N-B} \in P^{N-B}$) such that for all $m_*^1 \in M_*^1$ (i.e. for all $c^B \in C^B$), we have

$$H_*^1(m_*^1, p_*^2) \leq v(G_*)$$

(i.e., by (10),

(15) $$\sum_{i \in B} a^i H^i(c^B, p^{N-B}) \leq v(G_*)) .$$

Combining (5), (14), and (15), we obtain

(16) $$\sum_{i \in B} a^i H^i(c^B, p^{N-B}) \leq \sum_{i \in B} a^i h^i$$

for all $c^B \in C^B$.

From (16) it follows that

$$\sum_{i \in B} a^i (H^i(c^B, p^{N-B}) - h^i) \leq 0$$

for all $c^B \in C^B$. Combining this with (8) and (9), we obtain for each $c^B \in C^B$, the existence of at least one $i \in B$ for which

$$H^i(c^B, p^{N-B}) - h^i \leq 0$$

This completes the proof of the lemma.

The next lemma tells us that the non-negative integers can be partitioned into disjoint subsets whose asymptotic densities will yield an arbitrary finite set of non-negative real numbers adding up to 1.

LEMMA 5.2. *Let Z be a finite set, and let $y \in C(Z)$. For any mapping π from the set K of all non-negative integers into Z, and for any $k \in K$ and $z \in Z$, let $\rho_\pi(k; z)$ denote the number of $j \in K$ for which*

$$j \leq k$$

and

$$\pi(j) = z .$$

Then there is a π for which

$$\lim_{k \to \infty} \frac{\rho_\pi(k; z)}{k+1} = y(z)$$

for all $z \in Z$.

$\pi(j)$ will also be denoted by $\pi(j; y)$, and $\rho_\pi(k, z)$ by $\rho(k; z; y)$.

The proof is not difficult. It will be omitted.

THEOREM 5.3. *In a game G of perfect information,*

$$H(A_c) \subset H(S_{p^*}) .$$

Proof. In its main outlines, the proof is analogous to that of Theorem 3 of [1], which states that $H(A_c) \subset H(S_c)$. The details, however, differ considerably in the two cases. Both proofs are divided into three parts: Given an acceptable payoff vector h, we must first find a sequence of strategy vectors which will yield a payoff of h in the supergame (under the assumption that the players are all "loyal"). Next, we must find a way to determine which players, if any, are disloyal; and finally, we must find a way to punish the disloyal players. All these elements must be incorporated into a supergame strategy vector. In Theorem 3 of [1], the first of these tasks was accomplished by having the players play the same c-strategy vector on each play, namely the one that yields an expected payoff of h. Here this cannot be done, because the players must restrict themselves to pure strategies on each play. They must therefore play different pure strategy vectors on different plays in such a way so that the limiting payoff is h; to show that this can be done, use must be made of Lemma 5.2. As for the second task, this was accomplished in Theorem 3 of [1] by simply noting the make-up of the coalitions; here this cannot be done, because in supergame p^*-strategy vectors, there are no coalitions. Instead, use must be made of the perfect information that each player has. Finally, a group B of disloyal players could be punished in Theorem 3 of [1] by use of the c-strategy $(N-B)$-vector c^{N-B} provided in the definition of acceptability; here only pure strategy $(N-B)$-vectors may be used, so that recourse must be had to Lemma 5.1. For a more detailed intuitive statement of the proof, see § 10 of [1].

We now give the detailed proof of 5.3.

Let $h \in H(A_c)$, and suppose $\gamma^N \in A_c$ is such that

$$H(\gamma^N) = h \,. \tag{1}$$

Then by 4.3 of [1], for each $B \subset N$ there is a $c^{N-B} \in C^{N-B}$, such that for each $c^B \in C^B$, there is an $i \in B$ for which

$$H^i(c^B, c^{N-B}) \leq h^i \,.$$

Applying Lemma 5.1, we obtain for each $B \subset N$ a pure strategy $(N-B)$-vector γ^{N-B}, such that for each $c^B \in C^B$, there is an $i \in B$ for which

$$H^i(c^B, \gamma^{N-B}) \leq h^i \,. \tag{2}$$

For each $j \geq 1$, let W_j be a copy of W. W_j represents the set of possible outcomes of the jth play. Let

$$Q_k = W_1 \times \cdots \times W_k \,;$$

Q_k represents the set of possible outcomes for the first k plays, and as such is the domain of the function f^i_k.

Let g be any supergame p^*-strategy vector in G. We define a *compliance function* $\alpha(v_1, \cdots, v_k; g)$ for all $(v_1, \cdots, v_k) \in Q_k$ as follows:

DEFINITION (3). $\alpha(v_1, \cdots, v_k; g)$ is the *maximal* subset A of N for which

$$v_j \in \lambda(g^A_{j-1}(v_1, \cdots, v_{j-1}) \times P^{N-A}) \qquad \text{for } j = 1, \cdots, k \,.$$

For each member of Q_k, α tells which subset of N has been "loyal" to, or has complied with, the supergame p^*-strategy vector g.

It is not difficult to see that for each g, we have

$$\alpha(z_k(g); g) = N \qquad \text{for } k \geq 1 \,. \tag{4}$$

To show (4), it is sufficient to show that N is the maximal set satisfying (3), i.e. that we have

$$\begin{aligned} x_j(g) &= \lambda(g^N_{j-1}(z_{j-1}(g)) \times P^{N-N}) \,, && j = 1, \cdots, k \\ &= \lambda(g_{j-1}(z_{j-1}(g))) \,, && j = 1, \cdots, k \,. \end{aligned}$$

But this follows at once from 4.19.

Moreover, it follows from (3) that

$$\alpha = N \text{ when } k = 0 \,. \tag{5}$$

We are now ready to define a strong equilibrium p^*-point whose payoff is h.

For $k \geq 0$ and $q_k \in Q_k$, define

$$\begin{cases} f_k(q_k) = \pi(k;\gamma^N)\,, & \text{if } \alpha(q_k; f) = N \\ f_k^{\alpha(q_k;f)}(q_k) = \gamma^{\alpha(q_k;f)} & \Big\} \text{ otherwise} \,. \\ f_k^{N-\alpha(q_k;f)}(q_k) = \text{arbitrary} & \end{cases} \tag{6}$$

Definition (6) is a recursive definition; $\alpha(q_k; f)$ depends only on $f_0, \cdots, f_{k-1}$, not on f_k.

Set $z_k = z_k(f)$ for $k \geq 1$. We first prove

$$f_k(z_k) = \pi(k; \gamma^N) \text{ for } k \geq 0 \,. \tag{7}$$

For $k > 0$, (7) follows from (6) and (4); for $k = 0$, it follows from (6) and (5).

Combining (7) with 4.20 and 4.21, we obtain

$$H_{k+1}(f) = H(\pi(k; \gamma^N)) \text{ for } k \geq 0 \,. \tag{8}$$

Hence

$$\begin{aligned} \sum_{r=1}^{k+1} H_r(f) &= \sum_{r=0}^{k} H_{r+1}(f) \\ &= \sum_{r=0}^{k} H(\pi(r;\gamma^N)) \qquad \text{(by (8))} \\ &= \sum_{y \in P} \rho(k; y; \gamma^N) H(y) \qquad \text{(by 5.2)} \,. \end{aligned}$$

Hence

$$\begin{aligned} \lim_{k\to\infty} \frac{1}{k} \sum_{r=1}^{k} H_r(f) &= \lim_{k\to\infty} \frac{1}{k+1} \sum_{r=1}^{k+1} H_r(f) \\ &= \lim_{k\to\infty} \frac{1}{k+1} \sum_{y\in P} \rho(k; y; \gamma^N) H(y) \\ &= \sum_{y\in P} H(y) \lim_{k\to\infty} \frac{1}{k+1} \rho(k; y; \gamma^N) \\ &= \sum_{y\in P} \gamma^N(y) H(y) \qquad \text{(by 5.2)} \\ &= H(\gamma^N) \\ &= h \qquad \text{(by (1))} \,. \end{aligned}$$

Applying 6.8 of [1], we obtain

$$(9) \qquad H(f) = h\ .$$

By 3.17, f is also summable.

It remains to prove that f is a strong equilibrium p^*-point. Suppose not. Then there is a $B \subset N$ and a supergame p^*-strategy vector g satisfying 3.6 and 3.7. We must then have

LEMMA (10). *$\alpha(z_k(g); f)$ is monotone decreasing with k.*

Proof. By 4.18,

$$z_k(g) = (x_1(g), \cdots, x_k(g))\ .$$

The result now follows from (3).

From 4.19 we obtain

$$\begin{aligned} x_j(g) &= \lambda(g_{j-1}(z_{j-1}(g))) \\ &= \lambda(g_{j-1}^{N-B}(z_{j-1}(g)), g_{j-1}^{B}(z_{j-1}(g))) \\ &\in \lambda(g_{j-1}^{N-B}(z_{j-1}(g)) \times P^B) \\ &= \lambda(f_{j-1}^{N-B}(z_{j-1}(g)) \times P^B) \qquad \text{(by 3.6)} \\ &= \lambda(f_{j-1}^{N-B}(x_1(g), \cdots, x_{j-1}(g)) \times P^B) \qquad \text{(by 4.18)}\ . \end{aligned}$$

It now follows from (3) that

$$(11) \qquad N - B \subset \alpha(z_k(g); f) \text{ for } k \geq 1\ .$$

Combining (11) with (5), we obtain

$$(12) \qquad N - B \subset \alpha(z_k(g); f) \text{ for } k \geq 0\ .$$

From (10) we obtain the existence of a set $B(g) \subset N$ and a non-negative integer k_0 such that

$$(13) \qquad \alpha(z_k(g); f) = N - B(g) \text{ for } k \geq k_0\ .$$

Combining (12) and (13), we obtain

$$(14) \qquad B(g) \subset B\ .$$

If $B(g) = \phi$, then from (13) we obtain

$$\alpha(z_k(g); f) = N \text{ for } k \geq k_0\ ,$$

whence, using (10), we deduce that

$$(15) \qquad \alpha(z_k(g); f) = N \text{ for } k \geq 0\ .$$

Using (3) and 4.18, we deduce from (15) that

$$(16) \qquad x_k(g) = \lambda(f_{k-1}(z_{k-1}(g))) \text{ for } k \geq 1\ .$$

From (16) and 4.17 we deduce

$$x_1(g) = \lambda(f_0) = x_1(f) ,$$

and a simple inductive argument based on (16), 4.18 and 4.19 leads to the conclusion that

$$z_k(g) = z_k(f) \text{ for } k \geq 1 .$$

Applying 4.20 and 4.21, we obtain

$$H_k(g) = H_k(f) \text{ for } k \geq 1 . \tag{17}$$

From 6.8 of [**1**], 3.4, and (17) it follows that

$$\lim_{k \to \infty} S_k(g) = H(f) ,$$

which contradicts 3.6. Thus the assumption $B(g) = \phi$ has led to a contradiction, and we may conclude that

$$B(g) \neq \phi . \tag{18}$$

Combining (6), (13), and (18), we obtain

$$f_k^{N-B(g)}(z_k(g)) = \gamma^{N-B(g)} \text{ for } k \geq k_0 . \tag{19}$$

Let μ be the payoff function defined on W, so that

$$H = \mu \circ \lambda . \tag{20}$$

Our μ is what is called h in [**3**]; it may also be defined by

$$\mu = \psi \mid W ,$$

where ψ is as in § 6 of [**1**]. We then have

$$\begin{aligned} H_k(g) &= H(E_k(g)) \qquad \text{(by 4.21)} \\ &= H(g_{k-1}(z_{k-1}(g))) \qquad \text{(by 4.20)} \\ &= \mu(\lambda(g_{k-1}(z_{k-1}(g)))) \qquad \text{(by (20))} \\ &= \mu(x_k(g)) \qquad \text{(by 4.19)} . \end{aligned} \tag{21}$$

Now by (3), (13), and 4.19, we have

$$x_k(g) = \lambda(f_{k-1}^{N-B(g)}(z_{k-1}(g)), p_k^{B(g)}) , \tag{22}$$

where $p_r^{B(g)}$ is some member of $P^{B(g)}$.
Hence for $k > k_0$, we have

$$\begin{aligned} H_k(g) &= \mu(x_k(g)) \qquad \text{(by (21))} \\ &= (\mu \circ \lambda)(f_{k-1}^{N-B(g)}(z_{k-1}(g)), p_k^{B(g)}) \qquad \text{(by (22))} \\ &= H(\gamma^{N-B(g)}, p_k^{B(g)}) \qquad \text{(by (20) and (19))} . \end{aligned}$$

Hence for $k > k_0$, we have by the linearity of H that

$$\text{(23)}\qquad \frac{1}{k-k_0}\sum_{r=k_0+1}^{k} H_r(g) = \frac{1}{k-k_0}\sum_{r=k_0+1}^{k} H(\gamma^{N-B(g)}, p_r^{B(g)})$$

$$= H\left(\gamma^{N-B(g)}, \sum_{r=k_0}^{k+1}{}^{*} \frac{1}{k-k_0}\, p_r^{B(g)}\right)$$

Applying (2), we obtain the existence of an $i \in B(g)$ such that

$$\text{(24)}\qquad H^i\left(\gamma^{N-B(g)}, \sum_{r=k_0}^{k+1}{}^{*} \frac{1}{k-k_0}\, p_r^{B(g)}\right) - h^i \leq 0 \;.$$

Combining (23) and (24), we deduce that

$$\min_{i\in B(g)}\left(\left(\frac{1}{k-k_0}\sum_{r=k_0+1}^{k} H_r^i(g)\right) - h^i\right) \leq 0 \;;$$

from this and (14) it follows what

$$\text{(25)}\qquad \min_{i\in B}\left(\left(\frac{1}{k-k_0}\sum_{r=k_0+1}^{k} H_r^i(g)\right) - h^i\right) \leq 0 \;.$$

Now it follows easily from the boundedness of H that as $k \to \infty$,

$$\frac{1}{k-k_0}\sum_{r=k_0+1}^{k} H_r^i(g) = \frac{1}{k}\sum_{j=1}^{k} H_j^i(g) + O\left(\frac{1}{k}\right)$$

$$= S_k^i(g) + o(1) \qquad \text{(by 3.4)} \;.$$

Applying this to (25), we obtain that as $k \to \infty$,

$$\min_{i\in B}\,(S_k^i(g) - h^i) \leq o(1) \;,$$

whence

$$\text{(26)}\qquad \limsup_{k\to\infty}\,\min_{i\in B}\,(S_k^i(g) - h^i) \leq 0 \;.$$

Applying (9), we see that (26) contradicts 3.7. This completes the proof of 5.3.

THEOREM 5.4. *In a game G of perfect information,*

$$H(A_c) = H(S_p) = H(\tilde{S}_p) \;.$$

In particular, h is a c-acceptable payoff vector in G, if and only if there is a strong equilibrium c-point f in supergame pure strategies for which

$$H(f) = h \;.$$

Proof. We have

$$\begin{aligned} H(A_c) &\subset H(S_{p^*}) && \text{(by 5.3)} \\ &\subset H(S_p) && \text{(by 4.12)} \\ &\subset H(S_c) && \text{(by 3.13)} \\ &= H(A_c) && \text{(by Corollary 4 of [1])} \end{aligned}$$

Hence equality must hold throughout, and in particular,

$$(1) \qquad H(A_c) = H(S_p)\ .$$

Next, we have

$$\begin{aligned} H(S_p) &\subset H(\tilde{S}_p) && \text{(by 3.16)} \\ &\subset H(\tilde{S}_c) && \text{(by 3.12)} \\ &= H(A_c) && \text{(by Corollary 4 of [1])} \\ &= H(S_p) && \text{(by (1))}\ . \end{aligned}$$

Hence equality must hold throughout, and we deduce

$$(2) \qquad H(S_p) = H(\tilde{S}_p)\ .$$

(1) and (2) yield the first part of 5.4. The second part follows at once from 3.11 and the first part.

COROLLARY 5.5. *Every stable*[4] *game of perfect information has strong equilibrium c-points in supergame pure strategies.*

6. The converse of the main theorem. For two-person zero-sum games not involving chance, Von Neumann's theorem is known to "characterize" games of perfect information (see [4]). More precisely, if Γ is a game structure of the above type which has the property that every game that can be obtained from Γ (by adjunction of a payoff function μ) has optimal pure strategies, then Γ must be equivalent to a game structure of perfect information. What can be said in this regard for the theory presented in the previous sections?

For one thing, it is of interest to know that there are *some* games that do not satisfy our main theorem (Theorem 5.4). Indeed, "matching pennies" is such a game.
This game is given by

$$\begin{aligned} N &= (1,2) \\ P^1 &= (p_1^1, p_2^1) \\ P^2 &= (p_1^2, p_2^2) \end{aligned}$$

[4] That is, every game that has any c-acceptable points (or, equivalently, any strong equilibrium c-points). See § 11 of [1].

$$H^1(p_i^1, p_j^2) = \begin{cases} 1 & \text{if } i = j \\ -1 & \text{if } i \neq j \end{cases}$$
$$H^2(p) = -H^1(p)$$
$$u(p) = p \ .$$

It is a two-person zero-sum game with value 0; hence by Theorem 1 of [**1**], we have

$$H(A_c) = (0, 0) \ .$$

If 5.4 holds for this game, then it follows that

$$H(S_p) = (0, 0) \ ,$$

and in particular, there is a summable strong equilibrium p-point f such that

(1) $$H(f) = (0, 0) \ .$$

Define a supergame p-strategy vector g by

(2) $$g^2 = f^2$$

and

$$g_k^1(v_1, \cdots, v_k) = p_i^1, \text{ for } k \geq 0, \ (v_1, \cdots, v_k) \in J_1 \times \cdots \times J_k \ ,$$

where i is such that

$$f_k^2(v_1, \cdots, v_k) = p_i^2 \ .$$

It is then easily seen that

$$H_k^1(g) = 1 \qquad \text{for } k \geq 0 \ ,$$

whence it follows that

$$S_k^1(g) = 1 \qquad \text{for } k \geq 0 \ .$$

Combining this with (1), we see that g satisfies 3.7 for $B = (1)$. By (2), g satisfies 3.6 for $B = (1)$. Hence f cannot be a strong equilibrium p-point.

The above example constitutes a formalization of the familiar argument that states that no "scheme" for playing a long sequence of penny-matchings that involves only pure strategies can be optimal.

The general statement of the converse would be as follows:

CONJECTURE. *Let Γ be a game structure and suppose that every stable game that is obtained from Γ by adjunction of a payoff function μ has a strong equilibrium p-point. Then Γ is essentially equivalent* (in the sense of [4]) *to a game structure of perfect information.*

There is little doubt in my mind that this conjecture is true, if not in the given form, then at least in some other closely allied form.

7. Notation for non-cooperative games. We will make use of the notion of mixed strategies. Formally, the space M^i of mixed strategies of player i is defined to be identical with C^i. If B is a subset of N, then we define

$$M^B = \prod_{i \in B} M^i \; ; \tag{7.1}$$

the cartesian product is meant. It follows that

$$M^B \subset C^B \; ;$$

the opposite inequality is generally false. The prefix m- is an abbreviation for "mixed". The definitions relating to payoff remain unchanged.

8. Acceptable points for non-cooperative games. The non-cooperative game differs from the cooperative game chiefly in that the use of correlated strategy vectors that are not also mixed strategy vectors is forbidden. The definition of acceptability for non-cooperative games will therefore be the same as that for cooperative games (see [**1**], section 4), except that correlated strategy vectors must be replaced throughout by mixed strategy vectors. The intuitive reasoning behind the definition remains unchanged. It might be objected that the "concerted action" that is necessary to prevent a set of players B from obtaining a payoff that is higher than at an acceptable point, is forbidden under non-cooperative rules. In fact, such concerted action will probably arise anyway as part of a "silent gentlemen's agreement" among the players of $N - B$. The only restriction is that though the players may "cooperate" in this sense (indeed, they cannot be prevented from so doing), they may not correlate their mixed strategies before a play.

Further intuitive discussion of the notion of m-acceptability will be found in a subsequent paper, devoted exclusively to acceptable points in non-cooperative games.

The formal definitions are as follows:

DEFINITION 8.1. *Let $m_0 \in M$. m_0 is said to be m-acceptable if there is no $B \subset N$ such that for each $m^{N-B} \in M^{N-B}$, there is an $m^B \in M^B$ for which*

$$H^B(m^B, m^{N-B}) > H^B(m_0) \; .$$

The set of all m-acceptable m-strategy vectors is denoted by A_m. Like c-acceptability m-acceptability is a "global" notion (see [**1**], § 4).

DEFINITION 8.2. *A payoff vector h is said to be m-acceptable, if for some $m \in A_m$, we have*

$$H(m) = h .$$

The following is a trivial restatement of 8.2:

THEOREM 8.3. A payoff vector h is m-acceptable if and only if for each $B \subset N$, there is an $m^{N-B} \in M^{N-B}$, such that for all $m^B \subset M^B$, there is an $i \in B$ for which

$$H^i(m^B, m^{N-B}) \leq h^i .$$

We remark that as in the cooperative case, all two-person games have m-acceptable points. When we go beyond two-person games we find games that have no m-acceptable points. The example given in § 11 of [1] holds for the non-cooperative case as well, as does the intuitive discussion following the example.

We remark also that even in the two-person case, there are games of perfect information that have no m-acceptable points in pure strategies. See § 2 of this paper, which applies unchanged in its entirety to the non-cooperative case.

9. Equivalence of M-acceptability and C-acceptability in games of perfect information.

THEOREM 9.1. *In a game G of perfect information,*

$$H(M) = H(C)$$

Proof. $H(M) \subset H(C)$ follows at once from $M \subset C$. It remains to prove

$$H(C) \subset H(M) .$$

Instead of proving this, we will prove a more general version that we will need later. What we need for 9.1 follows from 9.2 if we set $B = N$.

LEMMA 9.2. *Let G be a game of perfect information. Then with each $c^B \in C^B$, we may associate an $m^B \in M^B$, such that for all $c^{N-B} \in C^{N-B}$ we have*

$$H(c^B, c^{N-B}) = H(m^B, c^{N-B}) .$$

Proof. Fix c^B. Because of the linearity of H, it is sufficient to prove that there is an m^B such that for all $p^{N-B} \in P^{N-B}$ we have

$$H(c^B, p^{N-B}) = H(m^B, p^{N-B}) . \tag{1}$$

Let b be the cardinality of B. With each $i \in B$, we may associate an $n-b+1$ person game G_i as follows: The players are 0 and the members of $N-B$. (Intuitively, 0 represents the coalition of all the members of B.) The set of pure strategies of 0 is P^B, while the set of pure strategies for a member j of $N-B$ is P^j. The payoff to 0 is given by E^i, to members j of $N-B$ by E^j. To avoid confusion, we will denote the payoff in G_i by E_i, the expected payoff by H_i. E_i and H_i are $((0) \cup N-B)$-vectors.

From the definition of G_i, we see that for all $p^{N-B} \in P^{N-B}$, we have

$$(2) \qquad \begin{cases} H^{N-B}(c^B, p^{N-B}) = H_i^{N-B}(c^B, p^{N-B}) \\ H^i(c^B, p^{N-B}) = H_i^0(c^B, p^{N-B}) \, . \end{cases}$$

In G_i, c^B is a mixed strategy of player 0. Let β^* be its behavior (see [9], § 5, which will be called (*) in the sequel; Definition 16). Since G_i depends on i only because of its payoff, and since the behavior of a mixed strategy has nothing to do with the payoff, β^* is independent of i. Since G is of perfect information, so is G_i, and hence in particular, G_i is of perfect recall. Noting that every pure strategy is also a behavior strategy, and in fact its own behavior, and applying Theorem 4 of (*), we obtain that for all $p^{N-B} \in P^{N-B}$,

$$(3) \qquad H_i(c^B, p^{N-B}) = H_i(\beta^*, p^{N-B}) \, .$$

Returning to the game G, define behavior strategies β^i for each $i \in B$ by

$$\beta^i = \beta^* \mid \mathscr{U}^i \, ,$$

where $\mathscr{U}^i$ is the set of information sets for player i.
Then from Definitions 14 and 15 of (*) it follows that for all $d^{N-B} \in P^{N-B}$,

$$(4) \qquad \begin{cases} H_i^0(\beta^*, p^{N-B}) = H^i(\beta^B, p^{N-B}) \\ H_i^{N-B}(\beta^*, p^{N-B}) = H^{N-B}(\beta^B, p^{N-B}) \, . \end{cases}$$

Combining (2), (3), and (4), we obtain that for all $p^{N-B} \in P^{N-B}$,

$$H^{N-B}(c^{N-B}, p^{N-B}) = H^{N-B}(\beta^B, p^{N-B})$$

and for all $i \in B$, $H^i(c^{N-B}, p^{N-B}) = H^i(\beta^B, p^{N-B})$; that is,

$$(5) \qquad H(c^{N-B}, p^{N-B}) = H(\beta^B, p^{N-B}) \, .$$

If m^i is the mixed strategy corresponding to β^i in accordance with Lemma 3 of (*), then it follows from Lemma 3 and Theorem 4 of (*) that for all $p^{N-B} \in P^{N-B}$,

$$(6) \qquad H(\beta^B, p^{N-B}) = H(m^B, p^{N-B}) \, .$$

Combining (5) and (6), we obtain (1).

The following theorem will not be used in the sequel. It is included for the sake of completeness.

COROLLARY 9.3. *In a game G of perfect information*

$$E(M) = E(C) .$$

Proof. It is clear that $E(M) \subset E(C)$. To prove $E(C) \subset E(M)$, let $c \in C$. If μ is the payoff function on W, we have

$$E(P) = \mu(W) ,$$

and indeed

(1) $$E = \mu \circ \lambda .$$

Hence if

$$c = \sum_{p \in P}^{*} c(p)p \in C ,$$

then

$$\begin{aligned} H(c) &= \sum_{p \in P} c(p)\mu(\lambda(p)) \\ &\in H(C) \subset H(M) \qquad \text{(by 9.1)} . \end{aligned}$$

It follows that there is a mixed strategy vector m such that

(2) $$\sum_{p \in P} c(p)\mu(\lambda(p)) = \sum_{p \in P}\Big(\prod_{i \in N} m^i(p^i)\Big)\mu(\lambda(p)) .$$

Let us fix the coefficients $c(p)$, and consider a game G' which is the same as G except for its payoff, which is such that the $\mu(w)$ form a set that is linearly independent over the field generated by the coefficients $c(p)$ over the rationals. For this game G', a mixed strategy vector m may be formed that satisfies (2). Both sides of (2) can then be considered as linear combinations of distinct terms of the form $\mu(w)$, and it follows from the way we have chosen G' that the coefficients of the same terms on both sides of (2) must be equal, i.e.,

(3) $$\sum_{p \in \lambda^{-1}(w)} c(p) = \sum_{p \in \lambda^{-1}(w)} \Big(\prod_{i \in N} m^i(p^i)\Big),\ w \in W .$$

Now (3) is seen to hold independent of the payoff; hence no matter how μ is defined, we may write

(4) $$\sum_{w \in W}^{*}\Big(\sum_{p \in \lambda^{-1}(w)} c(p)\Big)\mu(w) = \sum_{w \in W}^{*}\Big(\sum_{p \in \lambda^{-1}(w)}\Big(\prod_{i \in N} m^i(p^i)\Big)\Big)\mu(w) .$$

(note that the outer sum is to be considered a probability distribution rather than an ordinary sum). From (4) we deduce

$$\sum_{p\in P}^{*} c(p)\mu(\lambda(p)) = \sum_{p\in P}^{*}\Big(\prod_{i\in N} m^i(p^i)\Big)\mu(\lambda(p)) ,$$

whence, applying (1), we obtain

$$E(c) = \sum_{p\in P}^{*}\Big(\prod_{i\in N} m^i(p^i)\Big)E(p) = E(m) .$$

This completes the proof.

COROLLARY 9.4. *In a game G of perfect information,*

$$H(A_m) \subset H(A_c) .$$

Proof. Suppose $h \notin H(A_c)$. Then there is a $B \subset N$, such that for all $c^{N-B} \in c^{N-B}$, there is a $c^B \in C^B$ such that

$$(1) \qquad H^B(c^B, c^{N-B}) > h^B .$$

In particular, for all $m^{N-B} \in M^{N-B}$, there is a $c^B \in C^B$ such that

$$(2) \qquad H^B(c^B, m^{N-B}) > h^B .$$

If we let m^B be the mixed strategy B-vector associated with c^B in accordance with Lemma 9.2, then we have

$$(3) \qquad H(m^B, m^{N-B}) = H(c^B, m^{N-B}) .$$

Combining (2) and (3), we obtain that for each m^{N-B}, there is an $m^B \in M^B$ for which

$$H^B(m^B, m^{N-B}) > h^B .$$

Hence $h \notin H(A_m)$, and the corollary follows.

COROLLARY 9.5. *In a game G of perfect information,*

$$H(A_c) \subset H(A_m) .$$

Proof. Suppose $h \in H(A_c)$. Then for all $B \subset N$, there is a $c^{N-B} \in C^{N-B}$, such that for all $c^B \in C^B$, there is an $i \in B$ for which

$$(1) \qquad H^i(c^B, c^{N-B}) \leq h^i .$$

Let m^{N-B} be the mixed strategy $(N-B)$-vector associated with c^{N-B} in accordance with Lemma 9.2. It then follows from 9.2 that for all $c^B \in C^B$,

$$(2) \qquad H(c^B, m^{N-B}) = H(c^B, c^{N-B}) ,$$

and combining (1) and (2), we obtain that for all $c^B \in C^B$, there is an $i \in B$ for which

$$(3) \qquad H^i(c^B, m^{N-B}) \leq h^i \ .$$

In particular, for all $m^B \in M^B$, there is an $i \in B$ for which

$$H^i(m^B, m^{N-B}) \leq h^i \ ,$$

and since this holds for all $B \subset N$, it follows that $h \in H(A_m)$, q.e.d.

COROLLARY 9.6. *In a game G of perfect information,*

$$H(A_c) = H(A_m) \ .$$

COROLLARY 9.7. *In a game G of perfect information,*

$$A_m = A_c \cap M \ .$$

Proof. If $m \in A_m$, then certainly

$$(1) \qquad m \in M \ .$$

But from 9.5 it follows that $H(m) \in H(A_c)$. Since among c-strategy vectors, the property of c-acceptability is a global one, depending only on the payoff, it follows that

$$(2) \qquad m \in A_c \ .$$

Combining (1) and (2), we obtain $m \in A_c \cap M$.

Next, let $c \in A_c \cap M$. Then $c \in M$. We also have $H(c) \in H(A_m)$, and since among m-strategy vectors, the property of m-acceptability is a global one, depending only on the payoff, it follows that

$$c \in A_m \ .$$

This completes the proof.

Because of 9.6 and 9.7, we are justified in dropping the qualifying prefix from the word "acceptable" when discussing games of perfect information.

10. Supergame strategies in the non-cooperative case. A supergame strategy vector for a non-cooperative game is the same as a supergame strategy vector for a cooperative game, except that coalitions are forbidden. Formally, we have

DEFINITION 10.1. *A supergame m-strategy f^i for player i is a supergame c-strategy for which*

$$e(f^i_k(y)) = (i)$$

for all $k \geq 0$ and $y \in J^i_1 \times \cdots \times J^i_k$.

The following theorem follows at once from 10.1:

THEOREM 10.2. *For a supergame m-strategy vector f, we have*

$$c(f_k(y)) \in M$$

for all $k \geq 0$ *and* $y \in J_1 \times \cdots \times J_k$.

Parallel to the definition of strong equilibrium c-point for cooperative games (§ 7 of [**1**]), we may make the following definition for non-cooperative games:

DEFINITION 10.3. *Let f be a summable supergame m-strategy vector. f is a strong equilibrium m-point if there is no* $B \subset N$ *for which there is a supergame m-strategy vector g satisfying* 7.1 *and* 7.2 *of* [**1**].

The set of strong equilibrium m-points will be denoted by S_m. As in [**1**], it is possible to replace 7.2 of [**1**] by 7.3 of [**1**]. The set of points thus obtained will be denoted by $\tilde{S}_m$.

LEMMA 10.4. $F_p \cap S_c \subset S_m$.

Proof. Let

$$(1) \qquad f \in F_p \cap S_c .$$

Since $f \in F_p$, it follows in particular that f is a summable supergame m-strategy vector. Suppose

$$(2) \qquad f \notin S_m .$$

Then there is a $B \subset N$ and a supergame m-strategy vector g satisfying 7.1 and 7.2 of [**1**]. Since every supergame m-strategy vector is also a supergame c-strategy vector, it follows that there is a supergame c-strategy vector g satisfying 7.1 and 7.2 of [**1**]. Hence

$$f \notin S_c ,$$

contradicting (1). Hence (1) implies the falsity of (2), and our result is proved.

LEMMA 10.5. $F_p \cap S_m \subset S_p$.

Proof. The proof is word for word the same as that of the second part of Theorem 3.11 (the part beginning with the word "conversely"; the proof is given before the proof of the first part), except that the two occurrences of the prefix "c-" must be replaced by prefixes "m-". It is also necessary to remember that since g is pure, it is in particular mixed.

THEOREM 10.6. $F_p \cap S_m = S_p$.

Proof. We have

$$S_p = F_p \cap S_c \qquad \text{(by 3.11)}$$
$$\subset S_m \qquad \text{(by 10.4)}\ .$$

Since

$$S_p \subset F_p\ ,$$

it follows that

$$S_p \subset F_p \cap S_m\ .$$

Combining this with 10.5, we obtain 10.6.

THEOREM 10.7. $F_p \cap \tilde{S}_m = \tilde{S}_p$.

Proof. The proof is similar to that of 10.6.

11. The main theorem for non-cooperative games.

THEOREM 11.1. *In a game G of perfect information,*

$$H(A_m) = H(S_p) = H(\tilde{S}_p)\ .$$

In particular, h is an m-acceptable payoff vector in G, if and only if there is a strong equilibrium m-point f in supergame pure strategies for which

$$H(f) = h\ .$$

Proof. The first part follows from 5.3 and 9.6. The second part follows from 10.6 and from the first part.

COROLLARY 11.2. *Every stable game of perfect information has strong equilibrium m-points in supergame pure strategies.*

Finally, we remark that the discussion of § 6 applies unchanged to the non-cooperative case.

REFERENCES

1. R. J. Aumann, *Acceptable points in general cooperative n-person games*, Contributions to the Theory of Games IV, Annals of Mathematics Study 40, Princeton University Press, 1959, pp. 287–324.

2. C. Berge, *Théorie générale des Jeux a n personnes*, Mémorial des Sciences Mathématiques **138**, 1957.

3. ———, *Topological games with perfect information*, Contributions to the Theory of

Games III, Annals of Mathematics Study 39, Princeton University Press, 1957, pp. 165–178.

4. N. Dalkey, *Equivalence of information patterns and essentially determinate games*, Contributions to the Theory of Games II, Annals of Mathematics Study 28, Princeton University Press, 1953, pp. 217–244.

5. H. Everett, *Recursive games*, Contributions to the Theory of Games III, Annals of Mathematics Study 39, Princeton University Press, 1957, pp. 47–78.

6. D. Gale and F. M. Stewart, *Infinite games with perfect information*, Contributions to the Theory of Games II, Annals of Mathematics Study 28, Princeton University Press, 1953, pp. 245–266.

7. D. Gillette, *Stochastic games with zero stop probabilities*, Contributions to the Theory of Games III, Annals of Mathematics Study 39, Princeton University Press, 1957, pp. 179–187.

8. D. B. Gillies, *Some theorems on n-person games*, Ph. D. Thesis, Princeton University, 1953.

9. H. W. Kuhn, *Extensive games and the problem of information*, Contributions to the Theory of Games II, Annals of Mathematics Study 28, Princeton University Press, 1953, pp. 193–216.

10. ———, *Lectures on the theory of games*, issued as a report of the Logistics Research Project, Princeton University, 1952.

11. O. Morgenstern, and J. von Neumann, *Theory of games and economic behavior*, Princeton University Press, 1953.

12. J. F. Nash, *Non-cooperative games*, Ann. Math. **54** (1951), 286–295.

13. L. S. Shapley, *Stochastic games*, Proc. Nat. Acad. Sci. **39** (1953), 1095–1100.

14. P. Wolfe, *The strict determinateness of certain infinite games*, Pacific J. Math. **5** (1955), 891–897.

THE HEBREW UNIVERSITY, JERUSALEM, ISRAEL

FRACTIONAL POWERS OF CLOSED OPERATORS AND THE SEMIGROUPS GENERATED BY THEM

A. V. BALAKRISHNAN

Fractional powers of closed linear operators were first constructed by Bochner [2] and subsequently Feller [3], for the Laplacian operator. These constructions depend in an essential way on the fact that the Laplacian generates a semigroup. Phillips [6] in fact showed that these constructions (for positive indices less than one) were part of a more general one based on the Kolmogoroff-Levy representation theorem for infinitely divisible distributions. Finally, the present author constructed an operational calculus [1] for infinitesimal generators affording in particular a systematic study of the representation and properties of these operators.

In this paper we obtain a new construction for fractional powers in which it is not required that the base operator generate a semigroup; indeed its domain need not even be dense. Since the semigroup is not available, the previous constructions, based as they are on the Riemann-Liouville integrals, are not possible. However, we shall show, if the resolvent exists for $\lambda > 0$, and is $O(1/\lambda)$ for all λ, (a weaker condition will suffice at the origin, see § 7), then fractional powers may still be constructed, using an abstract version of the Stieltjes transform.

It is immediate that a closed operator A, for which $\| \lambda R(\lambda, A) \| < M$, does not necessarily generate a semigroup of any type. For a simple example, let the Banach space be $l_2(-\infty, \infty)$ and let A correspond to multiplying the nth coordinate by $n(1+i)$ say. Then for $\lambda > 0$, $\| R(\lambda, A) \| \leq \sqrt{2}/\lambda$, whereas A does not generate a semigroup, since no right-half plane is free of the spectra of A. An example in which A has no spectra in the right half plane and yet no semigroup is generated is given by Phillips [4].

The main motivation for the construction of fractional powers is the application to abstract Cauchy problems of the type:

$$\frac{d^n}{dt^n} u(t) \pm A u(t) = 0 \tag{1}$$

for $n \geq 2$, and it turns out that for the solution of (1.1), A itself need not be an infinitesimal generator. In this paper we study only the case $n = 2$, and we expect to consider the general case later.

The properties of newly constructed fractional powers are identical with those obtained in [1] for the case where A is a generator, with one important difference; namely that $-(-A)^{\alpha}$ generates semigroups in

Received May 15, 1959.

general only for $\alpha \leq 1/2$. On the other hand, these are the only exponents that matter in the application to Cauchy problems of the type (1.1).

1. Construction of fractional powers. Let A be a closed linear operator with domain and range in a Banach space X. Let each $\lambda > 0$ belong to its resolvent set and let (H_0)

$$\| \lambda R(\lambda, A) \| < M < \infty \ , \qquad \lambda > 0.$$

We have already noted that these conditions do not imply that A generate a semigroup of any kind. Let $x \in D(A)$. Then for $0 < \alpha < 1$, the integral

$$\int_0^\infty \lambda^{\alpha-1} R(\lambda, A) Ax d\lambda$$

where λ^α is taken positive, is convergent in the Bochner or absolute sense, since it can be expressed as

$$\int_0^1 \lambda^{\alpha-1}[R(\lambda, A)x - x] d\lambda + \int_1^\infty \lambda^{\alpha-1} R(\lambda, A) Ax d\lambda$$

and both of these intergals are absolutely convergent in view of (H_0). We define a linear operater J^α such that:

$$J^\alpha x = \frac{\sin \pi\alpha}{\pi} \int_0^\infty \lambda^{\alpha-1} R(\lambda, A)(-A) x d\lambda \ , \qquad 0 < \Re\alpha < 1. \tag{2.1}$$

For $0 < \Re\alpha < 2$, we define for each $x \in D(A^2)$

$$J^\alpha x = \frac{1}{\Gamma(1-\alpha)\Gamma(\alpha)} \int_0^\infty \lambda^{\alpha-1} \left[R(\lambda, A) - \frac{\lambda}{1+\lambda^2} \right] (-A) x d\lambda + \sin \pi\alpha/2(-A)x \ . \tag{2.2}$$

For $x \in D(A^2)$ definitions (2.1) and (2.2) coincide for overlapping ranges of α. More generally, for α such that $n - 1 < \Re\alpha < n$, we define, for $x \in D(A^n)$:

$$J^\alpha x = J^{\alpha-n+1}(-A)^{n-1} x \ . \tag{2.3}$$

For $n - 1 < \Re\alpha \leq n$, we define for $x \in D(A^{n+1})$

$$J^\alpha x = J^{\alpha-n+1}(-A)^{n-1} x \ . \tag{2.4}$$

These definitions are also evidently consistent. In (2.1), the principal value of λ^α is taken so that λ^α is positive for α positive.

We shall now obtain some properties of these operators which will qualify them to be recognized as fractional powers. First, if A does generate a semigroup, these coincide with the previous definitions in [1]. In particular, if we specialize $-A$ to be denote multiplication by the complex number s, non-negative, on the space of complex numbers, the definitions yield s^α, principal determination.

LEMMA 2.1. *The operators J^α can be extended to be closed linear.*

Proof. The operators J^α being linear, it is enough to show that for any sequence x_n, $x_n \in D(J^\alpha)$, converging to zero, the sequence $J^\alpha x_n$, if convergent, has zero limit also. To be specific, let $0 < \Re\alpha < 1$. Consider

$$y_n = R(\lambda, A)J^\alpha x_n \ ,$$

for fixed λ. Now $R(\lambda, A)x_n \in D(A) = D(J^\alpha)$ and it readily follow from (2.1) that

$$R(\lambda, A)J^\alpha x_n = J^\alpha R(\lambda, A)x_n \ . \tag{2.5}$$

Moreover, since $AR(\lambda, A)$ is bounded linear, so is $J^\alpha R(\lambda, A)$. Hence if limit $J^\alpha x_n = y$, we have from (2.5) that $R(\lambda, A)y$ is zero, hence y is zero also. The proof for other values of α is similar.

LEMMA 2.2. *For $x \in D(A^n)$, $J^\alpha x$ is analytic in α for $0 < \Re\alpha < n$.*

Proof. This may be directly verified from the definitions. In particular, it may be noted that for $x \in D(A^\infty)$, $J^\alpha x$ is analytic for $\Re\alpha > 0$.

For elements in certain domains larger than the ones in Lemma 2.2, we retain continuity. Thus

LEMMA 2.3. *Let $x \in D(A)$ and $Ax \in \overline{D(A)}$. Then for $0 < \Re\alpha < 1$ and α tending to 1 in a fixed sector about 1, $J^\alpha x \to -Ax$.*

Proof. We note that since $Ax \in \overline{D(A)}$, $\lambda R(\lambda, A)Ax \to Ax$ as $\lambda \to \infty$. Now

$$J^\alpha x - (-A)x = \int_0^\infty \frac{\sin \pi\alpha}{\pi}\lambda^{\alpha-1}\left[R(\lambda, A) - \frac{1}{\lambda + 1}\right](-A)xd\lambda \ ,$$

and the integral can be split into two parts, one from 0 to L and the other from L to infinity. For fixed L, the first part goes to zero since it is $0(|\sin \pi\alpha|)$. The second part in absolute value is

$$< L^{\sigma-1}\left|\frac{\sin \pi(1-\alpha)}{\pi(1-\sigma)}\right| \sup \|[\lambda R(\lambda, A) - I](-A)x\| + M\frac{L^{\sigma-2}}{|\sigma - 2|}\|Ax\| \ ,$$
$$\sigma = \Re\alpha.$$

and hence goes to zero also.

We do not in general have convergence to the identity at the origin, $\alpha = 0$. For if for some x, $Ax = 0$, $J^\alpha x$ is zero also. However, we can state:

LEMMA 2.4. *For any $x \in D(A)$ such that $\lambda R(\lambda, A)x \to 0$ as $\lambda \to 0$, $J^\alpha x \to x$ as $\alpha \to 0+$ in a fixed sector about* 0.

Proof. We have

$$J^\alpha x - x = -\frac{\sin \pi\alpha}{\pi}\int_0^\infty \left[R(\lambda, A)Ax + \frac{x}{\lambda+1}\right]\lambda^{\alpha-1}d\lambda$$
$$= -\frac{\sin \pi\alpha}{\pi}\int_0^\infty \lambda^\alpha \frac{R(\lambda, A)(x + Ax)}{\lambda+1}d\lambda$$

and the result follows from the second integral as a simple estimation shows.

LEMMA 2.5. *Let $x \in D(A^2)$. Then for $0 < \Re(\alpha+\beta) < 1$,*

$$J^{\alpha+\beta}x = J^\alpha J^\beta x\ . \tag{2.6}$$

Proof. For $x \in D(A^2)$, it is clear from (2.1) that $J^\beta x \in D(A) = D(J^\alpha)$. Moreover we have:

$$J^\alpha J^\beta x = \frac{\sin \pi\alpha}{\pi}\,\frac{\sin \pi\beta}{\pi}\int_0^\infty\int_0^\infty \lambda^{\beta-1}\mu^{\alpha-1}R(\lambda, A)R(\mu, A)A^2x d\lambda d\mu\ ,$$

where the double integral is absolutely convergent, and can be rewritten as

$$\frac{\sin \pi\alpha}{\pi}\,\frac{\sin \pi\beta}{\pi}\int_0^1(\sigma^{\beta-1}+\sigma^{\alpha-1})d\sigma\int_0^\infty R(\lambda\sigma, A)R(\lambda, A)A^2x\lambda^{\alpha+\beta-1}d\lambda\ .$$

Using the first resolvent equation, we have

$$R(\lambda\sigma, A)R(\lambda, A)A^2x = \frac{\sigma R(\lambda\sigma, A) - R(\lambda, A)}{1-\sigma}(-A)x$$

so that we have finally, after a change of variable:

$$J^\alpha J^\beta x = \text{Const.}\int_0^\infty \lambda^{\alpha+\beta-1}R(\lambda, A)(-A)xd\lambda\ .$$

where the constant

$$\frac{\sin \pi\alpha}{\pi}\,\frac{\sin \pi\beta}{\pi}\int_0^1\frac{(\sigma^{\beta-1}+\sigma^{\alpha-1}-\sigma^{-\alpha}-\sigma^{-\beta})}{(1-\sigma)}d\sigma$$

evaluates to $\sin \pi(\alpha + B)/\pi$, thus verifying (2.6), The semigroup property is readily extended to all exponents for $x \in D(A^\infty)$. For exponents less than some finite positive number this domain can be enlarged. First, however, we define:

$$(2.7) \qquad (-A)^\alpha = \text{Smallest closed extension of } J^\alpha .$$

We term this the principal value, even though we cannot, in general, claim that any other determination will differ from it by only a factor of $\exp i\pi k$ for integral k. For an example see [1]. On the other hand, the principal determination still enjoys a uniqueness property similar to the one obtained by Hille [4] for linear bounded operators. We state this as a Lemma:

LEMMA 2.5. *Let $x \in D(A^\infty)$. Then*

$$(2.8) \qquad \limsup_{|\eta| \to \infty} \frac{\log \| (-A)^{\sigma + i\eta} x \|}{|\eta|} < \pi , \qquad 0 < \sigma.$$

Moreover, it is not possible to find a determination of $(-A)^\alpha$ analytic in α for $x \in D(A^\infty)$, interpolating integral powers, and preserving the extremal property, other than the one given in (2.7).

Proof. A direct calculation yields (2.8). The uniqueness part follows as in [4, p. 496], using a classical result of F. Carlson.

We note that all these fractional powers are uniquely determined by their values on $D(A^n)$ each for a large enough n. [Actually on $D(A^\infty)$, the latter domain being dense in $\overline{D(A)}$. See § 3, Lemma 3.1.] The semigroup property (2.6) can be sharpened to read

$$(2.9) \qquad (-A)^{\alpha+\beta} = [(-A)^\alpha(-A)^\beta]_c ,$$

the right side being the smallest closure of $(-A)^\alpha(-A)^\beta$. This follows essentially from the fact that J^α (and hence $(-A)^\alpha$) commutes with $R(\lambda, A)$, as in [1].

3. Spectral theory. We next examine the spectra of the operators $(-A)^\alpha$. For this purpose we denote the second commutant of the set $\{R(\lambda, A), \lambda > 0\}$ by $\mathfrak{B}$. Then $\mathfrak{B}$ is a commutative Banach algebra with unit, and is strongly closed. Moreover, using the Gelfand theory, the linear multiplicative functionals over $\mathfrak{B}$ split into two classes $\mathfrak{M}_0$ and $\mathfrak{M}_1$. For any $m \in \mathfrak{M}_0$,

$$m[R(\lambda, A)] = 0$$

while for any $m \in \mathfrak{M}_1$, there is an $s \in \sigma[A]$, such that

$$m[R(\lambda, A)] = 1/(\lambda - s)$$

for every $\lambda > 0$.

For any function α of bounded variation on compact Borel sets of the half-line $[0, \infty)$ such that

$$\int_0^\infty \| R(\lambda, A)x \| \, d\,|\alpha| < \infty$$

it is clear that setting

$$\theta(\alpha)x = \int_0^\infty R(\lambda, A)x d\alpha$$

$\theta(\alpha) \in \mathfrak{B}$. We now collect some special functions we shall need in the sequel. Let μ be a complex number such that $\mu + |s|^\alpha \exp i\alpha\theta \neq 0$, for any $|s|$ and $-\pi \leq \theta \leq \pi$, and fixed α, $0 < \mathfrak{Re}\,\alpha < 1$. We can always find such a number for $|\alpha|$ sufficiently small, and we shall assume this is the case. Let

$$f(\lambda) = (1/\pi)(1/2i)[(\mu + \lambda^\alpha e^{-i\pi\alpha})^{-1} - (\mu + \lambda^\alpha e^{i\pi\alpha})^{-1}] \,. \tag{3.1}$$

Next, for each $t > 0$, let

$$g(\lambda; t) = (1/\pi)\,\mathfrak{Im}.[\exp(-t\lambda^\gamma \exp -i\pi\gamma)] \tag{3.2}$$

for some fixed γ, $0 < \gamma < 1/2$. Then

$$\theta(f) = \int_0^\infty R(\lambda, A) f(\lambda) dr$$

and

$$\theta(g(\lambda; t)) = S(t) = \int_0^\infty R(\lambda, A) g(\lambda; t) d\lambda$$

both belong to $\mathfrak{B}$, the integrals existing in the Bochner sense in the uniform topology. Moreover, for $m \in \mathfrak{M}_1$, with corresponding $-s \in \sigma(A)$, we have

$$m(\theta(f)) = \int_0^\infty \frac{f(\lambda)}{\lambda + s} d\lambda = [\mu + s^\alpha]^{-1} \,. \tag{3.3}$$

The integral on the right is of course the Stieltjes transform and exists for any s not on the negative real axis. [For the properties of Stieltjes transforms explicitly or implicitly used here see [7].] On the other hand, we note that the spectrum of A may be empty. Similarly,

$$m(S(t)) = \int_0^\infty \frac{g(\lambda; t)}{\lambda + s} d\lambda = \exp -ts^\gamma \,. \tag{3.4}$$

[Here and throughout, $s^\alpha = |s|^\alpha \exp i\alpha\theta$, $-\pi < \theta < \pi$].

LEMMA 3.1. *For every $x \in X$, $S(t)x \in D(A^\infty)$ for every $t > 0$. For*

$x \in \overline{D(A)}$, $\| S(t)x - x \| \to 0$ *as* $t \to 0+$.

Proof. We note that for every positive integer n,

$$\int_0^\infty \lambda^n | g(\lambda, t) | \| R(\lambda, A) \| d\lambda < \infty \tag{3.5}$$

and

$$\int_0^\infty \lambda^n g(\lambda; t) d\lambda = 0 \ , \ \text{including } n = 0. \tag{3.6}$$

For any $x \in X$,

$$AR(\lambda, A)x = \lambda R(\lambda, A)x - x \ .$$

Hence

$$\begin{aligned} AS(t)x &= \int_0^\infty R(\lambda, A) g(\lambda; t) d\lambda - x \int_0^\infty g(\lambda; t) d\lambda \\ &= \int_0^\infty R(\lambda, A) \lambda g(\lambda; t) x d\lambda \end{aligned}$$

using (3.5) and (3.6). In a similar manner we can extend this to any positive integer n.

$$A^n S(t)x = \int_0^\infty \lambda^n R(\lambda, A) x g(\lambda; t) d\lambda \ . \tag{3.7}$$

This shows that $S(t)x \in D(A^\infty)$. Next let $x \in D(A)$. Now

$$S(t)x - x = \int_0^\infty g(\lambda; t) \left[R(\lambda, A) - \frac{1}{\lambda} \right] x d\lambda \tag{3.8}$$

where we have used

$$\int_0^\infty \lambda^{-1} g(\lambda; t) d\lambda = 1 \ .$$

Since $x \in D(A)$, we can rewrite (3.8) as

$$S(t)x - x = \int_0^\infty g(\lambda, t) \lambda^{-1} R(\lambda, A) A x d\lambda$$

and the integral on the right is seen to go to zero with t. To see that the result is true for $x \in \overline{D(A)}$, we have only to note that (using H_0) $\| S(t) \| \leq M$.

Now, since $(-A)^\alpha$ is closed and $S(t)x \in D(A^\infty)$ for every x, $(-A)^\alpha S(t)$ is linear bounded and $\in \mathfrak{B}$. Actually, more is true. Thus:

LEMMA 3.2.

$$(-A)^\alpha S(t) = \int_0^\infty R(\lambda, \mathrm{A}) h(\lambda, t) d\lambda, \qquad 0 < \Re\alpha < 1, \tag{3.9}$$

where

$$h(\lambda, t) = \frac{1}{2\pi i}[e^{-i\pi\alpha}\lambda^{\alpha} \exp(-t\lambda^{\gamma} \exp -i\pi\gamma) - e^{i\pi\alpha}\lambda^{\alpha} \exp(-t\lambda^{\gamma} \exp +i\pi\gamma)] .$$

Proof. We note that for $m \in \mathfrak{M}_1$,

$$m\left[\int_0^\infty R(\lambda, A)h(\lambda, t)d\lambda\right] = \int_0^\infty \frac{h(\lambda, t)}{\lambda + s} d\lambda$$

and the Stieltjes transform on the right evaluates to

$$s^{\alpha} \exp(-ts^{\alpha}) .$$

While this is the same as $m[(-A)^{\alpha}S(t)]$, this does not necessarily constitute a proof of (3.9) since the radical of $\mathfrak{B}$ may be non-empty. However, a direct proof is possible. Thus we have

$$(-A)^{\alpha}S(t) = \int_0^\infty\int_0^\infty \frac{\sin \pi\alpha}{\pi} R(\lambda)(-A)R(\mu)\lambda^{\alpha-1}g(\mu, t)d\lambda d\mu$$

and changing variable of integration and using

$$R(\lambda\sigma, A)(-A)R(\lambda) = \frac{\sigma R(\lambda\sigma, A) - R(\lambda, A)}{1 - \sigma}$$

this can be written

$$\int_0^\infty R(\lambda, A)h(\lambda, t)d\lambda$$

where

$$h(\lambda, t) = g(\lambda, t)\int_0^1 \frac{\sigma^{-\alpha} - \sigma^{\alpha-1}}{(1-\sigma)} d\sigma + \int_0^1 \sigma^{-1}\frac{g(\lambda/\sigma) - g(\lambda\sigma)}{(1-\sigma)} d\sigma$$

which is readily verified to be the function required in (3.9).

LEMMA 3.3.

$$\theta(f)(-A)^{\alpha}S(t) = \int_0^\infty R(\lambda, A)q(\lambda, t)d\lambda \tag{3.10}$$

where

$$q(\lambda, t) = \frac{1}{2\pi i}\left[\frac{\lambda^{\alpha}e^{-i\pi\alpha} \exp(-t\lambda^{\gamma} \exp -i\pi\gamma)}{\mu + \lambda^{\alpha} \exp -i\pi\alpha} - \frac{\lambda^{\alpha}e^{i\pi\alpha} \exp(-t\lambda^{\gamma} \exp i\pi\gamma)}{\mu + \lambda^{\alpha} \exp i\pi\alpha}\right].$$

Proof. The Stieltjes tranform of $q(\lambda, t)$ is readily verified to be

$$\int_0^\infty \frac{q(\lambda, t)dt}{\lambda + s} = \frac{s^{\alpha} \exp -ts^{\gamma}}{\mu + s^{\alpha}} .$$

As in Lemma 3.2, this is not quite enough to prove (3.10). On the other hand, a direct proof may be given by double integration using the resolvent equation, and noting that

$$q(\lambda, t) = \int_0^1 (h(\lambda, t)(f(\lambda\sigma) - \sigma^{-1}f(\lambda/\sigma)) + f(\lambda)(h(\lambda\sigma, t) - \sigma^{-1}h(\lambda/\sigma, t)))\frac{d\sigma}{\sigma - 1}.$$

LEMMA 3.4.

$$\theta(f)S(t) = \int_0^\infty R(\lambda, A)r(\lambda, t)d\lambda \tag{3.11}$$

where

$$r(\lambda, t) = \frac{1}{2\pi i}\left[\frac{\exp(-t\lambda^\gamma \exp -i\pi\gamma)}{\mu + \lambda^\alpha \exp -i\pi\alpha} - \frac{\exp(-t\lambda^\gamma \exp i\pi\gamma)}{\mu + \lambda^\alpha \exp i\pi\alpha}\right].$$

Proof. The Stieltjes transform of $r(\lambda, t)$ is

$$\int_0^\infty \frac{r(\lambda, t)}{\lambda + s} d\lambda = \frac{\exp -ts^\gamma}{\mu + s^\alpha}.$$

As in Lemma 3.3, we can establish (3.11) by double integration, using the resolvent equation.

LEMMA 3.5. *Let* $x \in \overline{D(A)}$. *Then with* μ *as in Lemma* 3.1,

$$[\mu + (-A)^\alpha]\theta(f)x = x. \tag{3.12}$$

Proof. From the previous Lemmas it is immediate that

$$\mu\theta(f)S(t) + (-A)^\alpha\theta(f)S(t) = S(t)$$

for every $t > 0$. Let $x \in D(A)$. Since

$$[\mu + (-A)^\alpha]\theta(f)S(t)x = S(t)x$$

and by Lemma 3.2, $S(t)x \to x$ as $t \to 0$, (3.12) follows by letting $t \to 0$, and noting that $(-A)^\alpha$ is closed.

We are now ready to prove the spectral mapping theorem.

THEOREM 3.1. *Let* $D(A)$ *be dense in* X. *Then*

$$\sigma[(-A)^\alpha] = [\sigma(-A)]^\alpha, \qquad \Re\alpha > 0. \tag{3.13}$$

Proof. First let $|\alpha|$ be so small that we can find a μ as in Lemma 3.1. By Lemma 3.5, for $X \in D(A)$, (3.12) holds, and $D(A)$ being dense in X, continues to hold for any x. Since for $x \in D((-A)^\alpha)$, $\theta(f)(-A)^\alpha x = (-A)^\alpha\theta(f)x$, we see that $\theta(f)$ is a resolvent of $-(-A)^\alpha, \theta(f) = R(u, -(-A)^\alpha)$.

This is enough to prove that for these α, (3.13) holds. For, let δ be a number different from $+\mu$ such that $-\delta \neq [\sigma(-A)]^\alpha$. Then consider $[I + (\delta - \mu)R(\mu, -(-A)^\alpha)]$. This belongs to $\mathfrak{B}$, and

$$m[I + (\delta - \mu)R(\mu, -(-A)^\alpha)] = 1 \qquad m \in \mathfrak{M}_0$$
$$= \frac{\delta + s^\alpha}{\mu + s^\alpha} \qquad m \in \mathfrak{M}_1.$$

Hence this element has an inverse in $\mathfrak{B}$, and this is easily seen to be $R(\delta, -(-A)^\alpha)$. For other values of α, we note that for α such that $\rho[(-A)^\alpha]$ is not empty, we have a sharper version of (2.9):

$$(-A)^{n\alpha} = [(-A)^\alpha]^n \text{ for every integer } n.$$

Again, by the general spectral mapping theorem for closed operators with a non-empty resolvent set [4], we see that (3.13) holds for $n\alpha$. Finally, we note that for $\alpha = a + ib$, and $a^2 + b^2 < a$, it is always possible to find a μ such $\mu + s^\alpha \neq 0$ for all s not on the negative real axis. Hence (3.13) follows for all α, $\mathfrak{Re}\,\alpha > 0$.

4. **Some stability properties.** We shall call a property of A stable, if the same property holds for $-(-A)^\alpha$ at least for $0 < \alpha < 1$. We now state some stable properties of A.

4.1. Let A be linear bounded. Then $-(-A)^\alpha$ is also bounded for every α, $\mathfrak{Re}\,\alpha > 0$.

4.2. Let A^* be the adjoint of A. In view of hypothesis H_0 we can, following Phillips [4], define $A^{\odot}$ using his definition (Definition 14.3.1, p. 424). For $|\alpha|$ sufficiently small, we note that $-(-A)^\alpha$ also satisfies hypothesis H_0, so that we can also define $[(-A)^\alpha]^{\odot}$. We then have that

$$[(-A)^\alpha]^{\odot} = [(-A)^{\odot}]^\alpha .$$

4.3. Let A be the infinitesimal generator of a positive contraction semigroup. Then so is $-(-A)^\alpha$, for $0 < \alpha < 1$.

4.4. Let X be a Hilbert space. If A is dissipative, so is $-(-A)^\alpha$ for $0 < \alpha < 1$.

4.5. Let A be compact. Then so is $(-A)^\alpha$ for every α.

4.6. If for some $x \in X$, and s not on the negative real axis $Ax = -sx$, then $(-A)^\alpha x = s^\alpha x$ also.

5. **Generation of semigroups.** We now come to what is perhaps

the most important single property of these fractional powers (at least as far as applications to differential equations are concerned) viz., generation of semigroups. We shall show that for $0 < \alpha \leq 1/2$, $-(-)^{\alpha}$ generate strongly continuous semigroups, uniformly continuous away from the origin. We shall also obtain a representation for these semigroups in terms of $R(\lambda, A),$, $\lambda > 0$.

THEOREM 5.1. *Let $D(A)$ be dense, and let A satisfy H_0. Then for $0 < \alpha \leq 1/2$, $-(-A)^{\sigma}$ as defined by 2.7, generates a semigroup $S_{\alpha}(t)$, which is strongly continuous for $t \geq 0$, uniformly continuous for $t > 0$.*

Proof. First let $0 < \alpha < 1/2$. Let

$$S_{\alpha}(t) = \int_0^{\infty} R(\lambda, A) g(\lambda, t; \alpha) d\lambda \tag{5.1}$$

where

$$g(\lambda, t; \alpha) = (1/\pi) \sin (t\lambda^{\alpha} \sin \pi\alpha) \exp (-t\lambda^{\alpha} \cos \pi\alpha) .$$

Then (5.1) is a Bochner integral. The Stieltjes transform of $g(\lambda, t; \alpha)$ is, as we have noted before in § 3,

$$\int_0^{\infty} \frac{1}{\lambda + s} g(\lambda, t; \alpha) d\lambda = \exp -ts^{\alpha} .$$

However, this alone does not necessarily suffice to verify the semigroup property of $S_{\alpha}(t)$. A direct proof can be given, however, following the lines of Lemma 3.3. We shall next show that the infinitesimal generator of $S_{\alpha}(t)$ is $-(-A)^{\alpha}$, by showing that the resolvent of the latter for $\mu > 0$, is the Laplace transform of $S_{\alpha}(t)$. The Laplace transform can, further, be taken in the uniform topology, since $S_{\alpha}(t)$ is readily seen to be uniformly continuous for $t > 0$, by direct computation from (5.1). Now,

$$\begin{aligned} \int_0^{\infty} e^{-\mu t} S_{\alpha}(t) dt &= \int_0^{\infty} R(\lambda, A) \int_0^{\infty} e - {}^{\mu t} g(\lambda, t;) dt d\lambda \\ &= \int_0^{\infty} R(\lambda, A) \left[\frac{1}{\mu + \lambda^{\alpha} e^{-i\pi\alpha}} \right] d\lambda \end{aligned}$$

which as we have seen in Lemma 3.5, is the resolvent of $-(-A)^{\alpha}$ for $\mu > 0$. The strong continuity of $S_{\alpha}(t)$ has already been proved in Lemma 3.1.

Next let $\alpha = 1/2$. Let

$$S_{1/2}(t) = \int_0^{\rightarrow\infty} R(\lambda, A) \sin \sqrt{\lambda}\, t d\lambda , \tag{5.2}$$

where the integral is to be taken at infinity in the Cauchy sense. The convergence in the Cauchy sense can be seen as follows. By an integration by parts, we have, for each L,

$$\frac{1}{\pi}\int_0^L R(\lambda, A)\sin\sqrt{\lambda}\,t\,d\lambda = R(L, A)f(L, t) + \int_0^L R(\lambda, A)^2 f(\lambda, t)d\lambda\ ,$$

where

$$f(\lambda, t) = \frac{1}{\pi}\int_0^\lambda \sin\sqrt{\sigma}\,t\,d\sigma = \frac{1}{\pi}\left[\frac{2\sin\sqrt{\lambda}\,t}{t^2} - \frac{2\sqrt{\lambda}\cos\sqrt{\lambda}\,t}{t}\right].$$

Now the first term goes to zero as $L \to \infty$, and the second term is a convergent Bochner integral at infinity. Hence

$$S_{1/2}(t) = \frac{1}{\pi}\int_0^\infty R(\lambda, A)\sin\sqrt{\lambda}\,t\,d\lambda = \int_0^\infty R(\lambda, A)^2 f(\lambda, t)d\lambda\ . \tag{5.3}$$

Next, $S_{1/2}(t)$ is readily seen to be uniformly continuous for $t > 0$. A simple computation using (5.3) also shows that $\| S_{1/2}(t) \| \le$ Const. The semigroup property can be verified directly as before. Again, the Laplace transform of $S_{1/2}(t)$ is seen to be the resolvent of $-(-A)^{1/2}$. The strong continuity at the origin may be seen from:

$$\| S_{1/2}(t)x - x \| = \left\| \frac{1}{\pi}\int_0^\infty \left[R(\lambda, A) - \frac{1}{\lambda}\right] x \sin\sqrt{\lambda}\,t\,d\lambda \right\| .$$

For fixed L and t sufficiently small, the first term is $O(t)$. The second term is

$$\le \| Ax \| M \int_L^\infty \frac{|\sin\sqrt{\lambda}\,t|}{\lambda^2}\,d\lambda$$

and hence goes to zero also. Since $\| S_{1/2}(t) \|$ is bounded, strong continuity follows. This completes the proof of the theorem.

For values of $\alpha > 1/2$, $-(-A)^\alpha$ does not necessarily generate a semigroup of any type, as the following simple counter-example shows. Let $X = l_2(-\infty, \infty)$, and let A correspond to multiplying the nth coordinate by $(1 + i)n$. For $\alpha > 1/2$, no right half plane is free of spectra of $-(-A)^\alpha$, (as follows readily from Theorem 3.1) so that they cannot be generators of any semigroup.

We note in passing that (5.1) leads to a simple rigorous proof of Feller's expansion for the stable densities [3] for $0 < \alpha \le 1/2$. For, denoting the stable density by $F(\xi, t; \alpha)$, we have

$$F(\xi, t; \alpha) = \frac{1}{\pi}\int_0^\infty e^{-\xi\lambda}\,\mathfrak{Im}\,[\exp(-t\lambda^\alpha \exp -i\pi\alpha)]d\lambda$$

and expanding the second factor and interchanging integration and summation, which is obviously permissible, we have

$$F(\xi, t; \alpha) = -\frac{1}{\pi}\sum_{0}^{\infty}\frac{(-t)^n}{n!}\xi^{-n\alpha-1}\sin \pi\alpha n\Gamma(1+n\alpha)$$

which is Feller's expansion.

6. Application to abstract Cauchy problems. We shall next consider an application of the foregoing theory to a class of Cauchy problems. Indeed, this was the application which largely motivated the theory. This class may be considered a generalization of the abstract Cauchy problem of the type

$$u(t) = Au(t)$$

and is related to (though different from) the class treated by Hille [4, 5]. Thus we shall examine abstract Cauchy problems of the type

$$\frac{d^n u(t)}{dt^n} + (-1)^n Au(t) = 0 \,, \qquad n \geq 2. \tag{6.1}$$

More precisely, we shall phrase the problem as follows:

Given a complex Banach space X, and a closed linear operator A with domain dense in X and range in X, find a function $u(t)$ such that

(i) $u(t)$ is n times continuously differentiable in $[0, \infty)$

(ii) $u(t) \in D(A)$ for $t \geq 0$

(iii) $u(t)$ satisfies (6.1) for $t > 0$, and the initial conditions

$$\lim_{t\to 0} \| u^k(t) - u_k \| = 0 \text{ for prescribed } u_k,\ k = 0, 1, \cdots, r, \qquad r \leq n.$$

This is the reduced problem ('problème reduit') in the terminology of Hille [5, p. 42], $n - r$ being the defect ('default'). In addition to the existence of solutions [with some defect], we are of course interested in the uniqueness of the solutions. Now, if the operator A satisfies H_0, we are [by Theorem 5.1] assured of solutions for some suitable defect, but the question of uniqueness remains. On the other hand, if we do have unique solutions for some A, we would certainly like to know whether this implies that A satisfy H_0, since this would then characterize the solutions completely. In what follows we are concerned exclusively with the case $n = 2$. Our main result may be stated as follows:

THEOREM 6.1. *Let $n = 2$. Suppose A satisfies H_0. Then for each $u_0 \in D(A)$, the reduced problem with defect one has a solution such that*

$$\sup_{t\geq 0} \| u(t) \| < \infty \,. \tag{6.2}$$

Moreover, there is only one such solution, and it is given by

$$u(t) = S_{1/2}(t)u_0 \tag{6.3}$$

where the semigroup $S_{1/2}(t)$ *is uniformly continuous for* $t > 0$, *and has the representation* (5.3). *Further, for each* $t > 0$, *range of* $S_{1/2}(t) \subset D(A)$ *and* $S_{1/2}(t)$ *is analytic of class* $H(\phi, -\phi)$, $\phi > 0$. *Conversely, suppose for each* $u_0 \in D(A)$, *the reduced problem has a unique solution satisfying* (6.2) *for some* A. *Then setting* $u(t) = S(t)u_0$, *yields* $S(t)$ *as a strongly continuous semigroup. Suppose range of* $S(t) \subset D(A)$, *and* $S(t)$ *is analytic of class* $H(\phi, -\phi)$ *for some* $\phi > 0$. *Then* A *satisfies* H_0, *and* $S(t) = S_{1/2}(t)$ *as given by* (5.3).

We need some Lemmas.

LEMMA 6.1. *Let* A *satisfy* H_0. *Then* $R(\lambda, A)$ *exists in the sector* $-\phi < \arg \lambda < \phi$, *where* $(\operatorname{Tan} \phi)M = 1$.

Proof. Let $\varepsilon > 0$, and $\gamma_\varepsilon M = 1 - \varepsilon$. For any t, $|t| < \gamma_\varepsilon$, and $\sigma > 0$ consider $I + [(\sigma + it\sigma) - \sigma]R(\sigma, A)$. This element has an inverse given by the series

$$\sum_0^\infty [-it\sigma R(\sigma, A)]^n$$

and is clearly convergent, being majorized by the geometric series

$$\sum_0^\infty |t|^n M^n < \frac{1}{1 - \gamma_\varepsilon M} .$$

Moreover this also shows that the inverse is bounded in norm by $(1 - \gamma_\varepsilon M)^{-1}$. Now, by the first resolvent equation, it follows that

$$R(\sigma + it\sigma, A) = [I - [(\sigma + it\sigma) - \sigma]R(\sigma, A)]^{-1}R(\sigma, A)$$

and is in norm

$$\| R(\sigma + it\sigma, A) \| \leq \frac{M}{(1 - \gamma_\varepsilon M)\sigma} .$$

The assertion of the lemma follows readily from this.

LEMMA 6.2. *Suppose* A *satisfies hypothesis* H_0. *Then for each* α, $0 \leq \alpha \leq 1/2$, $\lambda \in \rho[-(-A)^\alpha]$ *for* $-\psi - \pi/2 < \arg \lambda < \pi/2 + \psi$ *for some* $\psi > 0$.

Proof. The proof is immediate from the spectral mapping theorem, Theorem 3.1, and Lemma 6.1 above. We can take $\psi = \pi/2 - \alpha(\pi - \varphi)$.

LEMMA 6.3. *Let* A *satisfy* H_0. *Then for* $0 < \alpha \leq 1/2$, *the semigroup* $S_\alpha(t)$, *defined by* 5.1 *and* (5.3) *is analytic, of class* $H(\phi_1, \phi_2)$ (Cf [4], p. 325, Definition 10.6.1), *with* $\phi_1 = -\alpha\varphi, \phi_2 = \alpha\varphi$, φ *being defined in Lemma* 6.1.

Proof. Let μ be such that $\Re\mathrm{e}\,\mu > 0$. Then for each α, $0 < \alpha \leq 1/2$, we know from Lemma 3.5 that $\mu \in \rho[-(-A)^\alpha]$ and

$$(6.4) \qquad R(\mu, -(-A)^\alpha) = \frac{1}{\pi}\int_0^\infty R(\lambda, A)\left[\frac{1}{\mu + \lambda^\alpha e^{-i\pi\alpha}} - \frac{1}{\mu + \lambda^\alpha e^{i\pi\alpha}}\right]d\lambda\,.$$

From this it readily follows that

$$(6.5) \qquad \| R(\mu, -(-A)^\alpha) | \leq \frac{M}{\Re\mathrm{e}\,\mu}$$

where the constant M is the same as in H_0. Next let $\varepsilon < 0$ be given. Then from Lemma 6.1 it follows that for $-\varphi + \varepsilon \leq \arg\lambda \leq -\varphi - \varepsilon$, there is a constant M_ε such that

$$(6.6) \qquad \| \lambda R(\lambda, A) \| \leq M_\varepsilon\,.$$

Let $-\varphi + \varepsilon \leq \psi \leq +\varphi - \varepsilon$. Let λ be >0. Then

$$R(\lambda, e^{-i\psi}A) = e^{i\psi}R(\lambda e^{i\psi}, A)\,.$$

Further it follows from (6.6) that

$$\| \lambda R(\lambda, e^{-i\psi}A) \| \leq M_\varepsilon\,,$$

so that $(A\exp - i\psi)$ satisfies H_0. Then we can define $(-e^{-i\psi}A)^\alpha$ using (2.7), and a simple contour integration shows that for $0 < \alpha < 1$,

$$(-e^{-i\psi}A)^\alpha = e^{-i\psi\alpha}(-A)^\alpha\,.$$

Moreover, applying (6.5) we know that for $\Re\mathrm{e}\,\mu > 0$,

$$\| R(\mu, -e^{-i\psi\alpha}(-A)^\alpha) \| \leq \frac{M_\varepsilon}{\Re\mathrm{e}\,\mu}\,.$$

But

$$\| R(\mu, -e^{-i\psi\alpha}(-A)^\alpha) \| = \| R(\mu e^{i\psi\alpha}, -(-A)^\alpha) \|\,.$$

Hence we obtain that for λ such that $\Re\mathrm{e}\,(\lambda e^{-i\psi\alpha}) > 0$,

$$\| R(\lambda, -(-A)^\alpha) \| \leq \frac{M_\varepsilon}{\Re\mathrm{e}\,(\lambda e^{-i\psi\alpha})}\,.$$

But this implies that the conditions for $S_\alpha(t)$ to be of class $H(\phi, \phi_2)$ as given by Hille ([4] p. 383, Theorem 12.8.1) are satisfied, thus proving lemma.

Proof of Theorem. We begin with the first part. Thus let A satisfy H_0. Setting $u(t) = S_{1/2}(t)u_0$, we get one solution satisfying (6.2). We shall now show that this solution is unique. Let $v(t)$ be a possibly different solution. By assumption $v'(t)$ is continuous at $t = 0$. Let v_1 be $v'(0)$. Let $w_0 = S(1)u_1$, and $w_1 = S(1)v_1$, where $S(t)$ is defined as in Lemma 3.1. Let $w(t) = S(1)v(t)$. Then $w(t)$, $w'(t) \in D(A^\infty)$ and

$$w''(t) + Aw(t) = 0 . \tag{6.7}$$

Let

$$L(\lambda, w) = \int_0^\infty e^{-\lambda t} w(t)dt , \qquad \lambda > 0.$$

Then since

$$w'(t) = w_1 + AS(1)\int_0^t u(\sigma)d\sigma$$

we get that $w'(t) \exp -\lambda t$ goes to zero at infinity, and hence by Laplace transforming (6.7) we have:

$$[\lambda^2 + A]L(\lambda, w) = \lambda w_0 + w_1 .$$

Since $B^2 = -A$, where we have written B for $-(-A)^{1/2}$, this can be rewritten

$$[\lambda I - B]L(\lambda, w) = R(\lambda w_0 + w_1) .$$

Since

$$R(\lambda, B) = \int_0^\infty e^{-\lambda t} S_{1/2}(t)dt$$

this yields

$$w'(t) + Bw(t) = S_{1/2}(t)Bw_0 + S_{1/2}(t)w_1 .$$

Hence

$$\frac{d}{dt}[S_{1/2}(t(w(t))] = S_{1/2}(2t)w_1 + S_{1/2}(2s)Bw_0$$

so that

$$\begin{aligned} S(t)w(t) &= w_0 + \int_0^t S_{1/2}(2t)w_1 dt + \int_0^t S_{1/2}(2t)Bw_0 dt \\ &= w_0 + \frac{1}{2}\int_0^{2t} S_{1/2}(\sigma)w_1 d\sigma + \frac{1}{2}(S_{1/2}(2t)w_0 - w_0) \\ &= S_{1/2}(2t)w_0 - \frac{1}{2}\int_0^{2t} S_{1/2}(\sigma)(w_1 - Bw_0)d\sigma . \end{aligned}$$

Hence

$$S_{1/2}(t)Bw(t) = S_{1/2}(2t)w_0 + \frac{1}{2}[S_{1/2}(2t) - I](w_1 - Bw_0) .$$

Now because of analyticity in a sector, zero does not belong to the point spectrum of $S_{1/2}(t)$ for any $t > 0$. Writing $S_{1/2}(-t)$ for the inverse, and using (6.2), we have that

$$\text{(6.8)} \qquad \operatorname{Sup} \| S_{1/2}(-t)(w_1 - Bw_0) \| < \infty .$$

We shall now show that for any element z such that $z \in \bigcap_t D(S_{1/2}(-t))$ and $\operatorname{Sup} \| S_{1/2}(-t)z \| < \infty$, $Bz = 0$. For this let

$$\text{(6.9)} \qquad F(\lambda) = \int_0^\infty e^{\lambda t} S_{1/2}(-t) z dt , \qquad \Re \lambda < 0 .$$

Then it is readily verified that

$$[\lambda I - B](-F(\lambda)) = z .$$

Now by Lemma 6.2, we know that $R(\lambda, B)$ exists for $-\psi - \pi/2 < \arg \lambda < \psi + \pi/2$ and hence there is a common domain where $-F(\lambda) = R(\lambda, B)z$. Hence for $\Re \lambda < 0$, $-F(\lambda)$ is the analytic continuation of $R(\lambda, B)$. Moreover, using the results of Lemma 6.3, it follows that $\| \lambda F(\lambda) \| \leq$ const., in a sector and $\| \lambda R(\lambda, B)z \| \leq$ const. in an intersecting sector, their union being the entire plane. Hence it follows that $\lambda R(\lambda, B)z = z$, since $\lambda R(\lambda, B)z \to z$ for $\lambda > 0$. Hence $Bz = 0$, as required. Hence $S_{1/2}(t)z = z$, so that

$$S(t)w(t) = S(2t)w_0 - t(w_1 - Bw_0) .$$

Hence using (6.2), $w_1 = Bw_0$. Since $u_0 \in D(A)$, using $S(1/n)$ in place of $S(1)$ and taking limits, we readily obtain that

$$v'(0) = Bu_0$$

and hence that $v(t) = S_{1/2}(t)u_0$. That range at $S_{1/2}(t) \subset D(A)$ follows from the representation (5.3).

We now proceed to the second part of the theorem. That $S(t)$ is a semigroup, strongly continuous at the origin with $\| S(t) \| \leq$ const., follows by arguments similar to the one used in Lemma 23.9.4 p. 627 of [4]. Let B be the infinitesimal generator of $S(t)$. Then for $x \in D(A)$, it is clear $B^2 x = -Ax$. For $x \in D(B^2)$ on the other hand, we note that since for $t > 0$, $S(t)x \in D(A)$, $B^2 S(t)x = -AS(t)x$, so that letting $t \to 0$, it follows that $B^2 x = -Ax$ also, since A is closed. Next we note that

$$\lambda^2 - A = (i\lambda - B)(i\lambda + B)$$

so that for $\lambda > 0$, $\lambda \in \rho(A)$ since $S(t)$ is analytic, and

$$R(\lambda, A) = R(i\sqrt{\lambda}, B) R(-i\sqrt{\lambda}, B) .$$

Again since $S(t)$ is analytic, of class $H(\phi, -\phi)$, it readily follows that $\| \lambda R(i\lambda, B) \| \leq$ Const., for λ real, from which we obtain that $\| \lambda R(\lambda, A) \| \leq$ Const., for $\lambda > 0$. Or, A satisfies hypothesis H_0. That $S(t) = S_{1/2}(t)$ is immediate from the first part of the theorem.

Additional properties of the solutions can of course be deduced from the representation (5.3). For instance, we note a rate of growth property: viz., for each $x \in X$, $\| S_{1/2}(t)x \| \to 0$ as $t \to \infty$, if $\lambda R(\lambda)x \to 0$ as $\lambda \to 0+$.

7. Some extensions. In this section we shall indicate some possible extensions of the foregoing theory.

The basic hypothesis H_0 concerning the operator A can be weakened. Thus suppose A satisfies H_1:

For each $\lambda > 0$, $\lambda \in \rho(A)$ and (H_1)

(i) $\| R(\lambda, A) \| = 0(1/\lambda)$ as $\lambda \to \infty$

(ii) $\int_0^1 \lambda^\sigma \| R(\lambda, A) \| d\lambda < \infty$, for some σ, $0 < \sigma < 1$.

Then it is possible to define $(-A)^\alpha$ for $\Re\alpha \geq \sigma$, still using definition (2.7). The hypothesis H_1 is satisfied for instance if A generates a semigroup $T(\xi)$ such that it is strongly continuous for $\xi \geq 0$, and

$$\int_1^\infty \| T(\xi) \| \xi^{-\sigma-1} d\xi < \infty \, . \tag{7.1}$$

The latter condition was used in [1], whereas hypothesis H_1 is similar to the one stated by Hille [4] (although of course the Hille condition is stronger since he considered only bounded operators). We shall show that for infinitesimal generators, H_1 and (7.1) are equivalent.

LEMMA 7.1. *Suppose A is the infinitesimal generator of a strongly continuous semigroup $T(\xi)$. Then if* (7.1) *holds, A satisfies H_1. Conversely, if A satisfies H_1, $T(\xi)$ satisfies* (7.1).

Proof. Suppose (7.1) holds. Then clearly

$$\lim_{\xi \to \infty} \frac{\log \| T(\xi) \|}{\xi} \leq 0$$

so that $R(\lambda, A)$ exists for $\lambda > 0$ and is of order $1/\lambda$ for $\lambda \to \infty$. Next

$$\begin{aligned}\int_0^1 \lambda^\sigma \| R(\lambda, A) \| d\lambda &\leq \int_0^1 \lambda^\sigma \int_0^\infty e^{-\lambda\xi} \| T(\xi) \| d\xi d\lambda \\ &\leq \text{const} + \int_1^\infty \| T(\xi) \| d\xi \int_0^1 e^{-\lambda\xi} \lambda^\sigma d\lambda \\ &\leq \text{const} \int_1^\infty \| T(\xi) \| \xi^{-\sigma-1} d\xi < \infty \, .\end{aligned}$$

To prove the converse we shall use some results from [1]. Let $S(\omega)$ be the B-algebra associated with $T(\xi)$ as in [1] $L(\omega)$ being the subspace of functions (Borel measurable) such that

$$\int_0^\infty \| T(\xi) \| \, | f(\xi) | \, d\xi < \infty .$$

Let A_w be the infinitesimal generator of the translation semigroup. Defining $(-A_w)^\alpha$ using (2.7) it follows for $f \in D(A_w)$, setting

$$(-A_w)^\alpha f = g ,$$

that

$$(-\lambda)^\alpha \phi(\lambda, f) = \phi(\lambda, g) , \qquad \Re\lambda \leq 0 ,$$

and hence that for any $\mu > 0$,

$$\frac{(-\lambda)^\alpha}{\mu - \lambda}$$

is a multiplier defined over all of $L(\omega)$. By the factor theorem (cf [**1**]), it follows that there is a corresponding function in $S(\omega)$ (actually in $L(\omega)$) and further an evaluation of this function shows that (7.1) is satisfied.

While for $\Re\alpha \geq \sigma$, we can define $(-A)^\alpha$, it is not possible to define, in general, $(-A)^\alpha$ for $\Re\lambda < \sigma$, at least not as a closed operator whose domain includes the domain of A. This may be seen as in the converse part of Lemma 7.1, using $\| T(\xi) \| = (1 + \xi)^\sigma$.

For A satisfying H_1 with $\sigma \leq 1/2$, $-(-A)^\alpha$ continues to generate strongly continuous semigroups for $\alpha \leq 1/2$, satisfying

$$\lim_{t\to\infty} \frac{\log \| S_\alpha(t) \|}{t} \leq 0 .$$

We do not know at present whether the semigroups are necessarily analytic of class $H(\phi_1, \phi_2)$.

References

1. A. V. Balakrishnan, *An operational calculus for infinitesimal generators of semigroups*, Trans. A.M.S., 1959.
2. S. Bochner, *Diffusion equation and stochastic processes*, Proc. Nat. Acad. Sciences, U.S.A., **35** (1949), 368–370.
3. W. Feller, *On a generalization of Marcel Riesz' potentials and the semigroups generated by them*, Communications du séminaire mathématique de l'université de Lund, tome supplimentaire (1952), dedie' a Marcel Riesz.
4. E. Hille and R. S. Phillips, *Functional analysis and semigroups*, Amer. Math. Soc. Colloq. Publ., Vol. **31**.
5. E. Hille, *Une generalization du problème de Cauchy*, Annales de L'Institut Fourier, Tome 4, 1952.
6. R. S. Phillips, *On the generation of semigroups of linear operators*, Pacific J. Math., (1952), 333–369.
7. D. V. Widder, *The Laplace transform*, Princeton University Press, 1946.

University of California at Los Angeles

BOUNDS FOR THE EIGENVALUES OF SOME VIBRATING SYSTEMS

DALLAS BANKS

1. Introduction. If a string with a non-negative integrable density $\rho(x)$, $x \in [a, b]$, is fixed at the points $x = a$ and $x = b$ under unit tension, then the natural frequencies of the string are determined by the eigenvalues of the boundary value problem

$$y'' + \mu\rho(x)y = 0 \ , \quad y(a) = y(b) = 0 \ . \tag{1.1}$$

Indicating their dependence on the function $\rho(x)$, we denote these eigenvalues by

$$\mu_1[\rho] < \mu_2[\rho] < \cdots \ . \tag{1.2}$$

We consider the set of all such strings which have the same total mass, $M = \int_a^b \rho(x)dx$. It is well known [5] that the eigenvalues (1.2) satisfy the inequality

$$\mu_n[\rho] \geq \frac{4n^2}{M(b-a)} \ , \qquad n = 1, 2, \cdots \ , \tag{1.3}$$

with equality when a mass of amount M/n is concentrated at the midpoint of each of n segments obtained by partitioning the string into n equal parts. If we place some restriction on $\rho(x)$ which prohibits such an accumulation of mass, then we can expect to get a larger bound than that of (1.3). M. G. Krein [8] has found that when $0 \leq \rho(x) \leq H < \infty$, the eigenvalues (1.2) satisfy the inequalities

$$\frac{4Hn^2}{M^2} X\left(\frac{M}{H(b-a)}\right) \leq \mu_n[\rho] \leq \frac{Hn^2\pi^2}{M^2} \ , \tag{1.4}$$

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

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

The inequality (1.4) is sharp and as $H \to \infty$, the lower bound approaches that of (1.3).

In this paper, we investigate the nature of the density functions for which the greatest lower bounds of the eigenvalues (1.2) are attained

Received June 8, 1959. This research was supported by the United States Air Force Office of Scientific Research.

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Carnegie Institute of Technology.

when other restrictions are placed on $\rho(x)$. For convenience, we may consider the eigenvalue problem

$$u'' + \lambda p(x)u = 0 \, , \quad u(0) = u(1) = 0 \, , \tag{1.5}$$

where $p(x) = (b - a)\rho[(b - a)x + a]$, $x \in [0, 1]$, instead of (1.1). We note that $\int_0^1 p(x)dx = M$. Denoting the eigenvalues of (1.5) by

$$\lambda_1[p] < \lambda_2[p] < \cdots ,$$

we see that

$$\lambda_n[p] = (b - a)\mu_n[\rho] \, . \tag{1.6}$$

We shall be concerned with determining bounds for the eigenvalues of the differential system (1.5) under various types of restrictions on $p(x)$. The principal restrictions we shall consider are:

(a) $p(x)$ is monotone in $[0, 1]$.

(b) $p(x)$ is convex, i. e., $p(x)$ satisfies the inequality

$$p(x) \leq \frac{x_2 - x}{x_2 - x_1} p(x_1) + \frac{x - x_1}{x_2 - x_1} p(x_2) \, .$$

where x_1 and x_2 are any two values such that $0 \leq x_1 \leq x_2 \leq 1$.

(c) $p(x)$ is concave, i. e., $-p(x)$ is convex.

These properties are invariant under the linear transformation used to obtain (1.5) from (1.1) so that $\rho(x)$, $x \in [a, b]$, will have the same properties as $p(x)$. Hence, no loss of generality is involved in using the system (1.5).

In § 2, 3 and 4, we obtain sharp lower bounds for $\lambda_1[p]$ in these three cases. For the higher eigenvalues we are able to obtain only general information concerning the density distributions which give the lower bounds. The ideas used also lead to results in the case of the more general Sturm-Liouville system

$$\begin{gathered} [r(x)u'] + [\lambda p(x) - q(x)]u = 0 \, , \\ u'(0) - h_0 u(0) = u'(1) + h_1 u(1) = 0 \, , \end{gathered} \tag{1.7}$$

where $p(x)$ and $q(x)$ are non-negative integrable functions, $r(x) \in C'$ is positive and $h_0 \geq 0$, $h_1 \geq 0$. In § 5, we obtain results under various assumptions about $p(x)$ and $q(x)$.

In § 6, we consider the vibrating rod of variable density and with clamped ends. The results we obtain are directly analogous to those obtained by Krein and to those derived in § 2, 3 and 4 for the first eigenvalue of (1.5). In § 7, we obtain results for the first eigenvalue of a membrane with fixed boundary in the case of bounded densities and in the case of concave densities on a convex domain.

The central idea used in finding lower bounds of $\lambda_1[p]$ is the following.

LEMMA 1.1. *If* $p(x)$ *of* (1.5) *can be expressed as*

$$p(x) = \int_0^1 K(x, t)g(t)df(t) \tag{1.8}$$

where

(i) $f(t)$ *is a monotone increasing bounded function,*
(ii) $g(t)$ *is non-negative and continuous,*
(iii) $K(x, t)$ *is non-negative and* $\int_0^1 K(x, t)dx = 1$,

then

$$\lambda_1[p] \geq \left[\int_0^1 p(x)dx\right]^{-1} \underset{t\in[0,1]}{\text{g. l. b.}}\ \lambda_1[K(x, t)] . \tag{1.9}$$

We use the fact that $\lambda_1[p]$ is the minimum of the Rayleigh quotient [4]

$$J(p, u) = \frac{\int_0^1 [u'(x)]^2 dx}{\int_0^1 p(x)[u(x)]^2 dx} , \tag{1.10}$$

where $u(x)$ ranges over all functions, with piecewise continuous first derivatives in $[0, 1]$, which satisfy the conditions $u(0) = u(1) = 0$. In view of (1.8), we have

$$\lambda_1^{-1}[p] = \max_u \frac{\int_0^1\left[\int_0^1 K(x, t)g(t)df(t)\right]u^2 dx}{\int_0^1 u'^2 dx} .$$

By the properties (i), (ii) and (iii) all terms are non-negative. Interchanging the order of integration, we find that

$$\lambda_1^{-1}[p] \leq \int_0^1 g(t)\left[\max_u \frac{\int_0^1 K(x, t)u^2 dx}{\int_0^1 u'^2 dx}\right]df(t) . \tag{1.11}$$

We note that

$$\lambda_1^{-1}[K(x, t)] = \max_u \frac{\int_0^1 K(x, t)u^2 dx}{\int_0^1 u'^2 dx} . \tag{1.12}$$

Hence, (1.11) and (1.12) yield

$$(1.13)\qquad \lambda_1^{-1}[p] \leq \int_0^1 g(t)df(t) \operatorname*{l.u.b.}_{t\in[0,1]} \lambda_1^{-1}[K(x,t)] .$$

From (1.8) and (iii), we have

$$\int_0^1 p(x)dx = \int_0^1 g(t)df(t) .$$

Hence (1.13) is equivalent to (1.9).

If the density $p(x)$ is normalized so that $\int_0^1 p(x)dx = 1$, then (1.13) reduces to

$$\lambda_1[p] \geq \operatorname*{g.l.b.}_{t\in[0,1]} \lambda_1[K(x,t)] .$$

To obtain results for the higher eigenvalues of (1.5) we use another approach.

LEMMA 1.2. *Let $p(x)$ and $q(x)$ be non-negative integrable functions defined for $x \in [a, b]$ and let $f(x)$ be non-negative, continuous and monotone increasing in $[a, b]$. Let $c \in (a, b)$ be such that $p(x) \geq q(x)$ for $x \in [a, c)$ and $p(x) \leq q(x)$ for $x \in (c, b]$. Then*

$$(1.14)\qquad \int_a^b p(x)dx = \int_a^b q(x)dx$$

implies that

$$(1.15)\qquad \int_a^b p(x)f(x)dx \leq \int_a^b q(x)f(x)dx .$$

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

By (1.14) we have

$$(1.16)\qquad \int_a^c [q(x) - p(x)]dx = \int_c^b [p(x) - q(x)]dx .$$

But $[p(x) - q(x)] \geq 0$ for $x \in [a, c)$ so that the generalized mean-value theorem gives

$$(1.17)\qquad \int_a^c [p(x) - q(x)]f(x)dx = f(x_1)\int_a^c [p(x) - q(x)]dx$$

for some $x_1 \in (a, c)$. Similarly, we have

$$(1.18)\qquad \int_c^b [q(x) - p(x)]f(x)dx = f(x_2)\int_c^b [q(x) - p(x)]dx$$

for some $x_2 \in (c, b)$. For a monotone increasing $f(x)$, we have $f(x_1) \leq f(x_2)$ so that (1.16), (1.17) and (1.18) imply

$$\int_a^c [p(x) - q(x)]f(x)dx \leq \int_c^b [q(x) - p(x)]f(x)dx .$$

Adding $\int_a^c q(x)f(x)dx$ and $\int_c^b p(x)f(x)dx$ to both sides, we obtain the desired result. If $f(x)$ is monotone decreasing, it is clear that the inequality has to reversed.

M. G. Krein has proved the following result which we will find useful. [8]

LEMMA 1.3. *Consider a family of density functions $p(x)$ on $[0, 1]$ such that $0 \leq p(x) \leq H < \infty$ and $\int_0^1 p(x)dx = M$. Let $\mu = \text{g.l.b.}\, \mu_1[p(x)]$ where the greatest lower bound is taken over this family. Then there is a function $p_0(x)$ in this family such that $\mu = \lambda_1[p_0]$.*

Krein's proof also holds for $\lambda_n[p]$, $n = 2, 3, \cdots$, and for the sum $\sum_{k=1}^n \lambda_k^{-1}[p]$.

2. Monotone densities. We first consider the system (1.5) when $p(x)$ is a monotone increasing function. We have the following result.

THEOREM 2.1. *Let $\lambda_1[p]$ be the lowest eigenvalue of a string of unit length with fixed end points whose density is an increasing function $p(x)$. Then*

$$\lambda_1[p]\int_0^1 p(x)dx \geq \lambda_0$$

where $\lambda_0 = 7.88\cdots$. The inequality is sharp and equality is attained for a string whose density is the step function

$$H(x, t_0) = \begin{cases} 0 \quad , & x \in [0, t_0] , \\ (1 - t_0)^{-1}, & x \in [t_0, 1] , \end{cases} \tag{2.1}$$

where $t_0 = 0.357\cdots$.

Since $p(x)$ is a positive, monotone increasing function, the Stieltjes integral

$$p(x) - p(0) = \int_0^x dp(t)$$

exists for $x \in [0, 1]$ except when $\lim_{x\to 1} p(x) = +\infty$. Even in this case the equality holds in a limiting sense. If we let

$$h(x, t) = \begin{cases} 0 , & 0 \leq x \leq t \leq 1 , \\ 1 , & 0 \leq t \leq x \leq 1 , \end{cases}$$

then we have

$$p(x) = \int_0^1 h(x, t)dp(t)$$

wherever $p(x)$ is continuous. Here we have replaced the original value of $p(x)$ at $x = 0$ by $p(0) = 0$; evidently this does not change our result. Since $p(x)$ is monotone, the set of discontinuity points is of zero measure. Hence, for purposes of integration, we may take the above equality to be true everywhere. If we let $H(x, t) = h(x, t)(1 - t)^{-1}$, we have

$$p(x) = \int_0^1 H(x, t)(1 - t)dp(t) . \tag{2.2}$$

By Lemma 1.1, we then have

$$\lambda_1[p] \geq \underset{t \in [0,1]}{\text{g.l.b.}} \lambda_1[H(x, t)] .$$

We find the values of t for which the greatest lower bound is attained by solving for $\lambda_1[H(x, t)]$ explicitly. If we solve (1.5) in the interval $[0, 1]$ with $p(x)$ replaced by $H(x, t)$ we find that over the interval $[t, 1]$, $u(x)$ must satisfy the differential system

$$u'' + \frac{\lambda}{1 - t}u = 0,\ tu'(t) = u(t),\ u(1) = 0 . \tag{2.3}$$

The eigenvalues of (2.3) will be equal to $\lambda_n[H(x, t)]$, $n = 1, 2, \cdots$. The eigenfunctions of (2.3) are

$$u_n(x) = \sin z_n x + \tan\left[\tan^{-1}\frac{tz_n}{1 - t} - \frac{tz_n}{1 - t}\right]\cos z_n t ,$$

$n = 1, 2, \cdots$, where $z_n = \sqrt{\lambda_n(1 - t)}$ is the nth positive roots of

$$\frac{\tan z}{z} = \frac{-t}{1 - t} . \tag{2.4}$$

Hence, the eigenvalues are

$$\lambda_n[H(x, t)] = \frac{z_n^2}{1 - t} , \qquad n = 1, 2, \cdots , \tag{2.5}$$

To find the value of t which minimizes $\lambda_n[H] = \lambda_n[H(x, t)]$ we replace (2.4) by

$$(1 - t)\sin z + tz\cos z = 0 . \tag{2.6}$$

This has the same positive zeros as (2.4). Since $\sin z$ and $z\cdot\cos z$ are positive for $0 < z < \pi/2$, (2.6) has no positive zeros in this interval for $t \in [0, 1]$. Over the interval $[\pi/2, \pi)$, $\sin z$ is positive while $z\cdot\cos z$ is negative for $z \in (\pi/2, \pi]$. Therefore, for $t \in (0, 1)$, the left side of (2.6) is positive at $z = \pi/2$ and negative at $z = \pi$. Hence, (2.6) has its first zero in the interval $[\pi/2, \pi]$. In fact, only the first one lies in this interval. For if we denote the left side of (2.6) by $F(t, z)$ then

$$F_z(t, z) = \cos z - tz \sin z$$

is negative so that for a particular value of $t \in (0, 1)$, $F(t, z)$ is monotone decreasing and hence has only one zero for $z \in [\pi/2, \pi]$. By (2.6)

$$\frac{dz_1}{dt} = \frac{-z_1}{1 - t + t^2 z_1^2} . \tag{2.7}$$

From (2.5), we have

$$\frac{d\lambda_1[H]}{dt} = \frac{z_1}{1-t}\left[2\frac{dz_1}{dt} + \frac{z_1}{1-t}\right] . \tag{2.8}$$

If we evaluate this at $t = 0$, we find $d\lambda_1/dt\,|_{t=0} = -\pi^2$. Furthermore, since $z_1(t)$ is finite, (2.4) implies that $\lambda_1[H] \to +\infty$ as $t \to 1$, so that $\lambda_1[H]$ has a minimum at some $t_0 \in (0, 1)$. Since we are considering only the first zero, $z_1(t)$, we will drop the subscript and write $z(t)$. At t_0 we must have $d\lambda_1/dt\,|_{t=t_0} = 0$ so that (2.8) implies

$$\left.\frac{dz}{dt}\right|_{t=t_0} = -\frac{1}{2}\frac{z(t_0)}{1 - t_0} .$$

From (2.7) we find that $z(t_0) = z'$ must satisfy

$$\frac{-z'}{1 - t_0 + t_0^2 z'^2} = -\frac{1}{2}\frac{z'}{1 - t_0} .$$

If we solve for $-t_0(1 - t_0)^{-1}$, it follows from (2.4) that

$$z' \tan z' = -t_0^{-1} .$$

Eliminating t_0 between this and (2.4) we find that z' must satisfy

$$\tan 2z' = 2z' . \tag{2.9}$$

The first zero of this equation is $z' \cong 2.25$. Hence, from (2.4) we find $t_0 = .357\cdots$ Now (2.9) has only one zero for $z \in [\pi/2, \pi]$ so that $\lambda_1[H]$ has only one relative extremum for $t \in (0, 1)$. But we know there is a minimum so that t_0 must be the value of t which minimizes $\lambda_1[H]$. From (2.5) we find this minimum to be approximately 7.88.

It does not appear possible to obtain lower bounds for the higher eigenvalues by the exclusive use of Lemma 1.1. We can, however, obtain a bound for the sum $\sum_{k=1}^{n}(\lambda_k^{-1}[p])$ with the help of a theorem of Courant [5], according to which

$$\sum_{k=1}^{n} \frac{\int_0^1 p(x) v_k^2 dx}{\int_0^1 v_k'^2 dx}$$

has the maximum $\sum_{k=1}^{n}(\lambda_k^{-1}[p])$ if the v_k, $k = 1, 2, \cdots, n$, range over all systems of mutually orthogonal functions with piecewise continuous derivatives in $[0, 1]$ such that $v_k(0) = v_k(1) = 0$.

THEOREM 2.2. *Let $\lambda_k[p]$, $k = 1, 2, \cdots, n$, be the first n eigenvalues of a string of unit length with fixed ends whose density is an increasing function $p(x)$. If $\lambda_k[H(x, t)]$, $k = 1, 2, \cdots, n$, are the first n eigenvalues of a string whose density is the step function defined by* (2.1), *then*

$$\left[\int_0^1 p(x)dx\right]^{-1} \sum_{k=1}^{n} \lambda_k^{-1}[p] \leq \sum_{k=1}^{n} \lambda_k^{-1}[H(x, t_0)]$$

where $H(x, t_0)$ is the step function (2.1) *and t_0 is a suitable value in* $[0, 1]$.

Evidently, the inequality is sharp.

By Courant's theorem, we have for the eigenvalues of the system (1.5)

$$\sum_{k=1}^{n} \lambda_k^{-1} = \max_{v_k} \left[\sum_{k=1}^{n} \frac{\int_0^1 p(x)v_k^2 dx}{\int_0^1 v_k'^2 dx} \right],$$

for suitable v_k. Using (2.2) and changing the order of integration we have

$$\sum_{k=1}^{n} \lambda_k^{-1} = \max_{v_k} \left\{ \int_0^1 (1-t) \left[\sum_{k=1}^{n} \frac{\int_0^1 H(x, t)v_k^2 dx}{\int_0^1 v_k'^2 dx} \right] dp(t) \right\}.$$

Since all the factors are positive we find

$$\sum_{k=1}^{n} \lambda_k^{-1} \leq \int_0^1 (1-t) \left[\max_{v_k} \sum_{k=1}^{n} \frac{\int_0^1 H(x, t)v_k^2 dx}{\int_0^1 v_k'^2 dx} \right] dp(t).$$

Again by Courant's theorem, we get

$$\sum_{k=1}^{n} \lambda_k^{-1} \leq \int_0^1 (1-t) \left\{ \sum_{k=1}^{n} [\lambda_k[H(x, t]^{-1} \right\} dp(t),$$

so that, as in the proof of Lemma 1.1, we have

$$\sum_{k=1}^{n} \lambda_k^{-1} \leq \underset{t \in [0,1]}{\text{l.u.b.}} \sum_{k=1}^{n} [\lambda_k[H(x, t)]]^{-1}.$$

We found in the proof of Theorem 2.1, that $\lambda_1[H(x, t)]$ becomes infinite as $t \to 1$. Hence, $\sum_{k=1}^{n}[\lambda_k[H(x, t)]]^{-1}$ approaches zero as $t \to 1$. Thus, there is a number $\delta > 0$ such that $t \geq 1 - \delta$ implies that

$$\text{l.u.b.} \sum_{k=1}^{n} [\lambda_k[H]]^{-1} \geq \sum_{k=1}^{n} [\lambda_k[H(x, t)]]^{-1} .$$

But for $t \in [0, 1 - \delta]$, $H(x, t)$ is uniformly bounded. Hence, Lemma 1.3 implies the desired result.

While it seems to be difficult to obtain exact numerical bounds for the higher eigenvalues of the system (1.5), when $p(x)$ is monotone, it is possible to give a geometric characterization of the function $p(x)$ which corresponds to the minimum value of $\lambda_n[p]$.

THEOREM 2.3. *Let $\lambda_n[p]$ be the* nth *eigenvalue of a string of unit length with fixed ends whose density is a monotone increasing function $p(x)$. Then there is a string with the same total mass whose density is a monotone increasing step function $q(x)$ with at most n jumps such that*

$$\lambda_n[p] \geq \lambda_n[q] , \tag{2.10}$$

where $\lambda_n[q]$ is the nth *eigenvalue of the string with density $q(x)$.*

Let $u_n(x)$ be the nth eigenfunction of (1.5) corresponding to $\lambda_n[p]$. It is well known that $u_n(x)$ has exactly $(n + 1)$ zeros in the closed interval $[0, 1]$. We denote these zeros by

$$x_0 = 0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = 1 .$$

In each open subinterval (x_k, x_{k+1}), $u_n(x)$ has only one maximum or one minimum so that $u_n^2(x)$ has a maximum there. We denote these n maximum points by $\bar{x}_1 < \bar{x}_2 < \cdots < \bar{x}_n$.

We will show presently that it is possible to construct a function $q(x)$ in such a way that over each of the intervals $(x_{k-1}, \bar{x}_k)$, $(\bar{x}_k, x_k)$, $k = 1, 2, \cdots, n$, $q(x)$ and $p(x)$ are related as indicated in Lemma 1.2. By Lemma 1.2, we will then have

$$\int_{x_{k-1}}^{\bar{x}_k} p(x)u_n^2(x)dx \leq \int_{x_{k-1}}^{\bar{x}_k} q(x)u_n^2(x)dx , \qquad k = 1, 2, \cdots, n ,$$

and

$$\int_{\bar{x}_k}^{x_k} p(x)u_n^2(x)\,dx \leq \int_{\bar{x}_k}^{x_k} q(x)u_n^2(x)dx , \qquad k = 1, 2, \cdots, n .$$

Upon adding these inequalities, we get

$$\int_{x_{k-1}}^{x_k} p(x)u_n^2(x)\,dx \leq \int_{x_{k-1}}^{x_k} q(x)u_n^2(x)dx , \qquad k = 1, 2, \cdots, n . \tag{2.11}$$

If we fix the string at the nodal points x_k, $k = 0, 2, \cdots, n$, then it is known [5] that

$$\lambda_n = \frac{\int_{x_{k-1}}^{x_k} u_n'^2 dx}{\int_{x_{k-1}}^{x_k} p(x) u_n^2 dx}\,, \qquad k = 1, 2, \cdots, n\,.$$

By (2.11), we have

$$\lambda_n \geq \frac{\int_{x_{k-1}}^{x_k} u_n'^2 dx}{\int_{x_{k-1}}^{x_k} q(x) u_n^2 dx} \geq \min_{u \in C'} \frac{\int_{x_{k-1}}^{x_k} u'^2 dx}{\int_{x_{k-1}}^{x_k} q(x) u^2 dx}$$

where $u(x_k) = 0, k = 1, 2, \cdots, n$. In particular, λ_n must satisfy

$$\lambda_n \geq \max_{1 \leq k \leq n} \left[\min_{u \in C'} \frac{\int_{x_{k-1}}^{x_k} u'^2 dx}{\int_{x_{k-1}}^{x_k} q(x) u^2 dx} \right].$$

But the quantity on the right is greater than the nth eigenvalue $\lambda_n[q]$ of a string with density $q(x)$ so (2.10) will hold (See [5]).

It remains to be shown that there exists a function $q(x)$ of the desired form.

We first consider the intervals $(\bar{x}_k, x_k]$, $k = 1, 2, \cdots, n$. Here we set

$$q(x) = (x_k - \bar{x}_k)^{-1} \int_{\bar{x}_k}^{x_k} p(x) dx = a_k\,, \qquad k = 1, 2, \cdots, n\,.$$

Since $p(x)$ is monotone increasing, the hypothesis of lemma (1.2) is obviously satisfied. For the intervals $(x_{k-1}, \bar{x}_k)$, we choose a point $t_k \in [x_{k-1}, \bar{x}_k]$ such that

$$(\bar{x}_k - t_k)(a_k - a_{k-1}) = \int_{x_{k-1}}^{\bar{x}_k} p(x) dx - a_{k-1}(\bar{x}_k - x_{k-1})$$

$k = 1, 2, \cdots, n$, where we take $a_0 = 0$ and set

$$q(x) = \begin{cases} a_{k-1}, & x \in (x_{k-1}, t_k]\,, \\ a_k\,, & x \in (t_k, \bar{x}_k]\,, \end{cases}$$

$k = 1, 2, \cdots, n$. By the definition of the a_k's, $p(x) \geq a_{k-1}$, $x \in [x_{k-1}, t_k)$ and $p(x) \leq a_k$ for $x \in [t_k, \bar{x}_k)$ so that the hypothesis of Lemma 1.2 is again satisfied. Thus, the function $q(x)$, $x \in [0, 1]$ may be taken to be

$$q(x) = a_k, x \in (t_k, t_{k+1})\,, \qquad k = 0, 1, 2, \cdots, n\,,$$

where we let $t_0 = 0$ and $t_{n+1} = 1$. This proves the theorem.

3. Convex densities. We now turn to the consideration of (1.5) in the case where $p(x)$ is convex. We have the following theorem.

THEOREM 3.1. *Let $\lambda_1[p]$ be the first eigenvalue of a string of unit length with fixed ends whose density is a continuous convex function $p(x)$. Then*

$$\lambda_1[p]\int_0^1 p(x)dx \geq \lambda_0$$

where $\lambda_0 = 9.397\cdots$. The inequality is sharp and equality is attained for a string whose density is the piecewise linear function

$$G(x, t_0) = \begin{cases} 0 & , \quad x \in [0, t_0] , \\ 2\dfrac{x - t_0}{(1 - t_0)^2} , & x \in [t_0, 1] , \end{cases} \tag{3.1}$$

where $t_0 = 0.104\cdots$.

We first note that any positive convex density $p(x)$ may be written as the sum $p_1(x) + p_2(x)$ where $p_1(x)$ is a positive monotone increasing convex function and $p_2(x)$ is a positive monotone decreasing convex function. In particular, we may define $p_1(x)$ such that $p_1(0) = 0$ and $p_2(x)$ such that $p_2(1) = 0$. If $\xi \in [0, 1]$ is a minimum point of $p(x)$, it is easily confirmed that the functions

$$p_1(x) = \begin{cases} p(\xi)x & , \quad x \in [0, \xi] , \\ p(x) - p(\xi)(1 - x) , & x \in [\xi, 1] , \end{cases}$$

and

$$p_2(x) = \begin{cases} p(x) - p(\xi)x , & x \in [0, \xi] , \\ p(\xi)(1 - x) \quad , & x \in [\xi, 1] , \end{cases}$$

have the required properties.

We may thus express $p(x)$ as $p(x) = \alpha\bar{p}_1(x) + \beta\bar{p}_2(x)$, where $M\alpha = \int_0^1 p_1(x)dx$, $M\beta = \int_0^1 p_2(x)dx$, $\bar{p}_1(x) = p_1(x)/\alpha$ and $\bar{p}_2(x) = p_2(x)/\beta$. From the Rayleigh quotient we then have

$$\begin{aligned} \lambda_1^{-1}[p] &= \max_{u\in C'} [J(p, u)]^{-1} \\ &\leq \max_{u\in C'} \alpha[J(\bar{p}_1, u)]^{-1} + \max_{n\in C'} \beta[J(\bar{p}_2\, u)]^{-1} . \end{aligned} \tag{3.2}$$

Let $\lambda_1[\bar{p}_1]$ and $\lambda_1[\bar{p}_2]$ be the first eigenvalues of strings with fixed end points and densities $\bar{p}_1(x)$ and $\bar{p}_2(x) (x \in [0, 1])$ respectively. Then, from (3.2),

$$\lambda_1[p] \geq \min_{(1, 2)} (\lambda_1[\bar{p}_1], \lambda_1[\bar{p}]) .$$

Because of the symmetry of the boundary conditions $u(0)=u(1)=0$, it is obvious that the bound for $\lambda_1[p]$ in the case of monotone increasing

$p(x)$ is the same as that for monotone decreasing $p(x)$. Hence, we need only consider monotone increasing functions. Furthermore, as shown above, we may assume that $p(0) = 0$.

Now set

$$g(x, t) = \begin{cases} 0 & , \quad x \in [0, t] , \\ x - t , & x \in [t, 1] , \end{cases}$$

where $t \in [0, 1]$. We first assume that the increasing function $p(x)$ is bounded and that the left-hand derivative $p'_-(x)$ is bounded for $x \in [0, 1]$. It then follows from integration by parts and the fact that

$$\int_0^x p'_-(t)dt = p(x) - p(0) ,$$

for such a function $p(x)$[**13**], that

$$p(x) = p'_-(0)x + \int_0^1 g(x, t)dp'_-(t) .$$

If we set $G(x, t) = (2/(1 - t)^2)g(x, t)$, we have

$$p(x) = \int_0^1 G(x, t)[1/2(1 - t)^2]dp'_-(t) .$$

Here we have replaced the original value of $p'_-(x)$ at $x = 0$, by $p'_-(0) = 0$; evidently this does not change our result. By Lemma 1.1, we then have

$$\lambda_1[p] \geq \underset{t \in [0, 1]}{\text{g. l. b.}} \lambda_1[G(x, t)] \left[\int_0^1 p(x) dx \right]^{-1} . \tag{3.3}$$

If $p(x)$ or its left derivative is not bounded in $[0, 1]$, we may consider the system

$$u'' + \lambda p(x) u = 0, u(0) = u(1 - \varepsilon) = 0 \tag{3.4}$$

where $\varepsilon > 0$ is arbitrarily small. In the interval $[0, 1-\varepsilon]$, $p(x)$ and $p'_-(x)$ are bounded, so transforming the system (3.4) to the unit interval we find, by (3.3), that

$$(1 - \varepsilon)\lambda_1[p]_\varepsilon \int_0^{1-\varepsilon} p(x)dx \geq \underset{t \in [0, 1]}{g. \text{l. b.}} \lambda_1[G(x, t)] ,$$

where $\lambda_1[p]_\varepsilon$ is the first eigenvalue of (3.4). Since ε is arbitrary and the eigenvalues are continuous functions of the length of the interval, it follows that (3.3) holds for any increasing convex density $p(x)$.

To find the values of t for which the greatest lower bound of $\lambda_1[p]$ is attained, we employ a procedure similar to that used in a corresponding situation in the proof of Theorem 2.1. Our problem is then seen to reduce to the computation of the lowest eigenvalue of the system

$$(3.5)\qquad u''(x) + \frac{2\lambda}{(1-t)^2}(x-t)u(x) = 0,\ tv'(t) = v(t),\ v(1) = 0\ .$$

The eigenfnnctions of (3.5) are [7]

$$u_n(x) = \sqrt{x-t}\left[J_{1/3}\left(z_n\left(\frac{x-t}{1-t}\right)^{3/2}\right) + \frac{t\Gamma(2/3)(6\lambda_n)^{1/3}}{(1-t)^{2/3}\Gamma(1/3)}J_{-1/3}\left(z_n\left(\frac{x-t}{1-t}\right)^{3/2}\right)\right]$$

where $J_{\pm(1/3)}(y)$ are Bessel functions of order $\pm 1/3$ and $z_n = (2/3)\sqrt{2\lambda}(1-t)^{-1}$ is the nth positive root of

$$(3.6)\qquad (1-t)\Gamma\left(\frac{4}{3}\right)\left(\frac{z}{2}\right)^{-1/3}J_{1/3}(z) + t\Gamma\left(\frac{2}{3}\right)\left(\frac{z}{2}\right)^{1/3}J_{-1/3}(z) = 0\ .$$

Hence, the eigenvalues of (3.5) are

$$(3.7)\qquad \lambda_n[G(x,t)] = \frac{9z_n^2}{8(1-t)}\ , \qquad n = 1, 2, \cdots\ .$$

We denote the left side of (3.6) by $F(t, z)$. To find the value t_0 which minimizes $\lambda_1[G]$, we must investigate some properties of this function. The first positive zeros of $J_{1/3}(z)$ and $J_{-1/3}(z)$ are $\xi_0 = 1.87$ and $\xi_1 = 2.90$, respectively [9]. In $(0, \xi_0)$, $F(0, z)$ and $F(1, z)$ are positive. Hence, $F(t, z)$ has no zeros in this interval for $t \in (0, 1)$. In $[\xi_0, \xi_1)$, $F(0, z)$ is positive, while $F(1, z)$ is negative in $(\xi_0, \xi_1]$. Accordingly, for $t \in (0, 1)$, $F(t, \xi_0)$ is negative while $F(t, \xi_1)$ is positive. Hence, $F(t, z)$ has its first zero in (ξ_0, ξ_1). Furthermore, there is only one zero in this interval for a given value of t, since

$$-F_z(t, z) = (1-t)\Gamma(4/3)(z/2)^{-1/3}J_{4/3}(z) + t\Gamma(2/3)(z/2)^{1/3}J_{2/3}(z)\ ,$$

and it is known that each of the terms on the right side is positive for $z \in (\xi_0, \xi_1)$. Thus, for a given t, $F(t, z)$ is monotone decreasing over this interval and thus has only one zero there.

Since we are considering only the first zero $z_1(t)$, we will drop the subscript and write $z(t)$. We have

$$(3.8)\qquad \frac{dz}{dt} = \frac{\Gamma(2/3)(z/2)^{1/3}J_{-1/3}(z) - \Gamma(4/3)(z/2)^{-1/3}J_{1/3}(z)}{(1-t)\Gamma(4/3)(z/2)^{-1/3}J_{4/3}(z) + t\Gamma(2/3)(z/2)^{1/3}J_{2/3}(z)}\ ,$$

and by (3.7),

$$(3.9)\qquad \frac{d\lambda_1[G]}{dt} = \frac{9z}{8(1-t)}\left[2\frac{dz}{dt} + \frac{z}{1-t}\right].$$

If we evaluate this at $t = 0$, we find $(d\lambda_1[G])/dt|_{t=0} = -1.38$. Furthermore, since $z(t)$ is finite for all $t \in [0, 1]$, (3.7) implies that for $t \to 1$, $\lambda_1[G] \to +\infty$ so that $\lambda_1[G]$ has a minimum at some value $t_0 \in (0, 1)$. At an extremum t' of $\lambda_1[G]$ we must have $(d\lambda_1[G])/dt|_{t'} = 0$, so that (3.9) implies

$$\left.\frac{dz}{dt}\right|_{t=t'} = -\frac{1}{2}\frac{z(t')}{1-t'} . \tag{3.10}$$

If we substitute (3.8) into (3.10) and then eliminate t' between this result and (3.5), we get

$$\Gamma(4/3)(z'/2)^{2/3}J_{1/3}(z')J_{2/3}(z') = \Gamma(2/3)(z'/2)^{1/3}[J_{1/3}(z')]^2$$
$$-\Gamma(4/3)(z'/2)^{-1/3}J_{1/3}(z')J_{-1/3}(z') - (z'/2)^{2/3}J_{4/3}(z')J_{-1/3}(z')] ,$$

where $z' = z(t')$. If we use the relations [7],

$$J_{4/3}(x) = (1/3)(x/2)^{-1}J_{1/3}(x) - J_{-2/3}(x) ,$$
$$J_{1/3}(x)J_{2/3}(x) + J_{-1/3}(x)J_{-2/3}(x) = \frac{2\sin\pi/3}{\pi x}$$

and

$$\Gamma\left(\frac{2}{3}\right)\Gamma\left(\frac{1}{3}\right) = \pi/\sin\pi/3 ,$$

we finally arrive at the equation

$$\left[\Gamma\left(\frac{2}{3}\right)\left(\frac{z'}{2}\right)^{1/3}J_{-1/3}(z')\right]\left[\Gamma\left(\frac{2}{3}\right)\left(\frac{z'}{2}\right)^{1/3}J_{-1/3}(z') - \frac{2}{3}\Gamma\left(\frac{4}{3}\right)\left(\frac{z'}{2}\right)^{-1/3}J_{1/3}(z')\right] - \frac{1}{3} = 0 . \tag{3.11}$$

We show that this equation has only one zero in (ξ_0, ξ_1), i.e., that $\lambda_1[G]$ has only one extremum for $t \in (0, 1)$. This will be the case if the derivative of the expression on the left of (3.11) is of one sign for $z' \in (\xi_0, \xi_1)$. The following statements relate to this interval.

(i) $[\Gamma(2/3)(z'/2)^{1/3}J_{-1/3}(z')]$, is negative and decreasing.

(ii) $-(2/3)\Gamma(4/3)(z'/2)^{-1/3}J_{1/3}(z')$, is negative so that the quantity in the second bracket also is negative.

(iii) The second bracket is decreasing.

The last statement requires verification. The derivative of the quantity in question is

$$-\left[\Gamma\left(\frac{2}{3}\right)\left(\frac{z'}{2}\right)^{1/3}J_{2/3}(z') - (2/3)\Gamma\left(\frac{4}{3}\right)\left(\frac{z'}{2}\right)^{-1/3}J_{4/3}(z')\right] .$$

(iii) will be verified if we show that

$$\frac{3\Gamma(2/3)}{2\Gamma(4/3)}(z'/2)^{2/3}J_{2/3}(z') - J_{4/3}(z') > 0 , \tag{3.12}$$

for $z \in (\xi_0, \xi_1)$. Since

$$\frac{d}{dz}(z/2)^{2/3}J_{2/3}(z) = (z/2)^{2/3}J_{-1/3}(z) < 0 \ ,$$

we have

$$(z'/2)^{2/3}J_{2/3}(z') \geq (\xi_1/2)^{2/3}J_{2/3}(\xi_1) \ .$$

Furthermore $\max_{\xi_0 \leq z \leq \xi_1} J_{4/3}(z) < .52$. Evaluation of (3.12) then gives the desired result.

We thus conclude that the left-hand side of (3.11) is increasing and hence, that $\lambda_1[G]$ has only one extremum for $t \in (0, 1)$. But we know that there is at least one minimum so that it must be determined by (3.11), i.e., $t_0 = t'$. From (3.11) we find by Newton's method that $z_1(t_0) = 2.73$ so that $t_0 = .104\cdots$. Therefore, we find that

$$\lambda_1[G(x, t)] \geq \lambda_1]G(x, t_0)] = 9.397\cdots.$$

Corresponding to Theorem 2.2, we have the following.

THEOREM 3.2. *Let $\lambda_k[p]$, $k = 1, 2, \cdots, n$ be the first n eigenvalues of a string of unit length with fixed ends whose density is a continuous convex function $p(x)$. If $\lambda_k[G(x, t)]$, $k = 1, 2, \cdots, n$ are the first n eigenvalues of a string whose density is the convex increasing function $G(x, t)$ defined in* (3.1), *then*

$$\left[\int_0^1 p(x)dx\right]^{-1} \sum_{k=1}^{n} \lambda_k^{-1}[p] \leq \sum_{k=1}^{n} \lambda_k^{-1}[G(x, t_0)]$$

where t_0 is a suitable value in $[0, 1]$.

Evidently, the inequality is sharp. The proof has the same formal relationship to that of Theorem 3.1 as the proof of Theorem 2.2 had to that of Theorem 2.1. Since no additional ideas are involved, we omit the proof.

Theorem 3.1 can be used to obtain an explicit lower bound for the second eigenvalue $\lambda_2[p]$ of the system (1.5) when $p(x)$ is convex.

THEOREM 3.3. *Let $\lambda_2[p]$ be the second eigenvalue of a string with convex density. Then*

$$\lambda_2[p]\int_0^1 p(x)dx \geq 4\lambda_0 \ ,$$

where λ_0 is the value defined in Theorem 3.1. *This inequality is sharp and equality is attained for a string with density $q(x)$ where*

$$q(x) = \begin{cases} G(1 - 2x, t_0), \ x \in [0, 1/2] \ , \\ G(2x - 1, t_0), \ x \in [1/2, 1] \ , \end{cases}$$

$G(x, t)$ and t_0 being defined as in Theorem 3.1.

Let $u_2(x)$ be the eigenfunction corresponding to $\lambda_2[p]$. Then $u_2(x)$ has exactly one nodal point $x_1 \in (0, 1)$. If we hold the string fixed at this point, we get two independent strings, each of which has the lowest eigenvalue $\lambda_2[p]$. By Theorem 3.1 and equation (1.6), the lowest eigenvalue of a string with a convex density satisfies the inequalities

$$\lambda_2[p] \geq \frac{\lambda_0}{x_1\int_0^{x_1} p(x)dx} = \mu' \tag{3.13}$$

and

$$\lambda_2[p] \geq \frac{\lambda_0}{(1 - x_1)\int_{x_1}^{1} p(x)dx} = \mu'' . \tag{3.14}$$

We may take the density function for which the bound in (3.13) is attained to be

$$\bar{p}(x) = \alpha G\left(\frac{x_1 - x}{x_1}, t_0\right), \ x \in [0, x_1] ,$$

where α is such that $\int_0^x \bar{p}(x)dx = \int_0^{x_1} p(x)dx$. For the second segment the bound is attained for the density function

$$\bar{p}(x) = \beta G\left(\frac{x - x_1}{1 - x_1}, t_0\right), \ x \in [x_1, 1]$$

where β is such that $\int_{x_1}^1 \bar{p}(x)dx = \int_0^{x_1} p(x)dx$.

Now, consider a string whose density function is defined by piecing together the above densities at x_1. This particular choice of $\bar{p}(x)$ then gives us a convex function. The second eigenvalue $\lambda_2[\bar{p}]$, of the resulting string satisfies the relationship $\lambda_2[\bar{p}] \leq \max(\mu', \mu'')$. Hence, (3.13) and (3.14) imply that

$$\lambda_2[p] \geq \lambda_2[\bar{p}] , \tag{3.15}$$

where $\lambda_2[\bar{p}]$ is the second eigenvalue of

$$u'' + \lambda\bar{p}(x)u = 0, \ u(0) = u(1) = 0 . \tag{3.16}$$

We now consider this system with $\bar{p}(x)$ replaced by

$$\bar{q}(x) = \begin{cases} \alpha G\left(\frac{x_1 - x}{x}, t_0\right) x \in [0, x_1] , \\ \beta G\left(\frac{x_1 - x}{1 - x_1}, t_0\right) x \in [x_1, 1] , \end{cases}$$

where α and β are now determined by the conditions $\int_0^1 \bar{q}(x)dx = M$, $\alpha \geq 0$, $\beta \geq 0$ and where $x_1 \in [0, 1]$ is a free parameter. It is clear that $\text{g.l.b.}_{(\alpha,\beta,x_1)} \lambda_2[\bar{q}] \leq \lambda_2[\bar{p}] \leq \lambda_2[p]$. We may let $\theta M = \int_0^{x_1} \bar{q}(x)dx$ so that $\int_{x_1}^1 \bar{q}(x)dx = (1-\theta)M$. We now show that smallest possible value of $\lambda_2[\bar{q}]$ is attained when $x_1 = 1/2$ and $\theta = 1/2$.

Let $\lambda_2[\bar{q}] = \lambda_2(x_1, \theta)$. We first show that $M\lambda_2(x_1, \theta) > 4\lambda_0$ if $x_1 \in [0, 1/4)$ or $x_1 \in (3/4, 1]$. Assume $x_1 \in [0, 1/4)$ and consider the case where x_1', the nodal point of the eigenfunction corresponding to $\lambda_2(x_1, \theta)$, lies in the interval $[0, 1/4)$. If we hold the string fixed at x_1', the resulting segments each have a lowest eigenvalue $\lambda_2(x_1, \theta)$. In particular, $\lambda_2(x_1, \theta)$ is the lowest eigenvalue of the segment $[0, x_1']$. By Theorem 3.1 and equation (1.6), we have

$$\lambda_2(x_1\, \theta) \geq \frac{\lambda_0}{x_1' \int_0^{x_1'} \bar{q}(x)dx} \geq \frac{\lambda_0}{x'M} .$$

But $x_1' < 1/4$ so that $M\lambda_2(x_1, \theta) > 4\lambda_0$. It follows in the same way that if $x_1' \in (3/4, 1]$ then $\lambda_2(x_1, \theta) \geq \lambda_0(1 - x_1')^{-1} \cdot M^{-1}$ and hence we again have $M\lambda_2(x_1, \theta) > 4\lambda_0$. Similarly if $x_1 \in (3/4, 1]$ we conclude that $M\lambda_2(x_1, \theta) > 4\lambda_0$ so that this inequality holds unless $x_1 \in [1/4, 3/4]$.

Hence, we consider the system (3.16) with $\bar{q}(x)$ in place of $\bar{p}(x)$ where $x_1 \in [1/4, 3/4]$. With x_1 restricted in this manner, it follows that the family of density functions $\bar{q}(x)$ is bounded uniformly. Hence, by Lemma 1.3, there are value x_1^0 and θ^0 such that $\lambda_2(x_1^0, \theta^0) = \min_{x_1,\theta} \lambda_2[x_1, \theta]$ for some density $\bar{q}^0(x)$.

We first note that x_1^0 must be a nodal point of the corresponding eigenfunction, for otherwise we could hold the string fixed at the nodal point and find another density which would give a lower second eigenvalue by the process which was used to obtain (3.15).

Thus, the lowest second eigenvalue is given by

$$M\lambda_2[x_1^0, \theta^0] = \frac{\lambda_0}{x_1^0\theta^0} = \frac{\lambda_0}{(1 - x_1^0)(1 - \theta^0)} .$$

Solving for θ^0, we find $\theta^0 = 1 - x_1^0$, so that $\lambda_2[x_1^0, \theta^0] = \lambda_0/(x_1^0(1 - x_1^0))$. This is obviously smallest when $x_1^0 = 1/2$, so that $\theta^0 = 1/2$. By (1.15) we then have the desired result.

We now consider the higher eigenvalues of the system (1.5) when $p(x)$ is convex. Unfortunately, we cannot use the technique just described for $\lambda_2[p]$, since the resulting function will, in general, not be convex. It is, however, possible to obtain a geometric characterization of the extremal density.

THEOREM 3.4. *Let $\lambda_n[p]$ be the* nth *eigenvalue of a string of unit length with fixed ends whose density is a continuous convex function $p(x)$. Then there is a string with the same total mass whose density is a piecewise linear convex function $q(x)$ with at most $n+1$ distinct linear segments such that*

$$\lambda_n[p] \geq \lambda_n[q]$$

where $\lambda_n[q]$ is the nth *eigenvalue of the string with density $q(x)$.*

Let $u_n(x)$ be the nth eigenfunction of (1.5) when $p(x)$ is convex. As in the proof of Theorem 2.3, we use the fact that $\lambda_n = J(p, u_n)$, where $J(p, u)$ is the Rayleigh quotient (1.10). If we construct a function $q(x)$ such that the inequality (2.11) is satisfied, then it follows as in the proof of Theorem 2.3 that $\lambda_n[q] \leq \lambda_n[p]$. Hence, it remains to be shown that such a function $q(x)$ exists.

We begin by carrying out a preliminary construction. As in Theorem 2.3, we denote the minimum points of $u_n^2(x)$ by x_k, $k = 0, 1, \cdots, n$, and the maximum points by $\bar{x}_k$, $k = 1, 2, \cdots, n$. We consider each of the intervals $(x_{k-1}, x_k) k = 1, 2, \cdots, n$ separately. Let $a(x)$ be any linear function such that $a(x) \leq p(x)$, $x \in [x_{k-1}, x_k]$. Then $r(x) = \max[a(x), 0]$ satisfies the inequality $0 \leq r(x) \leq p(x)$.

We now consider one of the intervals, say (x_{k-1}, x_k), where $1 \leq k \leq n$. Let c_k be any number such that $c_k \geq p(\bar{x}_k)$. Then there is a number a_k such that

$$\int_{x_{k-1}}^{\bar{x}_k} [a_k(x - x_k) + c_k]dx = \int_{x_{k-1}}^{\bar{x}_k} p(x)dx \; . \tag{3.18}$$

If $a_k(x - \bar{x}_k) + c_k \geq r(x)$, $x \in [x_{k-1}, \bar{x}_k]$, then we define

$$g_k(x, c_k) = a_k(x - \bar{x}_k) + c_k, \; x \in [x_{k-1}, \bar{x}_k] \; .$$

If $a_k(x - \bar{x}_k) + c_k < r(x)$ for some $x \in [x_{k-1}, \bar{x}_k]$, we determine a_k by the condition

$$\int_{x_{k-1}}^{x'_k} r(x)dx + \int_{x'_k}^{\bar{x}_k} [a_k(x - \bar{x}_k) + c_k]dx = \int_{x_{k-1}}^{\bar{x}_k} p(x)dx \; , \tag{3.19}$$

where x'_k is such that $a_k(x'_k - \bar{x}_k) + c_k = r(x'_k)$. We then define

$$g_k(x, c_k) = \begin{cases} r(x) & , \quad x \in [x_{k-1}, x'_k] \; , \\ a_k(x - \bar{x}_k) + c_k \; , & x \in [x'_k, \bar{x}_k] \; . \end{cases}$$

Likewise, we find b_k such that

$$\int_{\bar{x}_k}^{x_k} [b_k(x - \bar{x}_k) + c_k]dx = \int_{\bar{x}_k}^{x_k} p(x)dx \; . \tag{3.20}$$

If $b_k(x - \bar{x}_k) + c_k \geq r(x)$, $x \in [\bar{x}_k, x_k]$, we define

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

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

$$\int_{\bar{x}_k}^{x_k''} [b_k(x - \bar{x}_k) + c_k]dx + \int_{x_k''}^{x_k} r(x)dx = \int_{\bar{x}_k}^{x_k} p(x)dx \tag{3.21}$$

where x_k'' satisfies $b_k(x_k'' - \bar{x}_k) + c_k = r(x_k'')$, and define

$$h_k(x, c_k) = \begin{cases} b_k(x - \bar{x}_k) + c_k\ , & x \in [\bar{x}_k, x_k'']\ , \\ r(x) & ,\ \ x \in [x_k'', x_k]\ . \end{cases}$$

We may consider a_k and b_k to be functions of c_k, where $c_k \geq p(\bar{x}_k)$. In fact, they are continuous functions for any finite c_k, since a small change in c_k can cause only a small change in either a_k or b_k. We want to show that there is a value of c_k, say c_k', such that $a_k = b_k$, i.e., such that $a_k - b_k = 0$ for $c_k = c_k'$. If $c_k = p(\bar{x}_k)$, the convexity of $p(x)$ implies that the corresponding value of $a_k - b_k$ is non-positive. Furthermore, for c_k sufficiently large, the corresponding value of $a_k - b_k$ is positive. But $a_k - b_k$ is a continuous function of c_k so that the desired value c_k' exist.

We now let

$$q_k(x) = \begin{cases} g_k(x, c_k')\ , & x \in [x_{k-1}, \bar{x}_k]\ , \\ h_k(x, c_k')\ , & x \in [\bar{x}_k, x_k]\ . \end{cases}$$

From (3.18) or (3.19), whichever applies, we have

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

Hence, the convexity of $p(x)$ and the form of $q_k(x)$ imply (by Lemma 1.2) that

$$\int_{x_{k-1}}^{\bar{x}_k} q_k(x)u_n^2(x)dx \geq \int_{x_{k-1}}^{\bar{x}_k} p(x)u_n^2(x)dx\ . \tag{3.22}$$

Similarly, from (3.20) or (3.21) we have

$$\int_{\bar{x}_k}^{x_k} q_k(x)u_n^2(x)dx \geq \int_{\bar{x}_k}^{x_k} p(x)u_n^2(x)dx\ . \tag{3.23}$$

We are now able to construct the function $q(x)$ by complete induction. To avoid excessive detail, we carry out the proof only up to $n=3$. In (x_0, x_1), we set $r(x) = 0$, and form $q_1(x)$. In (x_1, x_2) we also form $q_2(x)$

with $r(x) = 0$. Then, comparing $q_1(x_1)$ and $q_2(x_1)$, we have the following alternatives:

(i) If $q_1(x_1) > q_2(x_1)$, we form a new function $q_2(x)$ with $r(x) = \max[q_1(x), 0]$, $x \in [x_1, x_2]$.

(ii) If $q_1(x_1) < q_2(x_1)$, we form a new function $q_1(x)$ with $r(x) = \max[0, q_2(x)]$, $x \in [x_0, x_1]$.

(iii) If $q_1(x_1) = q_2(x_1)$, we leave $q_1(x)$ and $q_2(x)$ as they are. Using whichever alternative applies, we define

$$q^{(1)}(x) = \begin{cases} q_1(x) , & x \in [x_0, x_1] , \\ q_2(x) , & x \in [x_1, x_2] . \end{cases}$$

Now, form $q_3(x)$ with $r(x) = 0$ for $x \in [x_2, x_3]$ and compare $q^{(1)}(x_2)$ and $q_3(x_2)$. We have the same alternatives as above, the only difference in procedure being that if $q^{(1)}(x_2) < q_3(x_2)$, we must redefine $q^{(1)}(x)$ with $r(x) = \max[0, q_3(x)]$ for $x \in [x_0, x_2]$.

It is clear that the induction can be carried out. Furthermore, the resulting function $q(x)$ will be convex, for by the above construction any two adjacent linear segments of the graph of $q(x)$ can only have a common point, such that the corresponding value of $q(x)$ is less than or equal to $p(x)$. Because of this convexity, there is at most one subinterval over which $q(x) = 0$. Hence, for each of the points $\bar{x}_k$, $k = 1, \cdots, n$, there is at most one vertex of the graph of $q(x)$, except for possibly the one just mentioned. Thus, the graph of $q(x)$ has at most $n + 1$ linear segments. That $q(x)$ satisfies the inequality (2.11) follows immediately from (3.22) and (3.23). Hence, our theorem follows as in the proof of Theorem 2.3.

4. Concave densities. In this section, we consider the system (1.5) when $p(x)$ is concave. We prove the following.

THEOREM 4.1. *Let $\lambda_1[p]$ be the first eigenvalue of a string of unit length with fixed ends whose density is a continuous concave function $p(x)$. Then*

$$\lambda_1[p]\int_0^1 p(x)dx \geq \lambda_0$$

where $\lambda_0 = 6.952\cdots$. The inequality is sharp, equality being attained for a string whose density is the symmetric triangular function

$$T(x) = \begin{cases} 4x , & x \in [0, 1/2] , \\ 4(1 - x) , & x \in [1/2, 1] . \end{cases}$$

We first assume that $p(x)$ has finite left and right derivatives in the closed interval $[0, 1]$. We define $f(x) = -p'_-(x)$, $x \in [0, 1]$ where we have set $p'_-(0) = p'_+(0)$. We then define

$$g(x, t) = \begin{cases} x(1-t)\ , & x \in [0, t]\ , \\ (1-x)t\ , & x \in [t, 1]\ . \end{cases}$$

It follows from integration by parts and the fact that

$$\int_0^1 f(t)dt = p(0) - p(1)$$

($p(t)$ is absolutely continuous) that

$$p(x) = p(0)(1-x) + p(1)x + \int_0^1 g(x, t)df(t)\ .$$

If we set $G(x, t) = 2/t(1-t)\, g(x, t)$, we get

$$p(x) = \int_0^1 G(x, t)[t(1-t)/2]df(t)\ .$$

Here, we have modified $f(t)$ so that the integral includes the terms $p(0)(1-x)$ and $p(1)x$. By Lemma 1.1, we have

$$\lambda_1[p]\int_0^1 p(x)dx \geq \min_{t\in[0,1]} \lambda_1[G(x, t)]\ .$$

The minimum exists by Lemma 1.3.

If either the right or left derivative is not finite in $[0, 1]$, we consider the system

$$u'' + \lambda_\varepsilon pu = 0\ , \quad u(\varepsilon) = u(1-\varepsilon) = 0\ ,$$

where $\varepsilon > 0$ is small. The above considerations then hold for this system and we have

$$(1-2\varepsilon)\lambda_1[p]_\varepsilon\int_\varepsilon^{1-\varepsilon} p(x)dx \geq \min_{t\in[0,1]} \lambda_1[G(x, t)]\ \cdot$$

Letting $\varepsilon \to 0$, we have the desired result.

To find the value of t for which $\lambda_1[G(x, t)]$ is a minimum, we consider the system

$$u'' + \lambda G(x, t)u = 0\ , \quad u(0) = u(1) = 0\ . \tag{4.1}$$

It is convenient to translate the system to the interval $[-1/2, 1/2]$. Thus, we consider

$$v'' + \lambda G_1(x, t)v = 0\ , \quad v(-1/2) = v(1/2) = 0\ , \tag{4.2}$$

where

$$G_1(x,t)=\begin{cases}2\dfrac{2x+1}{2t+1}\,, & -1/2\le x\le t\le 1/2\,,\\[2ex] 2\dfrac{1-2x}{1-2t}\,, & -1/2\le t\le x\le 1/2\,.\end{cases}$$

We show that $\lambda_1[G_1(x, t)]$ is a minimum when $t = 0$.

Following Hardy, Littlewood and Polya [6], we define the rearrangement of a non-negative integrable function $g(x)$, $x \in [-1/2, 1/2]$ into a symmetrically decreasing function $\bar{g}(x)$, $x \in [-1/2, 1/2]$. To do this, consider the set $S = \{x \mid g(x) \geq y\}$, where y is some number in the range of $g(x)$. Let $\mu\{S\}$ be the measure of the set S. We define the function

$$m(y) = \mu\{x \mid g(x) \geq y\}$$

and let

$$\bar{g}(x)=\begin{cases}m^{-1}(2x)\,, & x\in[0,1/2]\,,\\ \bar{g}(-x)\,, & x\in[-1/2,0]\,.\end{cases}$$

where m^{-1} denotes the inverse of $m(y)$. In particular, since $m(y)=1-y/2$ for $g = G_1$, $y \in [0, 2]$, we find that this symmetrization transforms $G_1(x, t)$ into $G_1(x, 0)$. Thus

$$G_1(x,0)=\bar{G}_1(x,t)=\begin{cases}2(1-2x)\,, & x\in[0,1/2]\,,\\ 2(1+2x)\,, & x\in[-1/2,0]\,.\end{cases}$$

By a result of Beesack and Schwarz [2], the first eigenvalue of (4.2) is greater than the first eigenvalue of

$$v'' + \lambda\bar{G}_1(x, t)v = 0\,, \quad v(-1/2) = v(1/2) = 0\,,$$

i. e., $\lambda_1[G(x, t)] \geq \lambda_1[G(x, 1/2)]$. Hence, $\lambda_1[p] \geq \lambda_1[G(x, 1/2)] = \lambda_1[T(x)]$.

If we solve (4.1) with $t = 1/2$ [7], we find the eigenvalues to be

$$\lambda_n[G(x, 1/2)] = (9/2)z_n^2\,,$$

where z_n is the nth positive root of the equation

$$J_{-2/3}(z) = 0$$

and $J_{-2/3}$ is the Bessel function of order $-2/3$. Numerical calculation gives the result

$$\lambda_1[G(x, 1/2)] = 6.952\cdots.$$

As in the case of monotone and convex densities, we have the following result for the first n eigenvalues of (1.5) when $p(x)$ is concave.

THEOREM 4.2. *Let $\lambda_k[p]$, $k = 1, 2, \cdots, n$ be the first n eigenvalues of a string of unit length with fixed ends whose density is a continuous* *concave function* $p(x)$. *If* $\lambda_k[G(x, t)], k = 1, 2, \cdots, n$, *are the first* n *eigenvalues of a string whose density is the triangular function*

$$G(x, t) = \begin{cases} 2\dfrac{x}{t} & , \; x \in [0, t] , \\ 2\left(\dfrac{1-x}{1-t}\right) , & x \in [t, 1] . \end{cases}$$

then

$$\left[\int_0^1 p(x)dx\right]^{-1} \sum_{k=1}^{n} \lambda_k^{-1}[p] \sum_{k=1}^{n} \leq \lambda_k^{-1}[G(x, t_0)]$$

for suitable t_0.

It is evident that the inequality is sharp. We omit the proof since it contains no new ideas.

For the higher eigenvalues of (1.5), when $p(x)$ is concave, we prove the following.

THEOREM 4.3. *Let* $\lambda_n[p]$ *be the* nth *eigenvalue of a string of unit length with fixed ends whose density is a continuous concave function* $p(x)$. *Then there is a string with the same total mass whose density is a piecewise linear concave function* $q(x)$, *where the graph of* $q(x)$ *has at most* $n + 1$ *linear segments and where* $q(0) = q(1) = 0$, *such that*

$$\lambda_n[p] \geq \lambda_n[q]$$

where $\lambda_n[q]$ *is the* nth *eigenvalue of the string with density* $q(x)$.

We use the same construction as in the proof of Theorem 3.4. This is possible, for if $p(x)$ is concave, $-p(x)$ is convex.

Let $u_n(x)$ be the nth eigenfunction of (1.5) when $p(x)$ is concave. As in § 2, we denote the nodal points of $u_n(x)$ by x_k, $k = 0, 1, 2, \cdots, n$, and the maximum points by $\bar{x}_k, k = 1, \cdots, n$. Then $-u_n^2(x)$ has the maximum points x_k, $k = 0, 1, \cdots, x$, and the minimum points $\bar{x}_k$, $k = 1, 2, \cdots, n$. Over each of the intervals $(\bar{x}_k, \bar{x}_{k+1})$, $k = 1, \cdots, n - 1$, we way define $-q_k(x, c_k)$ where now $-p(x_k) \leq c_k \leq 0$, as in Theorem 3.4. As before, there is a value of $c_k = c_k'$ such that $q_k(x, c_k')$ is linear at $x = x_k$. For the intervals $(0, \bar{x}_1)$ and $(\bar{x}_n, 1)$ we let $c_0 = 0$ and and $c_n = 0$ so as to define $-q_0(x, c_0)$ and $-q_n(x, c_n)$. Now using the same induction argument as in Theorem 3.4, we form the functions $-q_k(x, c_k)$ and obtain a function $-q(x)$ which is convex and satisfies the inequality

$$\int_0^1 p(x)u_n^2(x)dx \leq \int_0^1 q(x)u_n^2 dx \; . \tag{4.4}$$

Here, $r(x)$ is always a linear function and is always negative. The graph of $-q(x)$ consists of at most $n + 1$ linear segments, one for every point x_k, $k = 0, 1, \cdots, n$, and $q(x)$ is concave. By the argument used in the proof of Theorem 2.3, it then follows that $\lambda_n[p] \geq \lambda_n[q]$.

5. The general Sturm-Liouville system. We now turn to the Sturm-Liouville system

$$[r(x)u']' + [\lambda p(x) - q(x)]u = 0 , \qquad u'(0) - h_0 u(0) = u'(1) + h_1 u(1) = 0 , \tag{5.1}$$

where $p(x)$ and $q(x)$ are non-negative integrable functions, $r(x) \in C'$ is positive, and $h_0 \geq 0$, $h_1 \geq 0$. The lowest eigenvalue of this system is given by

$$\lambda_1[p] = \min_{u \in C'} \frac{\int_0^1 [r(x)u'^2 + q(x)u^2]dx}{\int_0^1 p(x)u^2 dx} , \tag{5.2}$$

where the functions $u(x)$ satisfy the appropriate bonndary conditions.

It is easy to see that the conclusion of Lemma 1.1 also applies to this differential system, i. e., if $p(x)$ can be expressed in the form (1.8), then

$$\lambda_1[p]\int_0^1 p(x)dx \geq \underset{t \in [0,1]}{\text{g.l.b.}}\ \lambda_1[K(x, t)] ,$$

where $\lambda_1[K(x, t)]$ is the first eigenvalue of the system (5.1) with $p(x)$ replaced by $K(x, t)$. Hence, it is possible to generalize Theorems 2.1, 3.1 and 4.1 to the system (5.1). We have

THEOREM 5.1. *The densities $p(x)$ minimizing the expression $\lambda_1[p]\int_0^1 p(x)dx$, where $\lambda_1[p]$ is the lowest eigenvalue of the system* (5.1) *under the assumptions*

(a) *$p(x)$ is monotone,*

(b) *$p(x)$ is convex,*

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

are of the same character as those discussed in Theorems 2.1, 3.1 *and* 4.1, *respectively.*

The proof of the theorem presents no new features. It should, however, be noted that the symmetrization argument used in the proof of Theorem 4.1, cannot be applied unless $q(x)$ and $r(x)$ have the symmetric property $q(x) = q(1 - x)$ and $r(x) = r(1 - x)$. In the general case, all that can be said is that the graph of the extremal density consists of two linear segments passing through (0, 0) and (0, 1), respectively. It should also be noted that the lack of symmetry in the boundary condition of (5.1) makes it impossible to tell, in general, whether the g.l.b.

$\lambda_1[p]$ for monotone convex $p(x)$ is approached in the case of increasing $p(x)$ or decreasing $p(x)$.

We now consider the system

$$u'' + [\lambda p(x) - q(x)]u = 0\ , \quad u(0) = u(1) = 0\ . \tag{5.3}$$

We denote the first eigenvalue of (5.3) by $\lambda_1[p, q]$. We have the following lemma.

LEMMA 5.1. *Let $q(x)$ be of the form*

$$q(x) = \int_0^1 K(x, t)g(t)df(t)$$

where $K(x, t)$, $g(t)$ and $f(t)$ are as defined in Lemma 1.1; then

$$\lambda_1[p, q] \geq \operatorname*{g.l.b.}_{t \in [0,1]} \lambda_1[p, QK(x, t)]\ ,$$

where $\lambda_1[p, QK(x, t)]$ is the lowest eigenvalue of

$$u'' + [\lambda p(x) - QK(x, t)]u = 0\ , \quad u(0) = u(1) = 0$$

and $Q = \int_0^1 q(x)dx$.

By (5.2), we have

$$\lambda_1[p, q] = \min_{u \in C'} \frac{\int_0^1 \left\{ u'^2 + \left[\int_0^1 K(x, t)g(t)df(t) \right] u^2 \right\} dx}{\int_0^1 pu^2 dx}\ .$$

Hence,

$$\begin{aligned}
\lambda_1[p, q] &\geq \min_{u \in C'} \int_0^1 g(t) \left\{ \frac{\int_0^1 [(u')^2/Q + K(x, t)u^2] dx}{\int_0^1 p(x)u^2 dx} \right\} df(t) \\
&\geq (1/Q) \int_0^1 g(t)\lambda_1[p, QK(x, t)]df(t) \\
&\geq (1/Q) \int_0^1 g(t)df(t) \operatorname*{g.l.b.}_{t \in [0,1]} \lambda_1[p, QK(x, t)]\ .
\end{aligned}$$

But $Q = \int_0^1 g(t)df(t)$, so that the conclusion of the lemma follows .

If $q(x)$ is concave, Lemma 5.1 yields the following result.

THEOREM 5.2. *The lowest eigenvalue $\lambda_1[p, q]$ of (5.3) when $q(x)$ is concave satisfies the inequality*

$$\lambda_1[p, q] \geq \min_{t \in [0,1]} \lambda_1[p, QG(x, t)]$$

where $\lambda_1[p, QG(x, t)]$ is the lowest eigenvalue of (5.3) with $q(x)$ replaced by $QG(x, t)$, $G(x, t)$ being defined by

$$G(x, t) = \begin{cases} 2\dfrac{x}{t} , & x \in [0, t] , \\ 2\dfrac{(1-x)}{(1-t)} , & x \in [t, 1] . \end{cases}$$

We first assume that $q'_+(0)$ and $q'_-(1)$ are finite. As in Theorem 4.1, we may express $q(x)$ as

$$q(x) = \int_0^1 G(x, t)[t(1-t)/2]df(t) .$$

By Lemma 5.1, we have

$$\lambda_1[p, q] \geq \underset{t\in[0,1]}{\text{g. l. b.}}\, \lambda_1[p, QG(x, t)] . \tag{5.4}$$

If $q'_+(0)$ or $q'_-(1)$ are not finite, we consider the system (5.3) with x restricted to the interval $[\varepsilon, 1-\varepsilon]$. Transforming this system to the unit interval, we see that (5.4) applies, so that letting $\varepsilon \to 0$, we find that (5.4) holds in general for concave $q(x)$.

Since $\lambda_1[p, QG(x, t)]$ is a continuous function of $t \in [0, 1]$, there must be a value of t for which the greatest lower bound is attained.

The same procedure can be made to yield corresponding results in the case of monotone and convex densities.

We close this section with some remarks about the system

$$u'' + \lambda p(x)u = 0 , \quad u'(0) = u(1) = 0 , \tag{5.4}$$

where $\int_0^1 pdx = 1$.

If $p(x)$ is monotone increasing, then λ_1, the lowest eigenvalue of (5.4), satisfies $\lambda_1 \geq \pi^2/4$.

This follows immediately from Lemma 1.2 if we compare $p(x)$ and $q(x) = 1$.

Similarly, if we compare a concave density $p(x)$ and $q(x) = 2x$, we find that $\lambda_1[p]$ satisfies the inequality

$$\lambda_1 \geq \lambda_0$$

where $\lambda_0 = 6.95\cdots/4$ is the lowest eigenvalue of (5.4) with $p(x) = 2x$.

6. The vibrating rod. The eigenvalue problem associated with a rod with clamped ends at $x = a$ and $x = b$ is

$$y^{iv} - \mu\rho(x)y = 0 , \quad y(a) = y'(a) = y(b) = y'(b) = 0 . \tag{6.1}$$

As in the case of the string, we may transform this system to the unit interval. We have

$$u^{iv} - \lambda p(x)u = 0 , \quad u(0) = u'(0) = u(1) = u'(1) = 0 , \tag{6.2}$$

where $p(x) = (b-a)\rho[(b-a)x+a]$. we note that $\int_0^1 p(x)dx = \int_a^b p(x)dx$. The eigenvalues $\mu_n[\rho]$, $n = 1, 2, \cdots$, of (6.1) are related to those of (6.2) by the equation

$$\lambda_n[p] = (b-a)^3 \mu_n[\rho] , \qquad n = 1, 2, \cdots$$

The first eigenvalue of (6.2) is equal to the minimum of the Rayleigh quotient

$$J(p, u) = \frac{\int_0^1 (u'')^2 dx}{\int_0^1 p(x) u^2 dx} \tag{6.3}$$

where $u(x)$ ranges over all functions $u \in C^2$ such that $u(0) = u'(0) = 0$ and $u(1) = u'(1) = 0$.

The following results correspond to Theorems 2.1, 3.1 and 4.1 for the string.

THEOREM 6.1. *Let $\lambda_1[p]$ be the lowest eigenvalue of a rod of unit length with clamped ends. From the assumptions that*

(a) *$p(x)$ is monotone,*
(b) *$p(x)$ is convex,*
(c) *$p(x)$ is concave,*

we have

$$\lambda_1[p]\int_0^1 p(x)dx \geq \lambda_1[K(x, t_0)]$$

where

$$(a') \quad K(x, t_0) = \begin{cases} 0 , & x \in [0, t_0] , \\ \dfrac{1}{1-t_0} , & x \in (t_0, 1] , \end{cases}$$

$$(b') \quad K(x, t_0) = \begin{cases} 0 , & x \in [0, t_0] , \\ \dfrac{2(x-t_0)}{(1-t_0)^2} , & x \in (t_0, 1] , \end{cases}$$

$$(c') \quad K(x, t_0) = \begin{cases} 2\dfrac{x}{t_0} , & x \in [0, t_0] , \\ 2\dfrac{1-x}{(1-t_0)} , & x \in [t_0, 1] \end{cases}$$

respectively, for suitable values of t_0.

There is nothing new involved in the proof over that of the corresponding theorems for the string. In fact, we need only replace the

Rayleigh quotient of (1.5) by (6.3) and the respective proofs for the corresponding string problem apply.

In the case of concave $p(x)$, it can be shown that $\lambda_1[p]$ takes its smallest value for $t_0 = 1/2$, i.e., we have the following result.

THEOREM 6.2. *If $\lambda_1[p]$ is the lowest eigenvalue of a rod with clamped ends whose density function is a positive concave function $p(x)$, then*

$$\lambda_1[p]\int_0^1 pdx \geq \lambda_1[p_0] ,$$

where

$$p_0(x) = \begin{cases} 4x , & x \in [0, 1/2] , \\ 4(1-x) , & x \in [1/2, 1] . \end{cases}$$

The proof will be based on the following result of Beesack [1].

THEOREM 6.3. *Let $p(x)$ be continuous and non-negative for $x \in [-1/2, 1/2]$ and let $\bar{p}(x)$ be the rearrangement of $p(x)$ into symmetrically decreasing order. Then the first eigenvalues of the system*

$$(6.4) \quad u^{iv} - \lambda p(x) = 0, \quad u(-1/2)=u'(-1/2) = u(1/2) = u'(1/2) = 0 ,$$

and

$$(6.5) \quad v^{iv} - \mu\bar{p}(x)v = 0, \quad v(-1/2) = v'(-1/2) = v(1/2) = v'(1/2) = 0 .$$

satisfy the condition

$$(6.6) \qquad \mu_1[\bar{p}] \leq \lambda_1[p] .$$

The rearrangement of $p(x)$ into symmetrically decreasing order is defined as above in Theorem 4.1.

The proof of Theorem 6.2 follows immediately from Theorems 6.1 and 6.3, since the symmetrization of

$$K(x, t) = \begin{cases} 2\dfrac{x}{t} , & x \in [0, t] , \\ 2\dfrac{1-x}{1-t} , & x \in [t, 1] , \end{cases}$$

is $K(x, 1/2)$.

Theorem 6.3 also leads to a result corresponding to that of Krein for a string with a bounded density function.

THEOREM 6.4. *Let $p(x)$ satisfy the condition $0 \leq p(x) \leq H < \infty$, $x \in [0, 1]$. Then the lowest eigenvalue $\lambda_1[p]$ of a rod with clamped ends* *and density $p(x)$ satisfies the inequality*

$$\lambda_1[p]\int_0^1 p(x)dx \geq \lambda_0 \,,$$

where λ_0 is the lowest eigenvalue of the rod with density

$$p_0(x) = \begin{cases} 0\,, & x \in \left[0,\, 1/2 - \dfrac{M}{2H}\right)\,, \\ H\,, & x \in \left[1/2 - \dfrac{M}{2H},\ 1/2 + \dfrac{M}{2H}\right]\,, \\ 0\,, & x \in \left(1/2 + \dfrac{M}{2H},\ 1\right]\,, \end{cases}$$

and where $M = \int_0^1 p(x)dx$.

Let $y_1(x)$ be the first eigenfunction of a rod with clamped ends and density $\bar{p}(x)$, (the function resulting from symmetrization of $p(x)$ about $x = 1/2$). Then it is clear that $p_0(x), \bar{p}(x)$ and $y_1^2(x)$ satisfy the hypothesis of Lemma 1.2 over the interval $[0, 1/2]$ so that

$$\int_0^{1/2} \bar{p}(x)y_1^2(x)dx \leq \int_0^{1/2} p_0(x)y_1^2(x)dx \,.$$

By symmetry, we have

$$\int_{1/2}^1 \bar{p}(x)y_1^2(x)dx \leq \int_{1/2}^1 p_0(x)y_1^2(x)dx \,.$$

Adding these two inequalities, we find

$$\lambda_1[\bar{p}] = \frac{\int_0^1 (y_1'')^2 dx}{\int_0^1 \bar{p}(x)y_1^2 dx} \geq \frac{\int_0^1 (y_1'')^2 dx}{\int_0^1 p_0(x)y_1^2 dx} \geq \lambda_0 \,.$$

Hence, by Theorem 6.2 we have $\lambda_1[p] \geq \lambda_0$.

We close this section with the remark that corresponding versions of Theorem 6.1 hold if we replace the boundary conditions (6.2) by any of the other boundary conditions used in the theory of the vibrating rod.

7. The vibrating membrane. We consider a vibrating membrane covering a simply connected domain D whose boundary is a Jordan curve C. Let $p(x, y)$ be the density of the membrane. We assume that $p(x, y)$ is measurable and that

$$0 \leq p(x, y) \leq H < \infty \,, \quad (x, y) \in D \,. \tag{7.1}$$

The eigenfunctions and the eigenvalues of this membrane, with the boundary fixed at $(x, y) \in C$, are determined by the integral equation [14]

$$u(x, y) = \lambda \iint_D G(x, y, \xi, \eta) p(\xi, \eta) u(\xi, \eta) d\xi d\eta ,$$

where $G(x, y, \xi, \eta)$ is the Green's function of the domain D. We denote the first eigenvalue by $\lambda_1[p]$.

We define M and R by the relations

$$M = \iint_D p(x, y) dx dy , \quad \pi R^2 = \iint_D dx dy \tag{7.3}$$

and let D^* be the circle $x^2 + y^2 \leq R^2$. In this section, we prove the following two theorems concerning $\lambda_1[p]$.

THEOREM 7.1. *The minimum of* $\lambda_1[p]$, *subject to the restrictions* (7.1) *and* (7.3), *is given by a membrane covering* D^* *with density*

$$\bar{p}_0(x, y) = \begin{cases} H , & x^2 + y^2 \leq \rho^2 \\ 0 , & \rho^2 \leq x^2 + y^2 \leq R^2 \end{cases} \tag{7.4}$$

where ρ *is defined by* $\pi \rho^2 H = M$.

Let D be a convex domain. $p(x, y)$ is concave in D if, for $(x_1, y_1) \in D$ and $(x_2, y_2) \in D$, we have

$$p\left(\frac{x_1 + x_2}{2} , \frac{y_1 + y_2}{2}\right) \leq (1/2)[p(x_1, y_1) + p(x_2, y_2)] .$$

For a concave density function, we have the following result.

THEOREM 7.2. *Let* $\lambda_1[p]$ *be the lowest eigenvalue of a membrane, with fixed edges, covering a convex domain* D, *whose density function is concave. Then*

$$\lambda_1[p] \iint_D p(x, y) dx dy \geq \pi \lambda_0 ,$$

where $\lambda_0 = 3.26\cdots$. *The inequality is sharp and equality is attained for a circular membrane of radius* R *and density*

$$p_0(x, y) = p_0(r) = \frac{3M}{\pi R^3}(R - r) , \quad 0 < r \leq R , \tag{7.5}$$

where $r^2 = x^2 + y^2$.

Krein [8] has conjectured the result of Theorem 7.1 for the case where $D = D^*$. The numerical value of the minimum is given by the least positive root of

$$J_0(\sqrt{\lambda H}\rho) - \rho J_0'(\sqrt{\lambda H}\rho)\sqrt{\lambda H}\, ln\, \rho/R = 0 .$$

To prove these theorems, we use the extremal characterization of $\lambda_1[p]$, i.e., the first eigenvalue $\lambda_1[p]$ of (7.2) is given by

$$\lambda_1[p] = \underset{u}{\text{g. l. b.}}\, J(p, u) , \tag{7.6}$$

where $J(p, u)$ is the Rayleigh quotient

$$J(p, u) = \frac{\iint_D |\operatorname{grad} u(x, y)|^2 dxdy}{\iint_D p(x, y)u^2(x, y)dxdy} \tag{7.7}$$

and where the greatest lower bound is taken over all continuous functions with piecewise continuous first derivatives, such that $u = 0$ on C.

As the following lemma shows, the same result is obtained if u is made subject to additional restrictions.

LEMMA 7.1. *The first eigenvalue* $\lambda_1[p]$ *of* (7.2) *is given by*

$$\lambda_1[p] = \underset{u}{\text{g. l. b.}}\, J(p, u) \tag{7.8}$$

where $J(p, u)$, *is the Rayleigh quotient* (7.7) *and where the greatest lower bound is taken over all analytic functions* $u(x, y)$ *with* $u = 0$ *on* C.

Since $p(x, y)$ is a measurable function, $\sqrt{p(x, y)}$ is measurable. Hence, there is a polynomial, $Q(x, y)$, such that for arbitrary numbers $\delta, \eta > 0$, we have

$$|\sqrt{p(x, y)} - Q(x, y)| < \delta \tag{7.9}$$

except on a set of measure less than η. Furthermore, $Q(x, y)$ may be chosen such that $q(x, y) = Q^2(x, y)$ is non-negative and is less than H.

We consider the membrane over the domain D with density $q(x, y)$. The eigenfunctions and eigenvalues will be determined by (7.2) with $p(x, y)$ replaced by $q(x, y)$. In particular, we denote the first eigenvalue by $\lambda_1[q]$. Since $q(x, y)$ is an analytic function in D, it is well known that

$$\lambda_1[q] = \min_{u \in C^\infty} J(q, u), \tag{7.10}$$

where the minimum is taken over all the indicated functions for which $u \equiv 0$ on C. It is also well known that the eigenvalues of (7.2) are continuous functionals of $p(x, y)$. More precisely, for any $\varepsilon > 0$, there is a $\delta_1 > 0$ such that

$$\iint_D \iint_D G^2(x, y, \xi, \eta)[\sqrt{p(x, y)p(\xi, \eta)} - \sqrt{q(x, y)q(\xi, \eta)}]^2 dxdy\, d\xi d\eta < \delta_1 \tag{7.11}$$

implies

$$\left| \frac{1}{\lambda_1[p]} - \frac{1}{\lambda_1(q)} \right| < \varepsilon . \tag{7.12}$$

It is easy to see that δ and η may be chosen so that (7.9) implies (7.11) and hence, also (7.12).

For any analytic function $u(x, y)$ such that $u = 0$ on C, we now show that there are values of δ and η such that (7.9) implies

$$| [J(p, u)]^{-1} - [J(q, u)]^{-1} | < \varepsilon \tag{7.13}$$

where $\varepsilon > 0$ is arbitrary and δ and η are independent of $u(x, y)$.

We have

$$A = | [J(p, u)]^{-1} - [J(q, u)]^{-1} | \leq [J(| p - q |, u)]^{-1}.$$

Now $| p - q | = | \sqrt{p} + \sqrt{q} \, | \, | \sqrt{p} - \sqrt{q} \, | \leq 2\sqrt{H}\delta$, except on a set S of measure less than η. Hence,

$$A \leq 2\sqrt{H}\delta[J(1, u)]^{-1} + H\eta \frac{\max_{(x,y)\in D} u(x, y)}{\iint_D | \operatorname{grad} u |^2 dxdy} .$$

By Rayleigh's theorem on the first eigenvalue of a homogeneous membrane [11], $J(1, u) \geq j_0^2/R^2$, where j_0 is the least positive zero of the Bessel function $J_0(x)$ and R is defined by (7.3). Furthermore, if we let $u_m = \max_{(x,y)\in D} u(x, y)$, then

$$\iint_D \left| \operatorname{grad} \frac{u}{u_m} \right|^2 dxdy \geq 4\pi c ,$$

where c is the capacity of an infinite circular cylinder of radius R with zero potential on the surface of the cylinder and potential one on the axis of the cylinder [11]. Hence, we have

$$A \leq 2\sqrt{H}\frac{R^2}{j_0^2}\sigma + \frac{H}{4\pi c}\eta ,$$

so that (7.13) follows.

Let $u_1(x, y)$ be the first eigenfunction corresponding to $\lambda_1[q]$. We may choose δ and η so that (7.12) and (7.13) hold simultaneously. Hence, we have

$$\begin{aligned} | \lambda_1^{-1}[p] - [J(p, u_1)] | &\leq | \lambda_1^{-1}[p] - \lambda_1^{-1}[q] | \\ &\quad + | [J(q, u_1)] - [J(p, u_1)] | \leq 2\varepsilon , \end{aligned} \tag{7.14}$$

for some function $u_1(x, y)$ which is analytic in D.

By (7.13) we have, for any analytic function $u(x, y)$ that

(7.15) $$[J(p, u)]^{-1} < [J(q, u)]^{-1} + \varepsilon \ .$$

(7.10) and (7.12) then give

$$[J(p, u)]^{-1} < \lambda_1^{-1}[q] + \varepsilon < \lambda_1^{-1}[p]\, 2\varepsilon \ .$$

Since ε is small, we finally conclude that $\lambda_1[p] \leq J(p, u)$. This with (7.14) gives (7.7).

We now introduce the symmetrization of $p(x, y)$ with respect to a line perpendicular to the (x, y)-plane, i. e., Schwarz symmetrization [**11**]. We may define it by considering the function

$$a(\rho) = \mu\{(x, y),\ p(x, y) \geq \rho\}$$

where μ denotes the measure of the set indicated and where ρ is some number between 0 and H. Then the symmetrization of $p(x, y)$ is

$$\bar{p}(x, y) = \bar{p}(r) = a^{-1}(\pi r^2)\ , \quad r \in [0, R] = D^*\ ,$$

where $r^2 = x^2 + y^2$.

We now prove the following.

LEMMA 7.2. *The lowest eigenvalue* $\lambda_1[p]$ *of* (7.2) *is bounded below by the lowest eigenvalue* $\lambda_1[\bar{p}]$ *of the membrane with fixed boundary over* D^* *and density* $\bar{p}(x, y)$.

B. Schwarz [**15**] has shown that when $p(x, y) \in C'$,

$$\lambda_1[p] \geq \underset{u \in C'}{\text{g. l. b.}}\ J[\bar{p}, u]\ ,$$

where now the Rayleigh quotient is defined over D^*. By Lemma 7.1, it follows that $\lambda_1[p] \geq \lambda_1[\bar{p}]$. The proof of Lemma 7.2 differs only in detail from the proof of the result of Schwarz.

By Lemma 7.1, there is an analytic function $u(x, y)$ such that

(7.16) $$\lambda_1[p] + \varepsilon \geq J(p, u)$$

where $\varepsilon > 0$ is arbitrary. Let $\bar{u}(x, y) = \bar{u}(r), r \in D^*$ be the above symmetrization of $u(x, y)$. Schwarz shows that such a symmetrization of an analytic function gives a function with piecewise continuous first derivatives and it is further known [**11**] that

$$\iint_D |\operatorname{grad} u|^2 dxdy \geq \iint_{D^*} |\operatorname{grad} \bar{u}|^2 r dr d\theta\ .$$

We also known [**6**], [**11**] that

$$\iint_D p(x, y) u^2(x, y) dxdy \leq \iint_{D^*} \bar{p}(r) \bar{u}^2(r) r dr d\theta\ .$$

Hence, we have from (7.16) and Lemma 7.1 that

$$\lambda_1[p] + \varepsilon \geq J(\bar{p}, \bar{u}) \geq \lambda_1[\bar{p}] .$$

But ε is arbitrary so that Lemma 7.2 follows.

We now prove Theorem 7.1. By Lemma 7.1, there is a symmetric and analytic function $\bar{u}(x, y) = \bar{u}(r)$, $(x, y) \in D^*$, such that for arbitrary $\varepsilon > 0$,

$$\lambda_1[\bar{p}] + \varepsilon \geq J(\bar{p}, \bar{u}) .$$

$\bar{u}(r)$ may be chosen such that it is the first eigenfunction of a membrane with a symmetric, analytic density $\bar{q}(x, y) = \bar{q}(r)$, $(x, y) \in D^*$. In this, case, the integral equation which gives $\bar{u}(r)$ is equivalent to the partial differential equation of this membrane. It is easily seen that $\bar{u}(r)$ must have its only maximum at $r = 0$. We now compare the integrals

$$\iint_{D^*} \bar{p}(r)\bar{u}(r) r dr d\theta$$

and

$$\iint_{D^*} \bar{p}_0(r)\bar{u}(r) r dr d\theta ,$$

where $\bar{p}_0(r) = \bar{p}_0(x, y)$ is defined by (7.4). From the definition of $\bar{p}(r)$ we have $0 \leq \bar{p}(r) \leq H$, $0 < r \leq R$. Hence, $\bar{p}(r)r$ and $\bar{p}_0(r)r$ satisfy the same relationship as $p(x)$ and $q(x)$ of Lemma 1.2. It then follows that

$$\lambda_1[\bar{p}] + \varepsilon \geq J(\bar{p}_0 \bar{u}) .$$

By Lemma 7.1, we have $\lambda_1[\bar{p}] \geq \lambda_1[\bar{p}_0]$, since ϵ is arbitrary. In view of Lemma 7.2, this proves Theorem 7.1.

To prove Theorem 7.2, we again consider $\bar{p}(x, y) = \bar{p}(r)$, $(x, y) \in D^*$. This function is obtained by Schwarz symmetrization from $p(x, y)$, $(x, y) \in D$, where D is a convex domain. We show that if $p(x, y)$ is concave, then so is $\bar{p}(x, y)$.

Consider the three dimensional set

$$S = \{(x, y, z) \mid (x, y) \in D , \quad 0 \leq z \leq p(x, y)\} .$$

This set is convex and Steiner symmetrization, i. e., symmetrization with respect to a plane, preserves convexity [3]. Furthermore, $\bar{p}(x, y)$ may be obtained by an infinite number of Steiner symmetrizations with respect to planes through the origin which are perpendicular to (x, y)-plane [3], [**11**]. This symmetrization of S gives

$$\bar{S} = \{(x, y, z) \mid (x, y) \in D , \quad 0 \leq z \leq \bar{p}(x, y)\} .$$

Clearly, $\bar{p}(x, y)$ will then be a concave function.

As in the proof of Theorem 7.1, there is an analytic function

$\bar{u}(x, y) = \bar{u}(r)$ whose only extremal value is the maximum at $r = 0$ such that $\lambda_1[\bar{p}] + \varepsilon \geq J(\bar{p}, \bar{u})$. Since $\bar{u}(r)$ is concave, $\bar{p}_0(r)$ and $\bar{p}(r)$ satisfy the relation $\bar{p}_0(r) \geq \bar{p}(r)$ for $r \in (0, r_0)$ where $r_0 \in (0, R)$ and $\bar{p}_0(r) \leq \bar{p}(r)$ for $r \in (r_0, R)$. Hence, $r\bar{p}_0(r)$ and $r\bar{p}(r)$ are related in the same way as $p(x)$ and $q(x)$ in Lemma 1.2. As in Theorem 7.1, we have $\lambda_1[\bar{p}] \geq \lambda_1[\bar{p}_0]$. By Lemma 7.2, Theorem 7.2 then follows.

Using well-known techniques for the computation of eigenvalues [4], we find that the lowest eigenvalue $\lambda_1[p]$ of (7.2), where $p(x, y)$ is concave, satisfies the relation

$$\iint_D p(x, y)dxdy \cdot \lambda_1[p] \geq (3.26\cdots)\pi .$$

Z. Nehari has shown [10] that if $p(x, y)$ is superharmonic, then $\lambda_1[p]$ satisfies

$$\iint_D p(x, y)dxdy\lambda_1[p] \leq \pi j_0^2$$

where j_0 is the least positive zero of $J_0(x)$, the Bessel function of order zero. But a concave function is superharmonic [12] so that the bound also applies in our problem. Thus, if $p(x, y)$ is concave in a convex domain,

$$3.26\pi \leq \iint_D p(x, y)dxdy\lambda_1[p] \leq \pi j_0^2 .$$

References

1. P. R. Beesack, *Isoperimetric inequalities for the nonhomogeneous clamped rod and plate,* J. Math. Mech. **8**, pp. 471-482.
2. P. R. Beesack and B. Schwarz, *On the zeros of solutions of second order linear differential equations,* Canadian J. Math., **8** (1956) 504-515.
3. W. Blaschke, *Kreis und Kugel,* Chelsea Publishing Co., 1949.
4. L. Collatz, *Eigenwertprobleme und ihre Numerische Behandlung,* Chelsea Publishing Co., 1948.
5. R. Courant and D. Hilbert, *Methods of Mathematical Physics,* Interscience Publishers, 1953.
6. G. H. Hardy, J. E. Littlewood, and G. Polya, *Inequalities,* Cambridge University Press, 1952.
7. E. Jahnke and Emde, *Tables of Functions,* Dover, 1945.
8. M. G. Krein, *On certain problems on the maximum and minimum of characteristic values and on the Lyapunov zones of stability,* Amer. Math. Soc. Trans., Series 2, **1**, (1955), 163-187.
9. National Bureau of Standards, *Tables of Bessel Functions of Fractional Order,* Vol. I, Columbia University Press, 1948.
10. Z. Nehari, *On the principal frequency of a membrane,* Pacific J. Math. **8**, (1958) 285-293.
11. G. Polya and G. Szego, *Isoperimetric Inequalities in Mathematical Physics,* Princeton University Press, 1951.

12. T. Rado, *Subharmonic Functions*, Chelsea Publishing Co., 1949.
13. E. C. Titchmarsh, *The Theory of Functions*, Oxford University Press, 1939.
14. F. G. Tricomi *Integral Equations*, Interscience Publishers, 1957.
15. B. Schwarz, *Bounds for the principal frequency of the non-homogeneous membrane*, Pacific J. Math. **7**, (1957), 1653–1676.

CARNEGIE INSTITUTE OF TECHNOLOGY

ON THE SUMMABILITY OF DERIVED FOURIER SERIES

B. J. BOYER

1. Introduction. Bosanquet ([1] and [2]) has shown that the $(C, \alpha + r), \alpha \geq 0$, summability of the rth derived Fourier series of a Lebesgue integrable function $f(x)$ is equivalent to the (C, α) summability at $t = 0$ of the Fourier series of another function $\omega(t)$ (see (4), §2) integrable in the Cesaro-Lebesgue (CL) sense. This result suggests the following question: Is there a class of functions, integrable in a sense more general than that of Lebesgue, which permits such a characterization for the summability of rth derived Fourier series and which is large enough to contain $\omega(t)$ also?

In this paper it will be shown that such a characterization is possible within the class of Cesaro-Perron (CP) integrable functions for a summability scale more general than the Cesaro scale (Theorems 1 and 2, §4). Theorem 3 provides sufficient conditions for the summability of the Fourier series of $\omega(t)$ in terms of the Cesaro behavior of $\omega(t)$ at $t=0$.

Integrals are to be taken in the CP sense and of integral order, the order depending on the integrand.[1] It will be convenient to define the $C_{-1}P$ integral as the Lebesgue integral.

2. Definitions. A series Σu_ν is said to be summable (α, β) to S if

$$\lim_{n\to\infty} B\sum_{\nu<n} (1 - \nu/n)^\alpha \log^{-\beta}\left(\frac{1}{1 - \nu/n}\right)u_\nu = S$$

for C sufficiently large, where $B = \log^\beta C$ and $C > 1$. (It is sufficient to say for every $C > 1$.[2])

The function $\lambda_{\alpha,\beta}(x)$ is defined by the equation:

$$\lambda_{\alpha,\beta}(x) + i\tilde{\lambda}_{\alpha,\beta}(x) = \frac{B}{\pi}\int_0^1 (1 - u)^{\alpha-1} \log^{-\beta}\left(\frac{C}{1 - u}\right)e^{ixu}du. \tag{1}$$

$$\varphi(t) \equiv \varphi(t, r, x) = \tfrac{1}{2}[f(x + t) + (-1)^r f(x - t)]. \tag{2}$$

$$P(t) \equiv P(t, r) = \sum_{i=0}^{[r/2]} \frac{a_{r-2i}}{(r - 2i)!} t^{r-2i}. \tag{3}$$

$$\omega(t) = t^{-r}[\varphi(t) - P(t)], \tag{4}$$

[1] Many properties of CP integration have been given by Burkill ([4], [5] and [6]) and by Sargent [7]. Other properties used in this paper can easily be verified by induction.

Received July 6, 1959.

[2] Bosanquet and Linfoot [3]. They have also shown the consistency of this scale for $\alpha' > \alpha$ or $\alpha' = \alpha, \beta' > \beta$.

for $-\pi < t < \pi$ and is of period 2π.

The rth derived Fourier series of $f(t)$ at $t = x$ will be denoted by $D_rFSf(x)$, and the nth mean of order (α, β) of $D_rFSf(x)$ by $S_{\alpha,\beta}(f, x, n)$. The kth iterated integral of $f(x)$ will be written $F_k(t)$ or $[f(t)]_k$.

3. **Lemmas.** The following result is due to Bosanquet and Linfoot [3]:

LEMMA 1. *For $r \geq 0$ and $\alpha = 0, \beta > 1$ or $\alpha > 0, \beta \geq 0$,*

$$\lambda^{(r)}_{1+\alpha,\beta}(x) = 0(|x|^{-1-\alpha}\log^{-\beta}|x|) + |x|^{-r-2}) \text{ as } |x| \to \infty .$$

LEMMA 2. *For $\alpha \geq 0, \beta \geq 0$ and $r \geq 0$,*

$$x^r\lambda^{(r)}_{1+\alpha+r,\beta}(x) = \sum_{i,j=0}^{r} B^r_{ij}(\alpha, \beta)\lambda_{1+\alpha+r-i,\beta+j}(x),$$

where the $B^r_{ij}(\alpha, \beta)$ are independent from x and have the properties:

(i) $B^r_{ij}(\alpha, 0) = 0$ *for* $j \geq 1$;
(ii) $\beta^r_{r0}(\alpha, \beta) \neq 0$.

Proof. Let us put $\gamma_{1+\alpha,\beta}(x) = \lambda_{1+\alpha,\beta}(x) + i\bar{\lambda}_{1+\alpha,\beta}(x)$. For $r = 0$ we take $B^0_{00}(\alpha, \beta) = 1$. For $r \geq 1$ an integration by parts and the identity $u^r = -u^{r-1}(1-u) + u^{r-1}$ yield the following recursion:

$$\begin{aligned} x^r\gamma^{(r)}_{1+\alpha+r,\beta}(x) &= -(\alpha+2r)x^{r-1}\gamma^{(r-1)}_{1+\alpha+r,\beta}(x) - \frac{\beta}{\log C}x^{r-1}\gamma^{(r-1)}_{1+\alpha+r,\beta+1}(x) \\ &\quad + (\alpha+r)x^{r-1}\gamma^{(r-1)}_{1+\alpha+r-1,\beta}(x) + \frac{\beta}{\log C}x^{r-1}\gamma^{(r-1)}_{1+\alpha+r-1,\beta+1}(x) . \end{aligned} \tag{5}$$

The lemma follows easily from successive applications of equation (5).

LEMMA 3. *For $n > 0$ and $\alpha = 0, \beta > 1$ or $\alpha > 0, \beta \geq 0$,*

$$\left(\frac{d}{dt}\right)^r\left\{\frac{1}{2\pi} + \frac{B}{\pi}\sum_{\nu\leq n}\left(1-\frac{\nu}{n}\right)^\alpha \log^{-\beta}\left(\frac{C}{1-\frac{\nu}{n}}\right)\cos\nu t\right\} = n^{r+1}\sum_{k=-\infty}^{\infty}\lambda^{(r)}_{1+\alpha,\beta}[n(t+2k\pi)] ,$$

for $r = 0, 1, 2, \cdots$.

Proof. Smith ([8], Lemma, 3.1) has shown that for every even periodic, Lebesgue integrable function $Z(t)$,

$$2n\int_0^\infty Z(t)\lambda_{1+\alpha,\beta}(nt)dt = S_{\alpha,\beta}(Z, 0, n) . \tag{6}$$

Using Lemma 1 and the properties of $Z(t)$, one can show in a straightforward manner that

$$(7)\qquad \int_0^\infty Z(t)\lambda_{1+\alpha,\beta}(nt)dt = \int_0^\pi Z(t)\sum_{k=-\infty}^{\infty}\lambda_{1+\alpha,\beta}[n(t+2k\pi)]dt\,.$$

Let us define $Z(t) = \begin{Bmatrix} 1 \text{ for } |t| \le x \\ 0 \text{ for } x < |t| \le \pi \end{Bmatrix}$. Equations (6) and (7) imply that for every x, $0 \le x \le \pi$,

$$(8)\qquad \int_0^x n\sum_{k=-\infty}^{\infty}\lambda_{1+\alpha,\beta}[n(t+2k\pi)]dt = \int_0^x\left\{\frac{1}{2\pi} + \frac{B}{\pi}\sum_{\nu\le n}\left(1-\frac{\nu}{n}\right)^\alpha \cdot \log^{-\beta}\left(\frac{C}{1-\frac{\nu}{n}}\right)\cos\nu t\right\}dt\,.$$

Since the integrands in (8) are continuous, even and periodic, the lemma is proven for $k = 0$.

To prove the lemma for $k \ge 1$, we need only to observe that the derived series are uniformly convergent in every closed interval by Lemma 1.

LEMMA 4. *Let $f(x) \in CP[-\pi, \pi]$ and be of period 2π. Then for $n > 0$ and $\alpha = 0, \beta > 1$ or $\alpha > 0, \beta \ge 0$,*

$$S^r_{\alpha,\beta}(f, x, n) = 2(-1)^r n^{r+1}\int_0^\pi \varphi(t)\sum_{k=-\infty}^{\infty}\lambda^{(r)}_{1+\alpha,\beta}[n(t+2k\pi)]dt\,.$$

Proof. This result can be verified by direct calculation using Lemma 3 and the properties of CP integration.

When $f(x)$ is Lebesgue integrable, Lemma 4 is equivalent to a slightly different representation given by Smith [8].

LEMMA 5. *Let $f(x) \in C_\mu P[-\pi, \pi]$ and be of period 2π. Let ξ, $0 \le \xi \le \mu + 1$, be an integer for which $\varphi_\xi(t) \in L[0, \pi]$. Then, for $r \ge 0$ and $\alpha = \xi, \beta > 1$ or $\alpha > \xi, \beta \ge 0$,*

$$S^r_{\alpha+r,\beta}(f, x, n) - a_r = 2(-1)^r n^{r+1}\int_0^\pi [\varphi(t) - P(t)]\lambda^{(r)}_{1+\alpha+r,\beta}(nt)dt + o(1)\,.$$

Proof. From Lemmas 1 and 4 we see that

$$\begin{aligned} S^r_{\alpha+r,\beta}(P, 0, n) &= 2(-1)^r n^{r+1}\int_0^\pi P(t)\sum_{k=-\infty}^{\infty}\lambda^{(r)}_{1+\alpha+r,\beta}[n(t+2k\pi)]dt \\ &= 2(-1)^r n^{r+1}\int_0^\pi P(t)\lambda^{(r)}_{1+\alpha+r,\beta}(nt)dt + o(1) \text{ as } n\to\infty\,. \end{aligned}$$

Since $(d/dt)^r P(t) = a_r$, then $S^r_{\alpha+r,\beta}(P, 0, n) \to a_r$ for $\alpha = 0$, $\beta > 1$ or $\alpha > 0, \beta \geq 0$.[3]

It remains to be shown that

$$(9) \qquad S^r_{\alpha+r,\beta}(f, x, n) = 2(-1)^r n^{r+1} \int_0^\pi \varphi(t) \lambda^{(r)}_{1+\alpha+r,\beta}(nt)dt + o(1) .$$

Successive integrations by parts give

$$(10) \qquad n^{r+1}\int_0^\pi \varphi(t) \sum_{k=-\infty}^{\infty}{}' \lambda^{(r)}_{1+\alpha+r,\beta}[n(t+2k\pi)]dt = \sum_{j=0}^{\xi-1}(-1)^j n^{r+1+j}\Phi_{j+1}(\pi)$$
$$\cdot \sum_{k=-\infty}^{\infty}{}' \lambda^{(r+j)}_{1+\alpha+r,\beta}[\pi n(2k+1)] + (-1)^\xi n^{r+1+\xi}\int_0^\pi \Phi_\xi(t) \sum_{k=-\infty}^{\infty}{}'$$
$$\cdot \lambda^{(r+\xi)}_{1+\alpha+r,\beta}[n(t+2k\pi)]dt$$

By Lemma 1 each of the integrated terms on the right side of (10) is $o(1)$ as $n \to \infty$, and

$$n^{r+1+\xi} \sum_{k=-\infty}^{\infty}{}' \lambda^{(r+\xi)}_{1+\alpha+r,\beta}[n(t+2k\pi)] = o(1)$$

uniformly in t, $0 \leq t \leq \pi$. Since $\Phi_\xi(t)$ is Lebesgue integrable, it follows that the left side of (10) is $o(1)$. This result and Lemma 4 prove (9) and complete the proof of the lemma.

It can be shown that Lemma 5 holds if $\int_0^\pi$ is replaced by $\int_0^\delta$, $\delta > 0$. Thus, for the values of α and β under consideration, the summability of $D_r FSf(x)$ is a local property of $f(x)$.

Having found an expression for $S^r_{\alpha,\beta}(f, x, n)$, let us estimate the integer ξ in the preceding lemma.

LEMMA 6. *If* $h(t) \in C_\mu P[0, a]$ *and* $t^r h(t) \in C_\lambda P[0, a]$, *then*

$$H_{1+\xi}(t) \in L[0, a], \text{ where } \xi = \min[\mu, \max(\lambda, r)] .$$

Proof. The case $\mu = -1$ is trivial by definition of $C_{-1}P$. Therefore, let us assume $\mu \geq 0$. We may also assume, by the consistency of CP integration, that $\lambda \geq r$.

It will be convenient to use the "integration by parts" formula:

$$(11) \qquad [t^r h(t)]_k = \sum_{j=0}^{r} C_j(k, r) t^{r-j} H_{k+j}(t), \qquad k = 1, 2, \cdots ,$$

where the $C_j(k, r)$ do not depend on t or the function h.

By the Cesaro continuity and consistency of CP integration, there exists an integer $k \geq \lambda + 1$ such that for $j \geq 0$,

$$(12) \qquad H_{k+1+j}(t) = o(t^{k+j-r}) \text{ as } t \to 0 .$$

[3] Smith [8], Theorem 3.1.

Since $k \geq \lambda + 1$, equations (11) and (12) imply

$$[t^r h(t)]_k = o(t^{k-1}) = t^r H_k(t) + \sum_{j=1}^{r} t^{r-j} o(t^{k+j-1-r}) ;$$

hence, $H_k(t) = o(t^{k-1-r})$. This result and (12) yield

$$H_{k+j}(t) = 0\,(t^{k-1+j-r}) \text{ as } t \to 0 \text{ for } j \geq 0 . \tag{13}$$

Since (13) is merely (12) with k replaced by $k-1$, this inductive process terminates with $H_{\lambda+1}(t) = o(t^{\lambda-r})$. Therefore, $H_{\lambda+1}(t) = o(1)$ as $t \to 0$ if $\lambda \geq r$.

But for $\eta > 0$, $h(t) \in C_\lambda P[\eta, a]$. Therefore, $H_{1+\xi}(t) \in L[0, a]$.

Lemmas 5 and 6 may be combined to give the following:

LEMMA 7. *Let $f(x) \in C_\lambda P[-\pi, \pi]$ and be of period 2π. If $\omega(t) \in C_\mu P[0, \pi]$, then for $\alpha = 1 + \xi, \beta > 1$ or $\alpha > 1 + \xi, \beta \geq 0$,*

$$S_{\alpha,\beta}(\omega, 0, n) = 2n\int_0^\pi \omega(t)\lambda_{1+\alpha,\beta}(nt)dt + o(1), \text{ where } \xi$$
$$= \min [\mu, \max (\lambda, r)] .$$

This section is concluded with two results of Tauberian nature.

LEMMA 8. *If $\alpha \geq 0, \beta > 0$, $\{b_i\}_{i=0}^{k}$* and *$\{a_\nu\}_{\nu=0}^{\infty}$ are sequences of real numbers with $b_0 \neq 0$, and if*

$$F_{\alpha,\beta}(n) = \sum_{i=0}^{k} b_i \sum_{\nu \leq n} \left(1 - \frac{\nu}{n}\right)^\alpha \log^{-(\beta+i)} \left(\frac{C}{1 - \frac{\nu}{n}}\right) a_\nu = o(1) \text{ as } n \to \infty ,$$

then $\sum_{\nu=0}^{\infty} a_\nu = o(a, \beta)$.

The proof of this result is too long to be given here. In general, however, this method is similar to one employed by Bosanquet and Linfoot.[4]

LEMMA 9. *Let $S_{\alpha,\beta}(u, n)$ denote the* nth *mean of order (α, β) of the series Σu_ν. For α, β and $r \geq 0$ and $i, j = 0, 1, \cdots, r$, let us assume that*

(i) *The constants $C_{ij}(\alpha, \beta, r)$ have properties* (i) *and* (ii) *of the $B^r_{ij}(\alpha, \beta)$ in Lemma 2;*

(ii) $\sum_{i,j=0}^{r} C_{ij}(k + \alpha, \beta, r) S_{k+\alpha+r-1,\beta+j}(u, n) = o(1), \qquad k = 0, 1, 2, \cdots ;$

(iii) $\sum_{\nu=0}^{\infty} u_\nu = 0(C)$.

[4] Bosanquet and Linfoot [**3**], Theorem 3.1.

Then $\sum_{\nu=0}^{\infty} u_\nu = 0(\alpha, \beta)$.

Proof. Let us consider the case $\beta > 0$. By (iii) of the lemma and the consistency of (α, β) summability, there exists an integer $K \geq 1$ such that $S_{\alpha+K+i\cdot\beta+j}(u, n) = o(1)$ as $n \to \infty$ for $i, j = 0, 1, 2, \cdots$. Putting $k = K - 1$ in (ii) above, we see that

$$\sum_{j=0}^{r} C_{rj}(K - 1 + \alpha, \beta, r) S_{K-1+\alpha,\beta+j}(u, n) + o(1) = o(1) .$$

Therefore, from (i) above and Lemma 8, $S_{K-1+\alpha,\beta}(u, n) = o(1)$. That is, for $K \geq 1$, $\sum_{\nu=0}^{\infty} u_\nu = o(\alpha + K, \beta)$ implies $\sum_{\nu=0}^{\infty} u_\nu = o(\alpha + K - 1, \beta)$. It follows immediately that $\sum_{\nu=0}^{\infty} u_\nu = o(\alpha, \beta)$.

The case $\beta = 0$, in which we deal with linear combinations of Riesz means, is proved similarly.

4. Theorems.

THEOREM 1. *Let* $f(x) \in C_\lambda P[-\pi, \pi]$ *and be of period* 2π. *If there exist constants* a_{r-2i}, $i = 0, 1, \cdots, [r/2]$, *such that*

(i) $\omega(t) \in C_\mu P]0, \pi]$ *for some integer* μ;

(ii) $FS\omega(0) = 0(\alpha, \beta)$ *for* $\alpha = 1 + \xi, \beta > 1$ *or* $\alpha > 1 + \xi, \beta \geq 0$, *where* $\xi = \min[\mu, \max(\lambda, r)]$,

then $D_r FSf(x) = a_r(\alpha + r, \beta)$.

THEOREM 2. *Let* $f(x) \in C_\lambda P[-\pi, \pi]$ *and be of period* 2π. *If* $D_r FSf(x) = a_r(\alpha + r, \beta)$ *for* $\alpha = 1 + \lambda, \beta > 1$ *or* $\alpha > 1 + \lambda, \beta \geq 0$, *then there exist constants* a_{r-2i}, $i = 0, 1, \cdots, [r/2]$, *such that*

(i) $\omega(t) \in C_\mu P[0, \pi]$ *for some integer* μ;

(ii) $FS\omega(0) = o(\alpha', \beta')$, *where*

$$\left.\begin{cases} \alpha' = 1 + \xi, \beta' > 1 \ \textit{if}\ 1 + \lambda \leq \alpha < 1 + \xi \ \textit{or}\ \alpha = 1 + \xi, \beta \leq 1 \\ \alpha' = \alpha, \beta' = \beta \ \textit{if}\ \alpha = 1 + \xi, \beta > 1 \ \textit{or}\ \alpha > 1 + \xi, \beta \geq 0 \end{cases}\right\} \text{and}$$

$\xi = \min[\mu, \max(\lambda, r)]$.

Before proving these theorems, let us observe that the existence of the a_{r-2i} in the theorems implies their uniqueness from the definition of $\omega(t)$. In fact, somewhat more is true. Observe that $\omega(t) = \omega(t, r) \in CP[0, \pi]$ implies $\omega(t, r - 2i) = o(1)(C)$ as $t \to 0$. Therefore, if $\omega(t, r) \in CP[0, \pi]$ and $FS\omega(0) = 0(C)$, then assuming the truth of Theorems 1 and 2, it is clear that the a_{r-2i} are given by the formula:

$$D_{r-2i} FSf(x) = a_{r-2i}(C),\ i = 0, 1, \cdots [r/2] .$$ [5]

Proof of Theorem 1. Lemma 7 and the consistency of (α, β) sum-

[5] Compare Bosanquet [**2**], eqn. 5.2, for $f(x) \in L[\pi, \pi]$.

mability give the relations:

$$2n\int_0^\pi \omega(t)\lambda_{1+\alpha+r-i,\beta+j}(nt)dt = S_{\alpha+r-i,\beta+j}(\omega, 0, n) + o(1) = o(1) ,$$

for $i, j = 0, 1, 2, \cdots, r$. Therefore,

$$2n\int_0^\pi \omega(t) \sum_{i,j=0}^{r} B_{ij}^r(\alpha, \beta)\lambda_{1+\alpha+r-i,\beta+j}(nt)dt = o(1) ,$$

which by Lemma 2 becomes

$$2n^{r+1}\int_0^\pi \omega(t)t^r\lambda_{1+\alpha+r,\beta}^{(r)}(nt)dt = o(1) . \tag{14}$$

Since $t^r\omega(t) = \varphi(t) - P(t)$, relation (14) and Lemma 5 imply that $S_{\alpha+r,\beta}^r(f, x, n) - a_r = o(1)$, i.e., $D_rFSf(x) = a_r(\alpha + r, \beta)$.

Proof of Theorem 2. Let us first prove part (i). Putting $P(t) \equiv 0$ in Lemma 5, we obtain

$$2(-1)^r n^{r+1}\int_0^\pi \varphi(t)\lambda_{1+\alpha+r,\beta}^{(r)}(nt)dt = S_{\alpha+r,\beta}(f, x, n) + o(1) . \tag{15}$$

If the left side of (15) is integrated by parts $\lambda + 1$ times, the integrated part is $o(1)$ as $n \to \infty$ by Lemma 1, and (15) becomes

$$2(-1)^{r+\lambda+1}n^{r+\lambda+2}\int_0^\pi \Phi_{\lambda+1}(t)\lambda_{1+\alpha+r,\beta}^{(r+\lambda+1)}(nt)dt = S_{\alpha+r,\beta}^r(f, x, n) + o(1) \tag{16}$$

Let us define $\Phi_{\lambda+1}(t)$ for $-\pi < t < 0$ to be an odd (even) function if $r + \lambda + 1$ is odd (even). Then (16) may be written

$$S_{\alpha+r,\beta}^{r+\lambda+1}(\Phi_{\lambda+1}, 0, n) = S_{\alpha+r,\beta}^r(f, x, n) + o(1) .$$

It follows that $D_{r+\lambda+1}FS\Phi_{\lambda+1}(0) = a_r(C)$.

Since $\Phi_{\lambda+1}(t) \in L[-\pi, \pi]$, a theorem of Bosanquet establishes the following result.[6] There exist constants $a^{r+\lambda+1-2i}$, $i = 0, 1, \cdots, [(r+\lambda+1)/2]$, with $a^{r+\lambda+1} = a_r$, such that

$$\gamma(t) \equiv \{\Phi_{\lambda+1}(t) - P_*(t)\}t^{-(r+\lambda+1)} \in CL[0, \pi] \text{ and } FS\gamma(0) = 0(C) , \tag{17}$$

where $P_*(t) = \sum_{i=0}^{[(r+\lambda+1)/2]} [a^{r+\lambda+1-2i}/(r + \lambda + 1 - 2i)!]t^{r+\lambda+1-2i}$.

For $\lambda = -1$, put $a^{r-2i} = a_{r-2i}$ in (17). Then (17) states that $\omega(t) \in CP[0, \pi]$ and $FS\omega(0) = 0(C)$.

Let us consider the case $\lambda \geq 0$, and define $h(u, m + 1) \equiv \{\Phi_{m+1}(u) - P_*^{(\lambda-m)}(u)\}u^{-(r+m+1)}$, $m = -1, 0, 1, \cdots, \lambda$. Then for $0 < \eta < t \leq \pi$, an integration by parts yields

[6] Bosanquet [**2**], Theorem 2. The superscript notation has been used here to distinguish these constants from those whose existence is to be proven.

$$(18)\qquad \int_\eta^t h(u, m)du = uh(u, m+1)\Big|_\eta^t + (r+m)\int_\eta^t h(u, m+1)du .$$

Let us assume for the moment that for some integer m, $0 \leq m \leq \lambda$,

$$(19)\qquad h(u, m+1) \in C_k P[0, t],\ k \geq \lambda + 1.$$

From (19) and a result due to Sargent[7], it follows that

$$\int_\eta^t h(u, m+1)du \in C_k P[0, t] \text{ and } (C, k+1) \lim_{\eta \to 0} \int_\eta^t h(u, m+1)du$$
$$= \int_0^t h(u, m+1)du .$$

Since $\eta h(\eta, m+1) \in C_k P[0, t]$ and is $o(1)(C, k+1)$ as $\eta \to 0$, the right side of (18) has a limit $(C, k+1)$ as $\eta \to 0$. Sargent's result (*ibid.*) and equation (18) imply

$$(20)\qquad h(u, m) \in C_{k+1} P[0, t] .$$

We infer from the recursive behavior of (19) and (20) that whenever (19) is true, then $h(u, 0) \in CP[0, t]$. But (19) is true for $m = \lambda$ by (17). Therefore,

$$(21)\qquad h(t, 0) = \{\varphi(t) - P_*^{(\lambda+1)}(t)\}t^{-r} \in C_\mu P[0, \pi] \text{ for some } \mu .$$

In the course of the argument above, it has also been shown that by taking C-limits of (18) we obtain

$$(22)\qquad \int_0^t h(u, m)du = th(t, m+1) + (r+m)\int_0^t h(u, m+1)du$$

for $m = 0, 1, \cdots, \lambda$.

If we now define $a_{r-2i} = a^{r+\lambda+1-2i}$, $i = 0, 1, \cdots, [r/2]$, it is easily verified that $P_*^{(\lambda+1)}(t) = P(t)$ and $h(t, 0) = \omega(t)$. Part (i) of the theorem follows immediately from (21).

Next it will be shown that $FS\omega(0) = 0(C)$ for $\lambda \geq 0$, the case $\lambda = -1$ having been settled already.

From equations (11) and (22), it is seen that

$$(23)\qquad [h(t, m)]_{k+1} = t[h(t, m+1)]_k + (r+m-k)[h(t, m+1)]_{k+1} .$$

If for some integer m, $0 \leq m \leq \lambda$, the statement

$$(24)\qquad h(u, m+1) = o(1)(C, k) \text{ for some integer } k$$

is true, then (24) is also true when $m+1$ and k are replaced by m and $k+1$, respectively, by (23). In this manner we arrive at the conclusion that $h(t, 0) = \omega(t) = o(1)(C)$ as $t \to 0$, which ensures that $FS\omega(0) =$

[7] Sargent [**7**], Lemma 1.

$0(C)$. However, $h(u, \lambda + 1) = \gamma(t)$ and $FS\gamma(0) = 0(C)$ from (17). Therefore, $\gamma(t) = o(1)(C)$[8], so that (24) is true for $m = \lambda$.

It remains only to prove the order relations in part (ii).

Having determined the polynomial $P(t)$, we may state, with the aid of Lemmas 2 and 5, that

$$S^r_{\alpha+r,\beta}(f, x, n) - a_r = (-1)^r \sum_{i,j=0}^{r} B^r_{ij}(\alpha, \beta) \left\{2n\int_0^\pi \omega(t)\lambda_{1+\alpha+r-i,\beta+j}(nt)dt\right\} + o(1) . \tag{25}$$

If $\lambda + 1 \leq \alpha < 1 + \xi$ or $\alpha = 1 + \xi, \beta \leq 1$, then for $\beta^* > 1$ and $k = 0, 1, 2, \cdots, S^r_{r+1+\xi+k,\beta^*}(f, x, n) - a_r = o(1)$. Equation (25) then implies

$$\sum_{i,j=0}^{r} B^r_{ij}(1 + \xi + k, \beta^*)\left\{2n\int_0^\pi \omega(t)\lambda_{2+\xi+k+r-i,\beta^*+j}(nt)dt\right\} = o(1) \tag{26}$$

Similarly, for $\alpha = 1 + \xi, \beta > 1$ or $\alpha > 1 + \xi, \beta \geq 0$, it can be shown that

$$\sum_{i,j=0}^{r} B^r_{ij}(\alpha + k, \beta)\left\{2n\int_0^\pi \omega(t)\lambda_{1+\alpha+k+r-i,\beta+j}(nt)dt\right\} = o(1) . \tag{27}$$

With the definition of (α', β') and by means of Lemma 7, both (26) and (27) may be combined into the single equation:

$$\sum_{i,j=0}^{r} B^r_{ij}(\alpha' + k, \beta')S_{\alpha'+k+r-i,\beta'+j}(\omega, o, n) = o(1), k = 0, 1, 2, \cdots . \tag{28}$$

Since $FS\omega(0) = 0(C)$, Lemma 9 and (28) yield part (ii) of the theorem at once.

These two theorems may be combined in several ways to give generalizations to known results. In what follows it is assumed that $f(x) \in C_\lambda P[-\pi, \pi]$ and is of period 2π, $\xi = \min[\mu, \zeta]$ and $\zeta = \max(r, \lambda)$.

COROLLARY 1. *If* $\omega(t) \in C_\mu P[0, \pi]$, *then for* $\alpha = 1 + \xi, \beta > 1$ *or* $\alpha > 1 + \xi, \beta \geq 0, D_rFSf(x) = a_r(\alpha + r, \beta)$ *if and only if* $FS\omega(0) = 0(\alpha, \beta)$.[9]

COROLLARY 2. *For* $\alpha = 1 + \zeta, \beta > 1$ *or* $\alpha > 1 + \zeta, \beta \geq 0, D_rFSf(x) = a_r(\alpha + r, \beta)$ *if and only if* $\omega(t) \in CP[0, \pi]$ *and* $FS\omega(0) = 0(\alpha, \beta)$.[10]

From Corollary 2 it follows that $D_rFSf(x) = a_r(C)$ if and only if $\omega(t) \in CP[0, \pi]$ and $FS\omega(0) = 0(C)$. Along with a result by Sargent[11]

[8] That $FSg(0) = 0(C)$ if and only if $g(t) = o(1)(C)$ as $t \to 0$ has been shown by Sargent [**7**], Theorem 6.

[9] For $\mu = -1$ compare Wang [**9**].

[10] For $\alpha \geqq r + 1$ and $\lambda = -1$ compare Bosanquet [**2**].

[11] Sargent [**7**], Theorem 6.

this gives a solution, in the sense of Hardy and Littlewood, to the Cesaro summability problem for $D_r FSf(x)$ within the class of CP integrable functions.

The last theorem of this section sharpens a well known sufficient condition for the summability of $FS\omega(0)$ without, however, destroying the CP integrablity of $\omega(t)$.

THEOREM 3. *Let* $\omega(t) \in C_\mu P[-\pi, \pi]$ *and be an even function of period* 2π. *For* $k \geq \mu$, *sufficient conditions that* $FS\omega(0) = 0(1 + k, \beta)$, $\beta > 1$, *are*

(i) $\omega(t) = 0(1)(C, k + 1)$ and
(ii) $\omega(t) = o(1)(C, k + 2)$.

Proof. The proof of this theorem is similar to the proof of the analogous theorem for Riesz summability when $\omega(t)$ is Lebesgue integrable. Starting with Lemma 7 and $k + 1$ integrations by parts, one obtains

$$S_{1+k,\beta}(\omega, 0, n) = (-1)^{k+1} 2n^{k+2} \int_0^\pi \Omega_{k+1}(t) \lambda_{2+k,\beta}^{(k+1)}(nt)dt + o(1) .$$

Writing $\int_0^\pi = \int_0^{K/n} + \int_{K/n}^{\delta} + \int_{\delta}^{\pi}$, it can be shown by straightforward calculations that for arbitrary $\varepsilon > 0$ and $K > e$,
$|S_{1+k,\beta}(\omega, 0, n)| \leq M_1(K) \cdot \varepsilon + M_2 \int_K^\infty (X^{-1} \log^{-\beta} X + X^{-2})dX + o(1)$, where M_2 is independent from ε, K and n. The theorem follows from the last inequality by letting $n \to \infty, \varepsilon \to 0$ and $K \to \infty$ in that order.

The theorems of this section can be illustrated by means of the following CP integrable functions:
$t^{-m} \sin t^{-1}$ and $t^{-m} \cos t^{-1}$, $m = 0, 1, 2, \cdots$. For example, from Theorems 1 and 3, $FS[t^{-1} \sin t^{-1}]_{t=0} = 0(1, \beta)$ and $D_1 FS[\sin t^{-1}]_{t=0} = 0(2, \beta)$ for $\beta > 1$.

REFERENCES

1. L. S. Bosanquet, *Note on differentiated Fourier series*, Quart. Journal of Math. (Oxford), **10** (1939), 67–74.

2. ———, *A solution of the Cesaro summability problem for successively derived Fourier series*, Proc. London Math. Soc. (2), **46** (1940), 270–289.

3. L. S. Bosanquet and E. H. Linfoot, *Generalized means and the summability of Fourier series*, Quart. Journal of Math. (Oxford), **2** (1931), 207–229.

4. J. C. Burkill, *The Cesaro-Perron integral*, Proc. London Math. Soc. (2), **34** (1932), 314–322.

5. ———, *The Cesaro-Perron scale of integration*, Proc. London Math. Soc. (2), **39** (1935), 541–552.

6. ———, *The Cesaro scales of summation and integration*, Journal London Math. Soc.,

10 (1935), 254–259.
7. W. L. C. Sargent, *On the summability* (*C*) *of allied series*, Proc. London Math. Soc., **50** (1949), 330–348.
8. A. H. Smith, *On the summability of derived series of the Fourier-Lebesgue type*, Quart. Journal of Math. Oxford, **4** (1933), 93–106.
9. F. T. Wang, *Cesaro summation of the successively derived Fourier series*, Tohoku Math. Journal, **39** (1934) 399–405.

PURDUE UNIVERSITY

AN ELEMENTARY PROOF OF THE PRIME NUMBER THEOREM WITH REMAINDER TERM

ROBERT BREUSCH

Introduction. In this paper, the prime number theorem in the form $\psi(x) \equiv \sum_{p^m \leqq x} \log p = x + o(x \cdot \log^{-1/6+\varepsilon} x)$, for every $\varepsilon > 0$, is established via a proof that in the well-known formula

$$(1) \qquad \rho(x) \equiv \sum_{p^m \leqq x} \frac{\log p}{p^m} = \log x + O(1) \equiv \log x + a_x ,$$

$a_x = -A_0 + o(\log^{-1/6+\varepsilon} x)$. ($A_0$ is Euler's constant.)

Throughout the paper, p and q stand for prime numbers, k, m, n, t, and others are positive integers, and x, y, and z are positive real numbers.

Some well-known formulas, used in the proof, are

$$(2) \qquad \sum_{n \leqq x} \frac{\log^k n}{n} = \frac{1}{k+1} \cdot \log^{k+1} x + A_k + O\left(\frac{\log^k x}{x}\right), \quad \text{for } k = 0, 1, \cdots$$

$$(2') \qquad \sum_{y < n \leqq z} \frac{1}{n} \cdot \log^k (n/y) = \frac{1}{k+1} \cdot \log^{k+1}(z/y) + O\left(\frac{1}{y} \cdot \log^k (z/y)\right), \quad \text{for } k = 0, 1, \cdots$$

$$(3) \qquad \sum_{n \leqq x} \log^k (x/n) = O(x), \quad \text{for } k = 1, 2, \cdots$$

$$(4) \qquad \sum_{p^m \leqq x} \log p \cdot \log^k (x/p^m) = O(x), \quad \text{for } k = 0, 1, \cdots$$

$$(5) \qquad \sum_{n \leqq x} \mu(n)/n = O(1) \quad (\mu(n) \text{ is Moebius' function.})$$

Two other formulas, used prominently, are

$$(6) \qquad \sigma(x) \equiv \sum_{p^m \leqq x} \frac{\log p}{p^m} \cdot \log (x/p^m) = \frac{1}{2} \cdot \log^2 x - A_0 \cdot \log x + g_x \ (g_x = O(1))$$

$$(7) \qquad \tau(x) \equiv \sum_{p^m \leqq x} \frac{\log p}{p^m} \cdot \log^2 (x/p^m) = \frac{1}{3} \cdot \log^3 x - A_0 \cdot \log^2 x + (2 \cdot A_0^2 + 4 \cdot A_1) \log x + O(1) .$$

With the help of (1), (2), and (4), (6) can be proved easily:

$$\sigma(x) = \sum_{p^m \leqq x} \frac{\log p}{p^m} \cdot \left(\sum_{n \leqq x/p^m} 1/n - A_0 + O(p^m/x) \right), \quad \text{or, with } k = n \cdot p^m ,$$

$$\sigma(x) = \sum_{k \leqq x} \frac{1}{k} \cdot \sum_{p^m / k} \log p - A_0 \cdot \log x + O(1)$$

$$= \sum_{k \leqq x} \frac{\log k}{k} - A_0 \log x + O(1) = \frac{1}{2} \cdot \log^2 x - A_0 \log x + O(1) .$$

Received December 31, 1958.

Also, again with $k = n \cdot p^m$,

$$\sum_{k \leq x} \frac{\log k}{k} \cdot \log \frac{x}{k}$$

$$= \sum_{k \leq x} \frac{1}{k} \cdot \log \frac{x}{k} \cdot \sum_{p^m/k} \log p = \sum_{p^m \leq x} \log p \cdot \sum_{n \leq x/p^m} \frac{1}{n \cdot p^m} \cdot \log\left(\frac{x}{n \cdot p^m}\right)$$

$$= \sum_{p^m \leq x} \frac{\log p}{p^m} \cdot \left\{ \log\left(\frac{x}{p^m}\right) \cdot \sum_{n \leq x/p^m} \frac{1}{n} - \sum_{n \leq x/p^m} \frac{\log n}{n} \right\}$$

$$= \sum_{p^m \geq x} \frac{\log p}{p^m} \cdot \left\{ \log^2 (x/p^m) + A_0 \log (x/p^m) - \frac{1}{2} \log^2 (x/p^m) - A_1 \right\} + O(1)$$

(by (2) and (4))

$$= \frac{1}{2} \cdot \tau(x) + A_0 \cdot \sigma(x) - A_1 \cdot \rho(x) + O(1) .$$

(7) follows now by (1), (2), and (6).

The proof now proceeds in the following steps: in part I, certain asymptotic formulas for a_n (see (1)) and g_n (see (6)) are derived; they suggest that "on the average," a_n is $-A_0$, and g_n is $A_0^2 + 2A_1$. In part II, formulas for a_n and g_n are derived which are of the type of Selberg's asymptotic formula for $\psi(x)$; part III contains the final proof.

Part I

First, the following five formulas will be derived; $K_1, K_2, \cdots$, are constants, independent of x.

$$(8) \qquad \sum_{n \leq x} \frac{1}{n} \cdot a_n = - A_0 \log x + g_x + K_2 + O\left(\frac{\log x}{x}\right)$$

$$(9) \qquad \sum_{n \leq x} \frac{1}{n} \cdot a_{x/n} = - A_0 \log x + K_3 + O\left(\frac{\log x}{x}\right)$$

$$(10) \qquad \sum_{p \ \leq x} \frac{\log p}{p^m} \cdot a_{p^m} = - A_0 \log x + g_x + \frac{1}{2} a_x^2 + K_4 + O\left(\frac{\log x}{x}\right)$$

$$(11) \qquad \sum_{n \leq x} \frac{1}{n} \cdot g_n = (A_0^2 + 2 \cdot A_1) \cdot \log x + O(1)$$

$$(12) \qquad \sum_{n \leq x} \frac{1}{n} \cdot g_{x/n} = (A_0^2 + 2 \cdot A_1) \cdot \log x + K_5 + O\left(\frac{\log^2 x}{x}\right) .$$

Proofs.

$$\sigma(x) = \sum_{n \leq x} \log \frac{x}{n} (\rho(n) - \rho(n-1)) = \sum_{n \leq x} \rho(n) \cdot \log \frac{n+1}{n} + O\left(\frac{\log x}{x}\right)$$

$$= \sum_{n \leq x} \frac{\rho(n)}{n} + K_1 + O\left(\frac{\log x}{x}\right)$$

$$= \sum_{n \leq x} \frac{\log n}{x} + \sum_{n \leq x} \frac{1}{n} a_n + K_1 + O\left(\frac{\log x}{x}\right) .$$

(8) follows now from (6) and (2).

Also

$$\sum_{n\le x}\frac{1}{n}\cdot a_{x/n}=\sum_{n\le x}\frac{1}{n}\cdot\Bigl(\sum_{p^m\le x/n}\frac{\log p}{p^m}-\log\frac{x}{n}\Bigr)$$
$$=\sum_{k\le x}\frac{1}{k}\sum_{p^m/k}\log p-\sum_{n\le x}\frac{1}{n}\log\frac{x}{n}\quad (k=n\cdot p^m)$$
$$=\sum_{k\le x}\frac{\log k}{k}-\sum_{n\le x}\frac{1}{n}\log\frac{x}{n}\,.\quad \text{which proves (9) by (2).}$$

And

$$\sum_{p^m\le x}\frac{\log p}{p^m}\cdot a_{p^m}=\sum_{p^m\le x}\frac{\log p}{p^m}\Bigl(\sum_{q^t\le p^m}\frac{\log q}{q^t}-\log(p^m)\Bigr)$$
$$=\frac{1}{2}\Bigl(\sum_{p^m\le x}\frac{\log p}{p^m}\Bigr)^2+\frac{1}{2}\sum_{p^m\le x}\frac{\log^2 p}{p^{2m}}-\log x\cdot\sum_{p^m\le x}\frac{\log p}{p^m}+\sum_{p^m\le x}\frac{\log p}{p^m}\log\frac{x}{p^m}\,.$$

Thus, by (1), (2) and (6),

$$\sum_{p^m\le x}\frac{\log p}{p^m}\cdot a_{p^m}=\frac{1}{2}(\log x+a_x)^2+K_4+O\Bigl(\frac{\log x}{x}\Bigr)-\log x\cdot(\log x+a_x)$$
$$+\frac{1}{2}\log^2 x-A_0\log x+g_x\,,\quad \text{which proves (10).}$$

In the next proof, use is made of the easily established fact that

$$\rho(n)\cdot\log\frac{n+1}{n}=\sigma(n+1)-\sigma(n)\,.$$

$$\tau(x)=\sum_{n\le x}\log^2\Bigl(\frac{x}{n}\Bigr)(\rho(n)-\rho(n-1))$$
$$=\sum_{n\le x}\rho(n)\Bigl(\log^2\Bigl(\frac{x}{n}\Bigr)-\log^2\Bigl(\frac{x}{n+1}\Bigr)\Bigr)+O(1)$$
$$=\sum_{n\le x}\rho(n)\log\frac{n+1}{n}\cdot\log\frac{x^2}{n(n+1)}+O(1)$$
$$=\sum_{n\le x}(\sigma(n+1)-\sigma(n))\cdot\log\frac{x^2}{n(n+1)}+O(1)$$
$$=\sum_{n\le x}\sigma(n)\cdot\log\frac{n+1}{n-1}+O(1)=\sum_{n\le x}\sigma(n)\cdot\frac{2}{n}+O(1)$$
$$=\sum_{n\le x}\frac{\log^2 n}{n}-2\cdot A_0\cdot\sum_{n\le x}\frac{\log n}{n}+2\cdot\sum_{n\le x}\frac{1}{n}\cdot g_n+O(1)\quad\text{(by (6)).}$$

This proves (11), with the help of (2) and (7).

Finally

$$\sum_{n\le x}\frac{1}{n}\cdot g_{x/n}=\sum_{n\le x}\frac{1}{n}\cdot\Bigl(\sum_{p^m\le x/n}\frac{\log p}{p^m}\cdot\log\frac{x}{n\cdot p^m}-\frac{1}{2}\log^2\Bigl(\frac{x}{n}\Bigr)+A_0\log\frac{x}{n}\Bigr)\,,$$

or, with $k=n\cdot p^m$,

$$\sum_{n\leq x}\frac{1}{n}\cdot g_{x/n}$$

$$=\sum_{k\leq x}\frac{1}{k}\cdot\log\frac{x}{k}\cdot\sum_{p^m/k}\log p-\frac{1}{2}\cdot\sum_{n\leq x}\frac{1}{n}\cdot\log^2\left(\frac{x}{n}\right)+A_0\sum_{n\leq x}\frac{1}{n}\cdot\log\frac{x}{n}$$

$$=\sum_{k\leq x}\frac{1}{k}\cdot\log\frac{x}{k}\cdot\log k-\frac{1}{2}\cdot\sum_{n\leq x}\frac{1}{n}\cdot\log^2\left(\frac{x}{n}\right)+A_0\sum_{n\leq x}\frac{1}{n}\cdot\log\frac{x}{n}\ .$$

(12) now follows by (2).

Formulas (8) through (12) suggest setting

$$(13)\qquad b_x\equiv a_x+A_0\ ,\qquad h_x\equiv g_x-(A_0^2+2A_1)\ .$$

In terms of b_x and h_x, the five formulas read

$$(8')\qquad \sum_{n\leq x}\frac{1}{n}\cdot b_n=h_x+K_6+O\left(\frac{\log x}{x}\right)$$

$$(9')\qquad \sum_{n\leq x}\frac{1}{n}\cdot b_{x/n}=K_7+O\left(\frac{\log x}{x}\right)$$

$$(10')\qquad \sum_{p^m\leq x}\frac{\log p}{p^m}\cdot b_{p^m}=A_0\cdot(-A_0+b_x)+A_0^2+2\cdot A_1+h_x+K_4+\frac{1}{2}\cdot(-A_0+b_x)^2+O\left(\frac{\log x}{x}\right)$$

$$=h_x+\frac{1}{2}\cdot b_x^2+K_8+O\left(\frac{\log x}{x}\right)$$

$$(11')\qquad \sum_{n\leq x}\frac{1}{n}\cdot h_n=O(1)$$

$$(12')\qquad \sum_{n\leq x}\frac{1}{n}\cdot h_{x/n}=K_9+O\left(\frac{\log^2 x}{x}\right)\ .$$

Next, it will be shown that

$$(14)\qquad \sum_{n\leq x}\frac{1}{n}\cdot b_n^2=\sum_{n\leq x}\frac{1}{n}\cdot b_{x/n}^2+O(1)\ ,$$

and

$$(15)\qquad \sum_{n\leq x}\frac{1}{n}\cdot b_n\cdot b_{x/n}=\sum_{p^m\leq x}\frac{\log p}{p^m}\cdot h_{x/p^m}+O(1)\ .$$

For a proof of (14), we know, by (10′), that

$$\frac{1}{n}\cdot b_n^2=\frac{2}{n}\cdot\sum_{p^m\leq n}\frac{\log p}{p^m}\cdot b_{p^m}-\frac{2}{n}\cdot h_n-\frac{2}{n}\cdot K_8+O\left(\frac{\log n}{n^2}\right),$$

and

$$\frac{1}{n}\cdot b_{x/n}^2=\frac{2}{n}\cdot\sum_{p^m\leq x/n}\frac{\log p}{p^m}\cdot b_{p^m}-\frac{2}{n}\cdot h_{x/n}-\frac{2}{n}\cdot K_8+O\left(\frac{1}{x}\cdot\log\frac{x}{n}\right).$$

Thus, by (3), (11′) and (12′),

$$\sum_{n\leqq x}\frac{1}{n}\cdot(b_n^2-b_{x/n}^2)=2\cdot\sum_{n\leqq x}\frac{1}{n}\Big(\sum_{p^m\leqq n}\frac{\log p}{p^m}\cdot b_{p^m}-\sum_{p^m\leqq x/n}\frac{\log p}{p^m}\cdot b_{p^m}\Big)+O(1)$$

$$=2\cdot\sum_{p^m\leqq x}\frac{\log p}{p^m}\cdot b_{p^m}\Big(\sum_{p^m\leqq n\leqq x}\frac{1}{n}-\sum_{n\leqq x/p^m}\frac{1}{n}\Big)+O(1)$$

$$=2\cdot\sum_{p^m\leqq x}\frac{\log p}{p^m}\cdot b_{p^m}\cdot\Big(\log(x/p^m)+O(1/p^m)-\log(x/p^m)$$

$$-A_0-O(p^m/x)\Big)+O(1)$$

$=O(1)$, by (10′) and (4). This proves (14).

Also

$$\sum_{n\leqq x}\frac{1}{n}\cdot b_n\cdot b_{x/n}=\sum_{n\leqq x}\frac{1}{n}\cdot b_n\cdot\Big(\sum_{p^m\leqq x/n}\frac{\log p}{p^m}-\log\frac{x}{n}+A_0\Big)$$

$$=\sum_{n\leqq x}\frac{1}{n}\cdot b_n\cdot\Big(\sum_{p^m\leqq x/n}\frac{\log p}{p^m}-\sum_{t\leqq x/n}\frac{1}{t}+2\cdot A_0+O\Big(\frac{n}{x}\Big)\Big)$$

$$=\sum_{p^m\leqq x}\frac{\log p}{p^m}\sum_{n\leqq x/p^m}\frac{1}{n}b_n-\sum_{t\leqq x}\frac{1}{t}\sum_{n\leqq x/t}\frac{1}{n}b_n+O(1)\,,\quad\text{by (8′)}$$

$$=\sum_{p^m\leqq x}\frac{\log p}{p^m}\cdot h_{x/p^m}+K_6\log x-\sum_{t\leqq x}\frac{1}{t}h_{x/t}-K_6\log x+O(1)$$

(by (8′), (1) and (4))

$$=\sum_{p^m\leqq x}\frac{\log p}{p^m}\cdot h_{x/p^m}+O(1)\,,\quad\text{by (12′)}\,.$$

From (14) and (15) it follows that

$$\sum_{n\leqq x}\frac{1}{n}\cdot(b_n\pm b_{x/n})^2=2\cdot\sum_{n\leqq x}\frac{1}{n}\cdot b_n^2\pm 2\cdot\sum_{p^m\leqq x}\frac{\log p}{p^m}\cdot h_{x/p^m}+O(1)\,,$$

and therefore

(16) $$\sum_{n\leqq x}\frac{1}{n}\cdot b_n^2\geqq\Big|\sum_{p^m\leqq x}\frac{\log p}{p^m}\cdot h_{x/p^m}\Big|+O(1)\,.$$

Part II

In the following, we shall employ the inversion formula

$$G(x)=\sum_{n\leqq x}g\Big(\frac{x}{n}\Big)\qquad\text{for all } x>0\Rightarrow g(x)=\sum_{n\leqq x}\mu(n)\cdot G\Big(\frac{x}{n}\Big)\,,$$

as well as

(17) $$\sum_{n\leqq x}\frac{\mu(n)}{n}\log\frac{x}{n}=O(1)\,.$$

For a proof of (17), we make use of the fact that $\sum_{n\leqq x}x/n=x\cdot\log x+A_0x+O(1)$; thus, by the inversion formula,

$$x = \sum_{n \leq x} \mu(n) \cdot \frac{x}{n} \log \frac{x}{n} + A_0 \cdot \sum_{n \leq x} \mu(n) \cdot \frac{x}{n} + O(x) .$$

(17) follows now by (5).

If $f(x)$ is defined for $x > 0$, then

$$\sum_{n \leq x} \left\{ \frac{x}{n} \cdot \log \frac{x}{n} \cdot f\left(\frac{x}{n}\right) + \frac{x}{n} \cdot \sum_{p^m \leq x/n} \frac{\log p}{p^m} \cdot f\left(\frac{x}{n \cdot p^m}\right) \right\}$$

$$= \sum_{n \leq x} \frac{x}{n} \cdot \log \frac{x}{n} \cdot f\left(\frac{x}{n}\right) + \sum_{k \leq x} \frac{x}{k} \cdot f\left(\frac{x}{k}\right) \cdot \sum_{p^m/k} \log p \qquad (k = n \cdot p^m)$$

$$= \sum_{n \leq x} \frac{x}{n} \cdot \log \frac{x}{n} \cdot f\left(\frac{x}{n}\right) + \sum_{k \leq x} \frac{x}{k} \cdot f\left(\frac{x}{k}\right) \cdot \log k = x \cdot \log x \cdot \sum_{n \leq x} \frac{1}{n} \cdot f\left(\frac{x}{n}\right) .$$

Thus, if we set

$$F(x) \equiv x \cdot \log x \cdot \sum_{n \leq x} \frac{1}{n} \cdot f\left(\frac{x}{n}\right) ,$$

then, by the inversion formula,

$$x \cdot \log x \cdot f(x) + x \cdot \sum_{p^m \leq x} \frac{\log p}{p^m} \cdot f(x/p^m) = \sum_{n \leq x} \mu(n) \cdot F\left(\frac{x}{n}\right) .$$ [1]

In particular, if

$$\sum_{n \leq x} \frac{1}{n} \cdot f\left(\frac{x}{n}\right) = K + O\left(\frac{\log^k x}{x}\right) ,$$

then

$$\sum_{n \leq x} \mu(n) \cdot F\left(\frac{x}{n}\right) = K \cdot \sum_{n \leq x} \mu(n) \cdot \frac{x}{n} \cdot \log\left(\frac{x}{n}\right) + O\left(\sum_{n \leq x} \log^{k+1}\left(\frac{x}{n}\right)\right) = O(x) ,$$

by (17) and (3), and thus

$$(18) \qquad f(x) \cdot \log x + \sum_{p^m \leq x} \frac{\log p}{p^m} \cdot f(x/p^m) = O(1) ,$$

$$\text{if} \quad \sum_{n \leq x} \frac{1}{n} \cdot f\left(\frac{x}{n}\right) = K + O\left(\frac{\log^k x}{x}\right) .$$

(Selberg's asymptotic formula for $\psi(x)$ corresponds to $f(x) \equiv \psi(x)/x - 1$.) By (9′) and (12′), $f(x) \equiv b_x$ and $f(x) \equiv h_x$ both satisfy the condition of (18), and thus

$$(19) \qquad b_x \cdot \log x + \sum_{p^m \leq x} \frac{\log p}{p^m} \cdot b_{x/p^m} = O(1)$$

$$(20) \qquad h_x \cdot \log x + \sum_{p^m \leq x} \frac{\log p}{p^m} \cdot h_{x/p^m} = O(1) .$$

[1] Compare K. Iseki and T. Tatuzawa, "On Selberg's elementary proof of the prime number theorem." Proc. Jap. Acad. 27, 340–342 (1951).

From (16) and (20) it follows that

$$\sum_{n \leqq x} \frac{1}{n} \cdot b_n^2 \geqq |h_x| \cdot \log x + O(1) . \tag{21}$$

If we add to (19)

$$(\log x - A_0) \cdot \log x + \sum_{p^m \leqq x} \frac{\log p}{p^m} \cdot (\log (x/p^m) - A_0) ,$$

which by (1) and (6) is equal to $3/2 \cdot \log^2 x - 3 \cdot A_0 \cdot \log x + O(1)$, we obtain

$$\rho(x) \cdot \log x + \sum_{p^m \leqq x} \frac{\log p}{p^m} \cdot \rho(x/p^m) = \frac{3}{2} \cdot \log^2 x - 3 \cdot A_0 \cdot \log x + O(1) .$$

If $0 < c < 1$, and $c \cdot x < y < x$, then it follows from the last equation that

$$\rho(x) \cdot \log x - \rho(y) \cdot \log y \leqq \frac{3}{2} \cdot (\log^2 x - \log^2 y) + O(1)$$

$$= \frac{3}{2} \cdot \log \frac{x}{y} \cdot (\log x + \log y) + O(1) ,$$

$$\log x \cdot (\rho(x) - \rho(y)) + \log \frac{x}{y} \cdot \rho(y) \leqq \frac{3}{2} \cdot \log \frac{x}{y} \cdot (\log x + \log y) + O(1) ,$$

or, since $\rho(y) = \log y + O(1)$,

$$\log x \cdot (\rho(x) - \rho(y)) \leqq \log \frac{x}{y} \cdot \left(\frac{3}{2} \cdot \log x + \frac{1}{2} \cdot \log y\right) + O(1)$$

$$< 2 \cdot \log \frac{x}{y} \cdot \log x + O(1) .$$

Thus

$$\rho(x) - \rho(y) < 2 \cdot \log \frac{x}{y} + O\left(\frac{1}{\log x}\right) ,$$

and, since $\rho(x) = \log x - A_0 + b_x$, it follows that $b_x - b_y < \log x/y + O(1/\log x)$. Also obviously $b_x - b_y \geqq - \log x/y$, because $\rho(x)$ is non-decreasing. Thus we obtain

$$|b_x - b_y| \leqq \log \frac{x}{y} + O\left(\frac{1}{\log x}\right) \quad \text{if } c \cdot x < y < x,\ 0 < c < 1 . \tag{22}$$

Part III

Let $B \geqq 1$ be an upper bound of $|b_n|$.

Since $b_n - b_{n-1}$ is either $-\log [n/(n-1)]$, or $\log p/n - \log[n/(n-1)]$, it cannot happen that $b_n = b_{n-1} = 0$.

Let the integers $r_1, r_2, \cdots, r_t, \cdots$ be the indices n for which the b_n change signs. Precisely:

$$(23)\qquad \begin{cases} r_1 = 1\,;\ n = r_t \quad \text{if} \quad b_n \cdot b_{n+1} \leqq 0, \text{ and } b_{n+1} \neq 0\,; \\ \text{if} \quad r_t < v \leqq w < r_{t+1} \text{ then } b_v \cdot b_w > 0\,; \text{ and} \\ |\, b_{r_t} \,| < (\log r_t)/r_t \text{ for } t > 1. \end{cases}$$

Let $\{s_k\}$ be a sequence of integers, determined as follows: every r_t is an s_k; if $\log(r_{t+1}/r_t) < 7 \cdot B$, and $r_t = s_k$, then $r_{t+1} = s_{k+1}$; if $\log(r_{t+1}/r_t) \geqq 7 \cdot B$, enough integers s_{k+v} are inserted between $r_t = s_k$ and $r_{t+1} = s_{k+m}$ such that $3 \cdot B \leqq \log(s_{k+v+1}/s_{k+v}) < 7 \cdot B$, for $v = 0, 1, \cdots, m-1$. If there is a last $r_{t_0} = s_{k_0}$, a sequence $\{s_{k_0+v}\}$ is formed such that $3 \cdot B \leqq \log(s_{k_0+v+1}/s_{k+v}) < 7 \cdot B$. Thus the s_k form a sequence with the following properties:

$$(24)\qquad \begin{cases} s_1 = 1\,;\ \log(s_{k+1}/s_k) < 7 \cdot B\,;\ \text{for } k > 1, \textit{ either} \\ \log(s_{k+1}/s_k) \geqq 3 \cdot B, \textit{ or } |\, b_{s_k} \,| \text{ and } |\, b_{s_{k+1}} \,| \text{ are both} \\ \text{less than } \dfrac{\log s_k}{s_k}\,;\ b_v \cdot b_w > 0 \text{ for } s_k < v \leqq w < s_{k+1}\,. \end{cases}$$

Assume now that $\alpha\,(0 < \alpha < 1/2)$ is such that

$$(25)\qquad \textit{not} \quad h_x = O(\log^{-\alpha} x).$$

Then $|\, h_x \,| \cdot \log^\alpha x$ is unbounded. Let x be large, and such that $|\, h_x \,| \cdot \log^\alpha x \geqq |\, h_y \,| \cdot \log^\alpha y$ for all $y \leqq x$. Let c and d be positive integers such that

$$(26)\qquad s_{c-1} < \log x \leqq s_c\,, \quad \text{and} \quad s_d \leqq x < s_{d+1}\,.$$

It will be shown that

$$\frac{1}{2}\cdot(1 - \alpha - o(1)) \cdot S(x) \leqq |\, h_x \,| \cdot \log x \leqq \frac{1}{3} \cdot (1 + o(1)) \cdot S(x)\,,$$

where

$$(27)\qquad S(x) \equiv \sum_{k=c+1}^{d} |\, h_{s_k} - h_{s_{k-1}} \,| \cdot \log(s_k/s_{k-1})\,.$$

From this it will follow that $\alpha \geqq 1/3$.

Clearly

$$|\, h_x \,| \cdot \log x = |\, h_x \,| \cdot \log^\alpha x \cdot \left\{\log^{1-\alpha} x - \log^{1-\alpha} s_d + \sum_{k=2}^{d} (\log^{1-\alpha} s_k - \log^{1-\alpha} s_{k-1})\right\}$$

$$\geqq \frac{1}{2} \cdot \sum_{k=c+1}^{d} (|\, h_{s_k} \,| \cdot \log^\alpha s_k + |\, h_{s_{k-1}} \,| \cdot \log^\alpha s_{k-1}) \cdot (\log^{1-\alpha} s_k - \log^{1-\alpha} s_{k-1})$$

$$\geqq \frac{1}{2} \cdot \sum_{k=c+1}^{d} |\, h_{s_k} - h_{s_{k-1}} \,| \cdot \log^\alpha s_{k-1} \cdot (\log^{1-\alpha} s_k - \log^{1-\alpha} s_{k-1})\,.$$

If $y < z$, it is easily shown by the mean value theorem that

$$y^{\alpha} \cdot (z^{1-\alpha} - y^{1-\alpha}) > (1-\alpha) \cdot \frac{y}{z} \cdot (z-y) > \left(1 - \alpha - \frac{z-y}{z}\right)(z-y) .$$

With $y = \log s_{k-1}$, $z = \log s_k$, and from the fact that $s_k > \log x$, $\log (s_k/s_{k-1}) < 7 \cdot B$, it follows by (27) that

$$(28) \qquad | h_x | \cdot \log x > \frac{1}{2} \cdot \left(1 - \alpha - \frac{7 \cdot B}{\log \log x}\right) \cdot S(x) .$$

For the next estimate, we need the following lemma.

LEMMA. *Let v and w be positive integers such that*

(1) $\log \dfrac{w}{v} = O(1)$;

(2) $b_n > 0$ *for* $v \leqq n \leqq w$;

(3) $b_v < \dfrac{\log v}{v}$

Then

$$\sum_{v \leqq n \leqq w} \frac{1}{n} \cdot b_n^2 \leqq \frac{2}{3} \cdot \log \frac{w}{v} \cdot \sum_{v \leqq n \leqq w} \frac{1}{n} b_n + O\left(\frac{\log(w/v)}{\log v}\right) .$$

Proof. If $b_n \leqq 1/3 \cdot \log w/v$ for every n in $[v, w]$, the lemma is obviously correct. Otherwise, let n_1 be such that

$$b_{n_1} \geqq \frac{1}{3} \cdot \log \frac{w}{v} , \qquad b_n < \frac{1}{3} \cdot \log \frac{w}{v} \quad \text{for} \quad v \leqq n < n_1 .$$

If $\log (n_1/v) > 1/3 \log (w/v)$, let z ($v \leqq z < n_1$) be such that $\log (n_1/z) = 1/3 \log (w/v)$; otherwise, let $z = v$. Thus by (22), in every case, $\log (n_1/z) = 1/3 \log (w/v) + O(1/\log v)$. Clearly $b_n - 2/3 \cdot \log w/v < 0$ for $v \leqq n \leqq z$. Thus

$$\begin{aligned} T &\equiv \sum_{v \leqq n \leqq w} \frac{1}{n} \cdot b_n^2 - \frac{2}{3} \cdot \log \frac{w}{v} \cdot \sum_{v \leqq n \leqq w} \frac{1}{n} \cdot b_n \\ &\leqq \sum_{z \leqq n \leqq w} \frac{1}{n} \cdot b_n^2 - \frac{2}{3} \cdot \log \frac{w}{v} \cdot \sum_{z \leqq n \leqq w} \frac{1}{n} \cdot b_n , \end{aligned}$$

$$T \leqq \sum_{z \leqq n \leqq w} \frac{1}{n} \cdot \left(b_n - \frac{1}{3} \cdot \log \frac{w}{v}\right)^2 - \frac{1}{9} \cdot \log^2 (w/v) \cdot \log (w/z) + O\left(\frac{\log (w/v)}{v}\right) .$$

By (22),

$$\begin{aligned} \left| b_n - \frac{1}{3} \log (w/v) \right| &= | b_n - b_{n_1} | + O\left(\frac{\log v}{v}\right) \\ &\leqq | \log (n_1/n) | + O\left(\frac{1}{\log v}\right) = | \log (n_1/z) - \log (n/z) | + O\left(\frac{1}{\log v}\right) , \end{aligned}$$

and thus

$$\left| b_n - \frac{1}{3} \cdot \log \frac{w}{v} \right| \leqq \left| \log \frac{n}{z} - \frac{1}{3} \log \frac{w}{v} \right| + O\left(\frac{1}{\log v}\right).$$

Thus

$$\begin{aligned} T &\leqq \sum_{z \leqq n \leqq w} \frac{1}{n} \cdot \left(\log \frac{n}{z} - \frac{1}{3} \cdot \log \frac{w}{v}\right)^2 \\ &\quad - \frac{1}{9} \cdot \log^2 (w/v) \cdot \log (w/z) + O\left(\frac{\log (w/v)}{\log v}\right) \\ &= \sum_{z \leqq n \leqq w} \frac{1}{n} \cdot \log^2 (n/z) - \frac{2}{3} \cdot \log \frac{w}{v} \cdot \sum_{z \leqq n \leqq w} \frac{1}{n} \cdot \log \frac{n}{z} + O\left(\frac{\log (w/v)}{\log v}\right) \\ &= \frac{1}{3} \cdot \log^3 (w/z) - \frac{2}{3} \cdot \log^2 (w/v) \cdot \frac{1}{2} \cdot \log^2 (w/z) + O\left(\frac{\log (w/v)}{\log v}\right), \end{aligned}$$

by (2′), and thus $T \leqq O(\log (w/v)/\log v)$. This completes the proof of the lemma.

COROLLARY 1. *If condition* (3) *is replaced by* $b_w < \log w/w$, *the conclusion still holds; if* $b_n < 0$ *in* $v \leqq n \leqq w$, *the conclusion holds if* b_n *is replaced by* $|b_n|$.

COROLLARY 2. *If instead of* (3) *it is known that* $b_v < \log v/v$ *and* $b_w < \log w/w$ *then*

$$\sum_{v \leqq n \leqq w} \frac{1}{n} \cdot b_n^2 \leqq \frac{1}{3} \cdot \log \frac{w}{v} \cdot \sum_{v \leqq n \leqq w} \frac{1}{n} \cdot |b_n| + O\left(\frac{\log (w/v)}{\log v}\right).$$

For a proof, we split $[v, w]$ *into two intervals by a division point at* $(v \cdot w)^{1/2}$, *and apply the lemma separately to each subinterval.*

COROLLARY 3.

$$(29) \quad \sum_{s_{k-1} < n \leqq s_k} \frac{1}{n} \cdot b_n^2 \leqq \frac{1}{3} \cdot \log (s_k/s_{k-1}) \cdot \sum_{s_{k-1} < n \leqq s_k} \frac{1}{n} \cdot |b_n| + O\left(\frac{\log (s_k/s_{k-1})}{\log s_k}\right).$$

Proof. If $\log (s_k/s_{k-1}) < 3 \cdot B$, this follows from (24) and Corollary 2; if $\log (s_k/s_{k-1}) \geqq 3B$, it is obvious, since $|b_n| \leqq B$.

By (26), $\sum_{n \leqq s_c} 1/n \cdot b_n^2 = O(\log \log x)$, and $\sum_{s_d < n \leqq x} 1/n \cdot b_n^2 = O(1)$; also

$$\sum_{k=c+1}^{d} \frac{\log (s_k/s_{k-1})}{\log s_k} \leqq \sum_{k=c+1}^{d} \log \left(\frac{\log s_k}{\log s_{k-1}}\right) \leqq \log \log x.$$

It follows from (29) that

$$\sum_{n \leqq x} \frac{1}{n} \cdot b_n^2 \leqq \frac{1}{3} \cdot \sum_{k=c+1}^{d} \log (s_k/s_{k-1}) \cdot \sum_{s_{k-1} < n \leqq s_k} \frac{1}{n} \cdot |b_n| + O(\log \log x).$$

By (8′) $\sum_{s_{k-1} < n \leqq s_k} 1/n \cdot |b_n| = |h_{s_k} - h_{s_{k-1}}| + O(\log s_k/s_k)$, and thus, by (21) and (27),

$$|h_x| \cdot \log x \leqq \frac{1}{3} \cdot S(x) + O(\log\log x) \,. \tag{30}$$

It follows from (28) and (30) that

$$\left[\frac{1}{3} - \frac{1}{2}\cdot\left(1 - \alpha - \frac{7 \cdot B}{\log\log x}\right)\right] \cdot S(x) \geqq O(\log\log x) \,,$$

and since by (25) and (30) $S(x) \geqq K \cdot \log^{1/2} x$, this implies that $\alpha \geqq 1/3$. Thus $h_x = o(\log^{-1/3+\varepsilon} x)$, for every $\varepsilon > 0$, and therefore, by (8′),

$$\sum_{s_{k-1} < n \leqq s_k} \frac{1}{n} \cdot |b_n| = o(\log^{-1/3+\varepsilon} s_k) \,. \tag{31}$$

In order to find a bound for $|b_x|$, we consider now a particular interval $I_k = (s_{k-1}, s_k]$; let us assume that $b_n > 0$ in I_k. Let $n_2 \in I_k$ be such that $b_{n_2} \geqq b_n$ for every $n \in I_k$. Let n_1 $(s_{k-1} \leqq n_1 < n_2)$ be such that

$$b_{n_1} \leqq \frac{1}{2} \cdot b_{n_2} < b_{n_1+1} \,.$$

Then

$$\sum_{n \in I_k} \frac{1}{n} \cdot b_n > \sum_{n=n_1+1}^{n_2} \frac{1}{n} \cdot b_n > \frac{1}{2} \cdot b_{n_2} \cdot \log(n_2/n_1) - O(1/s_k) \,.$$

But by (22),

$$\log(n_2/n_1) \geqq b_{n_2} - b_{n_1} - O\left(\frac{1}{\log s_k}\right) \geqq \frac{1}{2} \cdot b_{n_2} - O\left(\frac{1}{\log s_k}\right) .$$

Thus

$$\sum_{n \in I_k} \frac{1}{n} \cdot b_n > \frac{1}{4} \cdot b_{n_2}^2 - O\left(\frac{1}{\log s_k}\right) .$$

It follows from (31) that $b_{n_2}^2 = o(\log^{-1/3+\varepsilon} n_2)$, and thus

$$b_x = o(\log^{-1/6+\varepsilon} x) \,. \tag{32}$$

Finally,

$$\begin{aligned} \psi(x) &= \sum_{n \leqq x} n \cdot (\rho(n) - \rho(n-1)) = [x] \cdot \rho([x]) - \sum_{n \leqq x-1} \rho(n) \\ &= x \cdot (\log x - A_0 + b_x) - \sum_{n \leqq x} (\log n - A_0 + b_n) + O(\log x) \\ &= x \cdot \log x - A_0 \cdot x + b_x \cdot x - x \cdot \log x + x + A_0 \cdot x - \sum_{n \leqq x} b_n + O(\log x) \\ &= x + o(x \cdot \log^{-1/6+\varepsilon} x) + o\left(\sum_{n \leqq x} \log^{-1/6+\varepsilon} n\right) , \quad \text{by (32).} \end{aligned}$$

The last sum is easily seen to be $o(x \cdot \log^{-1/6+\varepsilon} x)$, and thus

$$\psi(x) = x + o(x \cdot \log^{-1/6+\varepsilon} x) \,. \tag{33}$$

HÖLDER CONTINUITY OF N-DIMENSIONAL QUASI-CONFORMAL MAPPINGS

E. DAVID CALLENDER

1. Introduction and main results. This paper is an extension of previous work on the Hölder continuity of two-dimensional mappings. We shall use the approach of Finn and Serrin[1] and prove analogous results in n dimensions. A two-dimensional quasi-conformal mapping is one which carries infinitesimal circles into infinitesimal ellipses of bounded eccentricity. An n-dimensional quasi-conformal mapping carries infinitesimal spheres into infinitesimal ellipsoids of bounded eccentricity. Finn and Serrin gave an elementary proof that a quasi-conformal mapping is uniformly Hölder continuous in compact subdomains and obtained the best possible Hölder exponent. Their proof makes extensive use of the Dirichlet integral. We obtain similar results in n dimensions using a modified Dirichlet integral suggested by C. Loewner. It is not clear whether the n-dimensional exponent is the best possible one.

Let $u(x, y)$ and $v(x, y)$ be continuously differentiable functions in a domain D of the complex z-plane. Then the function $w(z) = u + iv$ represents a quasi-conformal mapping if there exists a constant K such that

$$(1) \qquad |\nabla w|^2 \equiv u_x^2 + u_y^2 + v_x^2 + v_y^2 \leq 2K(u_x v_y - u_y v_x) ,$$

for all points of the domain of definition of w. If $K < 1$, the mapping functions are constant; if $K = 1$, they are conformal. The only case of interest is $K \geq 1$. Geometrically, (1) implies that infinitesimal circles map into infinitesimal ellipses for which the ratio of minor to major axis $\geq K - \sqrt{K^2 - 1}$.

Let $f = (u_1, \cdots, u_n)$ be an n-dimensional mapping of a domain A of E_n into E_n such that f is continuously differentiable, the Jacobian, J, of the transformation is non-negative and

$$(2) \qquad |\nabla f|^2 \equiv \sum_{i,j=1}^{n} u_{i,j}^2 \leq nKJ^{2/n} , \qquad \text{where } u_{i,j} = \partial u_i / \partial x_j$$

and K is a constant holding for all points of the domain A of definition.

If $K < 1$, the mapping functions are constant, if $K = 1$, the mappings are the conformal mappings of space. Geometrically the mapping $x \to f(x)$ is sense preserving and infinitesimal spheres map onto infinitesimal ellipsoids. In this paper the norm used is the usual one for E_n

Received June 1, 1959. This work was supported in part by the Office of Naval Research.

[1] "On the Hölder Continuity of Quasi-Conformal and Elliptic Mappings." *Transactions of American Mathematical Society*, Vol. 89, No. 1 (1958), pp. 1–15. See this paper for a bibliography of previous work.

and is denoted by $|x|$.

Finn and Serrin treat a class of mappings which they call elliptic mappings. This generalization of the notion of quasi-conformal mapping is due to L. Nirenberg. $w(z)$ is an ellistic mapping if it satisfied the conditions for a quasi-conformal mapping except that condition (1) is replaced by

$$|\nabla w|^2 \leq 2KJ + K_1 , \tag{3}$$

where K and K_1 are constant, $K \geq 1$ and $K_1 \geq 0$. A generalization of two-dimensional elliptic mappings is obtained by replacing condition (2) in the definition of n-dimensional quasi-conformal mappings by

$$|\nabla f|^n \leq (nK)^{n/2} J + K_1 , \tag{4}$$

where K and K_1 are constants, $K \geq 1$, and $K_1 \geq 0$. Such mappings we shall call near quasi-conformal mappings.

In two dimensions many important estimates are given in terms of the Dirichlet integral

$$D(r) = \iint_{C_r} |\nabla w|^2 \, dxdy ,$$

where C_r is a circle of radius r. We shall find that the appropriate n-dimensional analog of this integral is

$$D(r) = \int_{S_r} \left\{ \sum_{i,j=1}^{n} u_{i,j}^2 \right\}^{n/2} dV , \tag{5}$$

where S_r is an n-dimensional hypersphere of radius r. This integral was suggested by C. Loewner in a paper that will appear shortly in the Journal of Mathematics and Mechanics.

The proofs of Finn and Serrin make use of Morrey's lemma, which is based on the usual Dirichlet integral. By means of the modified Dirichlet integral, an analogous lemma is proved in n dimensions.

For the n-dimensional quasi-conformal mappings and the near quasi-conformal mappings the following two theorems are proved.

THEOREM 1. *Let f be a quasi-conformal mapping defined in a domain A of E_n. Assume $|f| \leq 1$. Then in any compact subregion B of A,*

$$|f(x_1) - f(x_2)| \leq C \frac{|x_1 - x_2|^\mu}{d^\mu} , \tag{6}$$

where d is the distance from B to the boundary of A; $\mu = \mu(n, K)$ and $0 < \mu \leq 1$; and $C = C(n, K)$, a constant depending only on the dimension of the space and K. (See equation (12) for definition of μ.)

THEOREM 2. *Let f be a near quasi-conformal mapping defined on a domain A of E_n. Let $|f| \leq 1$. Then in any compact subregion B of A*

$$|f(x_1) - f(x_2)| \leq H|x_1 - x_2|^{\mu}, \tag{7}$$

where H is a constant depending on n, K, K_1, and d (d is the distance from B to boundary of A) *and $\mu = \mu(n, K)$, $0 < \mu \leq 1$. μ is the same constant that appears in Theorem* 1.

2. Preliminary lemmas. To generalize the proofs of Finn and Serrin to n dimensions, several lemmas are needed. They are listed below and the more difficult proofs are given.

LEMMA 1. *The weak Maximum Principle holds for quasi-conformal mappings, i.e., if f is quasi-conformal in a bounded region A and continuous in $\bar{A}$, then the maximum of the norm* (and of the components) *is attained on the boundary $\dot{A}$ of A. The minimum of the components is also attained on $\dot{A}$.* (The proof is the same as in two dimensions.)

LEMMA 2. *Let u be a function defined in some domain A. If $u = 0$ on $\dot{S}_r$ where $\dot{S}_r$ is the surface of a sphere of radius r in A and n is the dimension of the space, then*

$$\int_{\dot{S}_r} |u|^n dA \leq Cr^n \int_{\dot{S}_r} |u_t|^n dA, \tag{8}$$

where u_t is the tangential component of the gradient of u on S_r and C is a constant depending only on the dimension of the space.

LEMMA 3. *For all $a, b \geq 0$, $\lambda > 0$ and $n \geq 2$,*

$$\frac{n}{(n-1)^{\frac{n-1}{n}}} a^{1/n} b^{\frac{n-1}{n}} \leq \frac{a}{\lambda^{n-1}} + \lambda b, \tag{9}$$

and the constant of this inequality cannot be improved.

LEMMA 4. *Let u be a function defined in a domain A and let $\omega = \omega(\dot{S}_r)$ be the oscillation of u on the surface of sphere of radius r in A where n is the dimension of the space. Then there exists a constant C depending only on the dimension of the space such that*

$$\frac{\omega^n(\dot{S}_r)}{r} \leq C \int_{\dot{S}_r} |u_t|^n dA. \tag{10}$$

LEMMA 5. *Let (a_{ij}) be an $n \times n$ matrix with real coefficients. Then*

$$(11) \qquad |\det(a_{ij})| \leq \frac{1}{n^{n/2}}\left(\sum_{i,j=1}^{n} a_{ij}^2\right)^{n/2}.$$

The constant in the inequality cannot be approved. (This lemma follow immediately from the proof of Hadamard's inequality.)

Morrey's lemma in n dimensions. Let B be a closed subregion of D and let $d =$ distance $(\dot{B}, \dot{D})$. Suppose there exist constants L, μ, r_0, where $0 < \mu$ and $r_0 \leq d$, such that for all spheres S_r with center in B, $r \leq r_0$,

$$D(r) \equiv \int_{S_r} |\nabla f|^n dV \leq Lr^{n\mu},$$

Then f satisfies a Hölder condition in B:

$$|f(x_1) - f(x_2)| \leq C_2 |x_1 - x_2|^\mu,$$

where

$$C_2 = \frac{1}{\pi^{n-1}}\left(\frac{nL}{n-1}\right)^{1/n}\left(\frac{2\pi C_1(n-1)}{\mu}\right)^{\frac{n-1}{n}}$$

and $C_1 = C_1(n)$.

Proof of Lemma 2. Let $n \geq 3$. Choose the coordinates such that $u = 0$ at the north pole. For given $(\theta_2, \cdots, \theta_{n-1})$, let $u_m = u_m(\theta_2, \cdots, \theta_{n-1})$ be the maximum of $|u|$ for $0 \leq \theta_1 \leq \pi$. We have $u = \int_0^\theta u_\theta d\theta$, which implies that

$$u_m \leq \int_0^\pi |u_\theta| \, d\theta \leq \left[\int_0^\pi |u_\theta|^n r^{n-1} \sin^{n-2}\theta_1 d\theta_1\right]^{1/n}\left[\int_0^\pi \frac{d\theta_1}{r\sin^{\frac{n-2}{n-1}}\theta_1}\right]^{\frac{n-1}{n}}$$

by Hölder's inequality. Let

$$C = \int_0^\pi \frac{d\theta_1}{\sin^{\frac{n-2}{n-1}}\theta_1}.$$

$C < \infty$. Hence

$$\begin{aligned} u_m^n r^{n-1} &\leq C^{n-1}\int_0^\pi |u_\theta|^n r^{n-1}\sin^{n-2}\theta_1 d\theta_1 \\ &\leq r^n C^{n-1}\int_0^\pi |u_t|^n r^{n-1}\sin^{n-2}\theta_1 d\theta_1 . \end{aligned}$$

Now

$$\int_{\dot{S}_r} |u|^n dA = r^{n-1}\int_0^{2\pi}\underbrace{\int_0^\pi \cdots \int_0^\pi}_{n-2} |u|^n \sin^{n-2}\theta_1 d\theta_1 d\omega_{n-1}$$

$$\le r^{n-1}\int u_m^n \sin^{n-2}\theta_1 d\theta_1 d\omega_{n-1}$$

$$\le r^{n-1}\sqrt{\pi}\,\frac{\Gamma\{(n-1)/2\}}{\Gamma(n/2)}\int u_m^n d\omega_{n-1} .$$

Combining the above results

$$\int_{\dot S_r} |u^n|\, dA \le \sqrt{\pi}\,\frac{\Gamma\{(n-1)/2\}}{\Gamma(n/2)}\, C^{n-1} r^n \int_{\dot S_r} |u_t|^n dA .$$

Proof of Morrey's lemma. Denote the points x_1, x_2 by P and 0, respectively. Let $|x_1 - x_2| \le r_0$ and let $r = |x_1 - x_2|$. Let M be a perpendicular bisector of $\overline{PQ}$. Select a point S on M such that $\overline{PS} = \overline{QS} \le \overline{PQ} \le r_0$. Then

$$f(P) - f(Q) = \int_{PS} f_r dr - \int_{QS} f_r dr$$

which implies

$$|f(P) - f(Q)| \le \int_{PS} |f_r|\, dr + \int_{QS} |f_r|\, dr .$$

Hence

$$\int_0^{2\pi}\underbrace{\int_0^{\pi}\cdots\int_0^{\pi}}_{n-3}\int^{\pi/3} |f(P) - f(Q)|\, d\theta_1 \cdots d\theta_{r-1}$$

$$\le 2\int_0^R\int_0^{2\pi}\int_0^{\pi}\cdots\int_0^{\pi}\int^{\pi/3} |\nabla f|\, d\theta_1 \cdots d\theta_{n-1} dr .$$

So

$$|f(P) - f(Q)| \le \frac{3}{\pi^{n-1}}\int_{S_r} |\nabla f|\, d\theta_1 \cdots d\theta_{n-1} dr$$

$$\le \frac{3}{\pi^{n-1}}\left[\int_{S_r} |\nabla f|^n r^{n-1-\mu} \sin^{n-2}\theta_1 \cdots \sin\theta_{n-2} dr d\theta_1 \cdots d\theta_{n-1}\right]^{1/n} \cdot I^{\frac{n-1}{n}} ,$$

where

$$I = \int_{S_r} r^{-1+\frac{\mu}{n-1}} \sin^{-\left(\frac{n-2}{n-1}\right)}\theta_1 \sin^{-\left(\frac{n-3}{n-1}\right)}\theta_2 \cdots \sin^{-\frac{1}{n-1}}\theta_{n-2} dr d\theta_1 \cdots d\theta_{n-1} .$$

$$I = r^{\frac{\mu}{n-1}}\left(\frac{n-1}{\mu}\right) 2\pi C_1$$

where

$$C_1 = \int_0^{\pi}\cdots\int_0^{\pi} \sin^{-\left(\frac{n-2}{n-1}\right)}\theta_1 \sin^{-\left(\frac{n-3}{n-2}\right)}\theta_2 \cdots \sin^{-\frac{1}{n-1}}\theta_{n-2} d\theta_1 \cdots d\theta_{n-2} < \infty .$$

$$\int_{S_r} |\nabla f|^n r^{n-1-\mu} d\omega_n = r^{-\mu} D(r) \,,$$

$$u\int_0^r D(r) r^{-\mu-1} dr \le r^{-\mu} L r^{n\mu} + \mu L \int_0^r r^{\mu(n-1)-1} dr = \frac{n}{n-1} L r^{(n-1)\mu} \,,$$

since by hypothesis

$$D(r) \le L r^{n\mu} \,.$$

Combining

$$|f(P) - f(Q)| \le \frac{3}{\pi^{n-1}} \left(\frac{n}{n-1} L r^{(n-1)\mu} \right)^{1/n} \left\{ r^{\frac{\mu}{n-1}} \left(\frac{n-1}{\mu} \right) (2\pi C_1) \right\}^{\frac{n-1}{n}} = C_2 r^{\mu} \,,$$

where

$$C_2 = \frac{3}{\pi^{n-1}} \left(\frac{nL}{n-1} \right)^{1/n} \left(\frac{2\pi C_1 (n-1)}{\mu} \right)^{\frac{n-1}{n}} \,.$$

Proof of Lemma 4.[2] The surface of the n dimensional hypersphere of radius r can be mapped onto a $n-1$ dimensional hyperplane by a stereographic mapping. Under such a transformation

$$\int_{\dot{S}_r} |u_t|^{n-1} dS = \int_V |\nabla u|^{n-1} dV \,,$$

and

$$\int_{\dot{S}_r} (1 - \cos\theta_1) |u_t|^n dS = \int_V |\nabla u|^n dV \,,$$

where the variables on the surface of the sphere are $(\theta_1, \theta_2, \cdots, \theta_{n-1})$, on the hyperplane are $(\rho, \theta_2, \cdots, \theta_{n-1})$, and domains of integration are mapped onto one another. Hence

$$\int_V |\nabla u|^n dV \le \int_{\dot{S}_r} |u_t|^n dS \,.$$

In the hyperplane

$$\int_{|x|\le\rho} |\nabla u|^{n-1} dV \le \left[\int |\nabla u|^n dV \right]^{\frac{n-1}{n}} \left[\int dV \right]^{1/n}$$
$$\le C_1 \left[\rho \int_{\dot{S}_r} |u_t|^n dS \right]^{\frac{n-1}{n}} \,,$$

where latter integration is taken over the whole surface of the n dimensional hypersphere and

[2] The author is indebted to R. Finn for suggesting this proof which strengthens and simplifies the author's original proof.

$$c_1^n = \int_0^\pi \cdots \int_0^\pi \int_0^{2\pi} \sin^{n-3}\theta_2 \sin^{n-4}\theta_3 \cdots \sin\theta_{n-2}\, d\theta_1 \cdots d\theta_{n-1} .$$

Hence by Morrey's lemma applied in the $n-1$ dimensional hyperplane

$$\omega^n \leq C\rho \int_{\dot{S}_r} |u_t|^n dS ,$$

where C is a constant depending only on n. It follows immediately that

$$\frac{\omega^n(\dot{S}_r)}{r} \leq 2C \int_{\dot{S}_r} |u_t|^n dS .$$

3. Proof of Theorems 1 and 2. The proof of Theorem 2 will be given before that of Theorem 1, and Theorem 1 will follow as an immediate corollary of Theorem 2. Then an alternate method of proof for Theorem 1 will be given. This second proof uses a modulus of continuity instead of Morrey's lemma.

Proof of Theorem 2. It must be shown that it f is a near quasi-conformal mapping, then $D(r) \leq \text{constant}\cdot r^{n\mu}$ for r sufficiently small. Then the conclusion will follow by Morrey's n dimensional lemma. By (4)

$$D(r) \leq (nK)^{n/2} \int_{S_r} J dV + \omega_n r^n K_1 .$$

$$\int_{S_r} J dV = \int_{\dot{S}_r} u_1 du_2 \cdots du_n = \int (u_1 - \bar{u}_1) du_2 \cdots du_n$$
$$= \int_{\dot{S}_r} (u_1 - \bar{u}_1) \frac{\partial(u_2 \cdots u_n)}{\partial(s_2 \cdots s_n)} ds_2 \cdots ds_n ,$$

where $\bar{u}_1$ is the mean value of u_1 over S_r, $ds_2 = r d\theta_1$, $ds_3 = s \sin\theta_1 d\theta_2$, $ds_4 = r \sin\theta_1 \sin\theta_2 d\theta_3, \cdots,$ and $ds_n = r \sin\theta_1 \cdots \sin\theta_{n-2} d\theta_{n-1}$. Hence by Lemma 5, Lemma 3, and Lemma 2

$$\int_{S_r} J dV \leq \frac{1}{(n-1)^{\frac{n-1}{2}}}$$
$$\times \int_{\dot{S}_r} |u - \bar{u}_1| [u_{2,s_2}^2 + \cdots + u_{2,s_n}^2 + u_{3,s_2}^2 + \cdots + u_{n,s_n}^2]^{\frac{n-1}{2}} dA$$
$$\leq \frac{(n-1)^{\frac{n-1}{n} - \frac{n-1}{2}}}{n} \left\{ \int_{\dot{S}_r} \frac{|u_1 - \bar{u}_1|^n}{r^{n-1}} dA + r \int_{\dot{S}_r} [u_{2,s_2}^2 + \cdots + u_{n,s_n}^2]^{n/2} dA \right\}$$
$$\leq \frac{(n-1)^{\frac{n-1}{n} - \frac{n-1}{2}}}{n} r \left\{ C \int_{\dot{S}_r} |u_{1,t}|^n dA + \int_{\dot{S}_r} [u_{2,s_2}^2 + \cdots + u_{n,s_n}^2]^{n/2} dA \right\} ,$$

where $C = C(n)$ is the constant of Lemma 2. Hence

$$\int_{S_r} J dV \leq C' r \int_{\dot{S}_r} |f_t|^n dA ,$$

where

$$C' = \frac{(n-1)^{\frac{n-1}{n} - \frac{n-1}{2}}}{n} C ,$$

and finally

$$\int_{S_r} J dV \leq C' r \frac{dD}{dr} .$$

The Hölder exponent μ is defined by the equation

(12) $$\frac{1}{\mu} = C n^{n/2} K^{n/2} (n-1)^{-\frac{n^2-n+2}{2n}} ,$$

where C is the constant of Lemma 2.

Combining above results

(13) $$D(r) \leq \frac{r}{n\mu} \frac{dD}{dr} + \omega_n r^n K_1 ,$$

where ω_n is the area of the unit sphere in n dimensions.

Let B be a closed subregion of A, and let d be the distance from B to $\dot{A}$. Let S_r be a sphere whose center is in B. For such a sphere

$$D(r) \leq \frac{r}{n\mu} \frac{dD}{dr} + \omega_n r^n K_1 , \qquad \text{for } 0 < r < d.$$

Hence

$$-\frac{d}{dr}(r^{-n\mu} D) \leq n\mu K_1 \omega_n r^{n-1-n\mu} ,$$

and integrating

(14) $$D(\rho) \leq \{D(t) + K_2\} \left(\frac{\rho}{t}\right)^{n\mu} , \qquad \rho \leq t \leq d$$

where

$$K_2 = \frac{\pi K_1}{1-\mu} \omega_n t^n .$$

We now wish to estimate $D(t)$. We know

$$D(t) \leq \frac{(nK)^{n/2}}{(n-1)^{\frac{n-1}{2}}} \int_{\dot{S}_r} |u_1| [u_{2,s_2}^2 + \cdots + u_{n,s_n}^2]^{\frac{n-1}{2}} dA + r^n \omega_n K_1$$

$$\leq K_3\left[\int_{\dot{S}_r} |u_1|^n dA\right]^{1/n}\left[\int_{\dot{S}_r} |f_t|^n dA\right]^{\frac{n-1}{n}} + r^n\omega_n K_1$$

$$\leq K_4\left[r\int_{\dot{S}_r} |f_t|^n dA\right]^{\frac{n-1}{n}} + r^n\omega_n K_1$$

$$\leq K_4\left[r\frac{dD}{dr}\right]^{\frac{n-1}{n}} r^n\omega_n K_1 ,$$

where

$$K_3 = \frac{(nK)^{n/2}}{(n-1)^{\frac{n-1}{2}}} \quad \text{and} \quad K_4 = K_3\omega_n^{1/n} .$$

We have also used the fact that $|f| < 1$. The immediately preceeding result implies

$$(D(r) - K_5)^{\frac{n}{n-1}} \leq K_4^{\frac{n}{n-1}} r\frac{dD}{dr} ,$$

for $r < d$ and where $K_5 = d^n\omega_n K_1$.

Now suppose $D(t) > K_5$ for some t. Then $D(r) > K_5$ for all $r \geq t$. Hence

$$(D(t) - K_5)^{\frac{1}{n-1}} \leq \frac{K_4^{\frac{n}{n-1}}(n-1)}{\log(d/t)} ,$$

or

$$D(t) - K_5 \leq \frac{K_6}{\{\log(d/t)\}^{n-1}} ,$$

where

$$K_6 = K_4^n(n-1)^n .$$

So

$$D(t) \leq \frac{K_6}{\{\log(d/t)\}^{n-1}} + d^n\omega_n K_1 . \tag{15}$$

This inequality also holds if $D(r) \leq K_5$.

Now let $t = de^{-\nu}$ where $\nu = 1/n\mu$. Combining (14) and (15) we obtain $D(\rho) \leq H\rho^{n\mu}$ where H is a constant depending only on n, K, K_1, and d. We can now conclude that for $x_1, x_2 \in D$ and $|x_1 - x_2| \leq de^{-\nu}$ that $|f(x_1) - f(x_2)| \leq H|x_1 - x_2|^{\mu}$. Because of the bound on $|f|$, we get a similar result when $|x_1 - x_2| > de^{-\nu}$.

Alternate proof of Theorem 1. Here we do not use Morrey's lemma, instead a modulus of continuity on f is obtained in terms of $D(r)$.

Proof. For any point set $T \subset A$, let $\omega(t) = \text{l.u.b.}\, |f(x_1) - f(x_2)|$ where $x_1 x_2 \in T$. Since f is quasi-conformal, it satisfies the weak maximum principle. Let S_r be a sphere of radius r such that its center is at least a distance ρ from A. Then $\omega(S_r) = \omega(\dot{S}_r)$. By Lemma 4,

$$\frac{\omega^n(S_s)}{r} \leq C\frac{dD}{dr} \qquad \text{for } s \leq r.$$

Hence

$$\omega^n(S_s) \log \frac{\rho}{s} \leq CD(\rho) \qquad \text{for } s < \rho.$$

This implies

$$\omega(S_s) \leq \left[\frac{CD(\rho)}{\log(\rho/s)}\right]^{1/n},$$

where C depends only on the dimension of the space.

$D(\rho)$ can be estimated by the technique used in the proof of Theorem 2.

$$D(\rho) \leq e(n\mu)^{n-1}(nK)^{n^2/2}\omega_n(n-1)^{\frac{n+1}{2}}\left(\frac{\rho}{d^*}\right)^{n\mu},$$

where e is the base of the natural logs, μ is defined as in proof of Theorem 2 and $\rho \leq d^*e^{-\nu}$. This is valid for all spheres of radius ρ whose centers are at least a distance d^* from $\dot{A}$.

Let x_1 and x_2 be two points in B such that $|x_1 - x_2| = 2s < de^{-n\nu} = de^{-1/\mu}$. The midpoint of the line segment $\overline{x_1x_2}$ is at least a distance $d^* = d/2$ from $\dot{A}$. Consequently

$$|f(x_1) - f(x_2)| \leq \omega(S_s) \leq \frac{C(n, K)}{[\log(\rho/s)]^{1/n}}\left(\frac{\rho}{d^*}\right)^{\mu}$$

for $s \leq \rho \leq d^*e^{-\mu}$.

Let

$$\rho = se^{\nu}.$$

Then

$$|f(x_1) - f(x_2)| \leq C(n, K)\frac{|x_1 - x_2|^{\mu}}{d^{\mu}}.$$

On the other hand, if $|x_1 - x_2| \leq de^{-\nu\mu}$, we again get a Hölder estimate since $|f| \leq 1$.

4. Additional results. Theorems 3 and 4 are on removable singularities. The final theorem is concerned with one-to-one quasi-conformal mappings.

THEOREM 3. *Let $f(x)$ satisfy the hypothesis of Theorem 1 or 2 for all points x in the domain A except on a set T of isolated points in A. Then f can be defined at the points of T such that the resulting function is continuous in A and satisfies the conclusion of Theorems 1 or 2.*

Proof. To prove Theorem 3 it is sufficient to show that $D(r)$ exists and satisfies

$$D(r) \leq (nK)^{n/2}\int_{\dot{S}_r} u_1 du_2 \cdots du_n + \omega_n r^n K_1 \,,$$

for all spheres whose surface contains no points of T. Then all the previous statements are valid and hence f satisfies a Hölder condition in $B - T$. Finally f can be defined on T such that resulting function is continuous in A and satisfies a Hölder condition throughout A.

Let S be a sphere of radius r. Let S_r contain exactly one point x_0 of T. Let S^σ be a sphere of radius σ with center x_0.

$$D(\sigma, r) \equiv \int_{S_r - S_\sigma} |\nabla f|^n dV \,.$$

Hence

$$D(\sigma, r) \leq -(nK)^{n/2}\int_{\dot{S}_r} u_1 du_2 \cdots du_n + I \,,$$

when

$$I = (nK)^{n/2}\int_{\dot{S}_r} u_1 du_2 \cdots du_n + \omega_n r^n K_1 \,. \qquad |f| \leq 1.$$

Hence

$$(D(\sigma, r) - I) \leq K_4\left[-\sigma\frac{dD}{d\sigma}\right]^{\frac{n-1}{n}} ,$$

which implies

$$(D(\sigma, r) - I)^{\frac{n}{n-1}} \leq K_4^{\frac{n}{n-1}}\left(-\sigma\frac{dD(\sigma, r)}{d\sigma}\right) .$$

Suppose $D > I$ for some value of σ, say $\sigma = \sigma_2$. Then $D > I$ for all $\sigma < \sigma_2$. There we may integrate from σ_1 to σ_2 and obtain

$$\log\frac{\sigma_2}{\sigma_1} \leq \frac{C(n, K)}{(D(\sigma_2, r) - I)^{\frac{1}{n-1}}} \,.$$

Let σ_1 approach zero. A contradiction is then obtained. Therefore $D(\sigma, r) \leq I$. Let σ approach zero, and we obtain $D(r) \leq I$.

Since there at most a finite number of points of T in any compact subset of A, the desired result can be obtained.

THEOREM 4. *Let f be a continuously differentiable function defined in the region $0 < |x| \leq 1$. Suppose that*

$$|\nabla f|^n \leq (nK)^{n/2}J + K_1 |x|^{-n\lambda},$$

where K, K_1, and d are constants such that $K \geq 1$, $K_1 \geq 0$ and $0 \leq \lambda > 1$. Also assume $u_1 = o(|x|^{-\mu})$ as $x \to 0$ where $\mu = \mu(n, K)$ as defined in Theorems 1 and 2. Then w can be defined at $x = 0$ such that the resulting function is continuous in $0 \leq |x| \leq 1$, and in any closed subregion of $|x| < 1$, f satisfies a uniform Hölder condition with exponent μ.

Proof. If S_r is any sphere in $|x| \leq 1$ whose surface does not contain the origin, the $D(r)$ exists and satisfies

$$D(r) \leq (nK)^{n/2}\int_{\dot{S}_r} u_1 du_2 \cdots du_n + K_1\int_{S_r} \|x\|^{-n\lambda} dV . \tag{16}$$

If S_r does contain the origin, then let S_σ denote a sphere of radius σ and center $x = 0$.

Then as in proof of Lemma 3,

$$D(\sigma, r) \leq -(nK)^{n/2}\int_{\dot{S}_\sigma} u_1 du_2 \cdots du_n + B , \tag{17}$$

where B denotes the right hand side of (16).

By hypotheses

$$|u_1|^n \leq \varepsilon(|x|)|x|^{-n\mu} ,$$

where

$$\varepsilon(|x|) \to 0 \quad \text{as} \quad |x| \to 0 .$$

Without loss of generality we may assume the $\varepsilon(|x|)$ is monotonically increasing.

$$\begin{aligned}(D(\sigma, r) - J)^{\frac{n}{n-1}} &\leq (nK)^{\frac{n^2}{2(n-1)}}\left[\int_{\dot{S}_\sigma} u_1 du_2 \cdots dn_n\right]^{\frac{n}{n-1}} \\ &\leq -C(n, K)\varepsilon(\sigma)\sigma^{1-n\mu}\frac{dD}{d\sigma} .\end{aligned}$$

Now suppose $D > J$ for $\sigma = \sigma_0$. Then $D > J$ for all $\sigma < \sigma_0$. Hence

$$\sigma^{n\mu-1} \leq -\frac{C(n, K)\varepsilon(\sigma)}{(D - J)^{\frac{n}{n-1}}}\frac{dD}{d\sigma} \qquad \text{for } \sigma < \sigma_0.$$

Integrate from σ_1 to σ_2 where $\sigma_1 < \sigma_2 < \sigma_0$,

$$\frac{1}{n\mu}[\sigma_2^{n\mu} - \sigma_1^{n\mu}] \leq C(n, K)\varepsilon(\sigma_2)(D - J)^{-\frac{1}{n-1}} .$$

Let σ_1 approach zero. Hence

$$(D - B)^{\frac{1}{n-1}} \leq C(n, K)\varepsilon(\sigma)\sigma^{-n\mu} . \tag{18}$$

As in the proof of Theorem 2, the inequality of the hypothesis implies

$$D - B \leq -\frac{\sigma}{n\mu}\frac{dD}{d\sigma} + \frac{\omega_n K_1}{n - n\lambda}\sigma^{n-n\lambda} .$$

It follows that

$$-\frac{d}{d\sigma}[\sigma^{-n\mu}(D(\sigma, r) - B(r)] \leq C(n, K, \lambda)\sigma^{-n\lambda-n\mu+n-1} .$$

Hence, for $\sigma_1 < \sigma$,

$$\sigma_1^{-n\mu}(D(\sigma_1, r) - B) \leq \sigma^{-n\mu}(D(\sigma, r) - B) + C(n, K, \lambda)\sigma^{n(1-\mu-\lambda)} ,$$

and finally

$$D(\sigma, r) - B \geq [D(\sigma_1, r) - B - C\sigma^{n(1-\mu-\lambda)}\sigma_1^{n\mu}]\frac{\sigma^{n\mu}}{\sigma_1^{n\mu}} .$$

Let $\sigma < \sigma_0$. For fixed σ, σ_1 may be chosen small enough such that

$$D(\sigma_1, r) - B - C\sigma^{n(1-\mu-\lambda)}\sigma_1^{n\mu} > 0 .$$

For small enough σ this contradicts (16). Hence $D(\sigma, r) \leq B$ which implies $D(r) \leq B$.

Now proceed as in the proof of Theorem 2. Let B be an arbitrary compact subregion of $|x| < 1$ and let d = distance from B to $|x| = 1$. For any sphere with center in B,

$$D(r) \leq \frac{r}{n\mu}\frac{dD}{dr} + K_1\int_{S_r}\rho^{-n\lambda}dV .$$

This implies

$$-\frac{d}{dr}(r^{-n\mu}D) \leq K_1 r^{-n\mu-1}\int_{S_r}\rho^{-n\lambda}dV = \frac{K_1\omega_n}{n - n\lambda}r^{-1+n(1-\mu-\lambda)} .$$

Integrating from ρ to d,

$$\rho^{-n\mu}D(\rho) \leq d^{-n\mu}D(d) + Cd^{n(1-\mu-\lambda)} .$$

Note that $D(\rho)$ is bounded by

$$(nK)^{n/2}\int_{S_1}JdV + K_1\int_{S_1}\rho^{-n\lambda}dV < \infty .$$

So $D(\rho) \leq$ constant $\rho^{n\mu}$.

By Morrey's n-dimensional lemma, f is uniformly Hölder continuous on B with exponent μ.

THEOREM 5. *Let $f(x)$ be a one-to-one quasi-conformal mapping of $|x| < 1$ onto $|f| < 1$ and such that $f(0) = 0$. The f can be extended to a one-to-one continuous mapping of $|x| \leq 1$ onto $|f| \leq 1$ satisfying $|f(x_1) - f(x_2)| \leq H|x_1 - x_2|^\mu$ where $H = H(n, K)$ and $\mu = \mu(n, K)$. $0 < \mu < 1$.*

The proof of this theorem is an immediate generalization of the proof of the 2-dimensional theorem of the Finn and Serrin paper. All new ideas have already been introduced. Hence the proof will not be given.

5. **Weakened differentiability requirements.** The previous theorems remain true if instead of $f \in C^1$ and $|\Delta f|^2 \leq nKJ^{2/n}$, f satisfies

(i) $f \in C$ in A, $f = (u_1, u_2, \cdots, u_n)$,

(ii) u_i is absolutely continuous in x_j for almost all values of the other $n-1$ variables $i, j = 1, \cdots, n$,

(iii) the derivatives $u_{i,j}$ (which exist almost everywhere by (ii) should be nth integrable,

(iv) $|\nabla f|^2 \leq nKJ^{2/n}$ almost everywhere or $|\nabla f|^n \leq (nK)^{n/2}J + K_1$ almost everywhere.

To prove the above theorems it suffices to show that the following inequalities on the growth of the modified Dirichlet integral of f remain valid under the weakened hypotheses

$$D(\rho) \leq \{D(t) + K_2\}\left(\frac{\rho}{t}\right)^{n\mu}, \tag{19}$$

for $\rho \leq t \leq d$ and $K_2 = \dfrac{\mu K_1}{1-\mu}\omega_n t^n$.

$$D(t) \leq \frac{K_6}{\{\log(d/t)\}^{n-1}} + d^n\omega_n K_1 , \tag{20}$$

for $t < d$ and where

$$K_6 = \omega_n(nK)^{n/2}(n-1)^{\frac{n+1}{2}} .$$

We shall prove (19) in the case where $|\nabla f|^2 \leq nKJ^{2/n}$. The other statements are proved in a similar manner.

Let f be approximated in the nth integral norm of its derivative by a sequence of functions $f^{(h)} \in C^1$. Thus $\int_A |\nabla(f - f^{(h)})|^n dV$ and $\sup_A |f - f^{(h)}|$ approach zero as h approaches zero. For $f^{(h)}$, (let $J^{(h)}$ be its Jacobian), $Q^{(h)}$ is defined to be $\int_S J^{(h)} dV$.

$$\int_{S_r} J^{(h)} dV \leq \frac{1}{(n-1)^{\frac{n-1}{2}}} \int_{\dot{S}_r} |u_1| [u^2_{2,s_2} + \cdots + u^2_{n,s_n}]^{\frac{n-1}{2}} dA$$
$$\leq \frac{\omega_n^{1/n}(1+\varepsilon)}{(n-1)^{\frac{n-1}{2}}} \left[r\frac{dD^{(h)}}{dr} \right]^{\frac{n-1}{n}},$$

since $|f| \leq 1$. ε approaches zero as h approaches infinity, and

$$D^{(h)}(r) = \int_{S_r} |\nabla f^{(h)}|^n dV .$$

Hence

$$\int_r^{r+\lambda} [Q^{(h)}(\rho)]^{\frac{n-1}{n}} d\rho \leq \left(\frac{1+\varepsilon}{(n-1)^{\frac{n-1}{2}}} \right) \omega_n (r+\lambda)[D^{(h)}(r+\lambda) - D^{(h)}(r)] .$$

Let h approach infinity. Thus

$$\int_r^{r+\lambda} [Q(\rho)]^{\frac{n}{n-1}} d\rho \leq \frac{\omega_n}{(n-1)^{n/2}} (r+\lambda)[D(r+\lambda) - D(r)] ,$$

where

$$Q(\rho) = \int_{S_\rho} J dV .$$

We know $|\nabla f|^2 \leq nKJ^{2/n}$ almost everywhere. Hence

$$D(r) \leq (nK)^{n/2} Q(r) .$$

Therefore

$$\int_r^{r+\lambda} [D(\rho)]^{\frac{n}{n-1}} d\rho \leq \frac{\omega_n (nK)^{\frac{n^2}{2(n-1)}}}{(n-1)^{n-2}} [D(r+\lambda) - D(r)] .$$

Let

$$F(r) = \int_r^{r+\lambda} D(\rho) d\rho .$$

Then

$$F(r) \leq \left[\int_r^{r+\lambda} [D(\rho)]^{\frac{n}{n-1}} d\rho \right]^{\frac{n-1}{n}} \left[\int_r^{r+\lambda} d\rho \right]^{1/n} ,$$

which implies

$$[F(r)]^{\frac{n}{n-1}} \leq \lambda^{\frac{1}{n-1}} \int_r^{r+\lambda} [D(\rho)]^{\frac{n}{n-1}} d\rho .$$

So

$$[F(r)]^{\frac{n}{n-1}} \leq C(n, K)^{\frac{1}{n-1}} (r+\lambda) F'(r) ,$$

which implies

$$\frac{dr}{(r+\lambda)} \leq \lambda^{\frac{1}{n-1}} C(n, K) \frac{dF}{[F(r)]^{1+\frac{1}{n-1}}} .$$

Hence

$$\log \frac{R+\lambda}{r+\lambda} \leq \lambda^{\frac{1}{n-1}} C(n, K)\left[\frac{1}{F(r)^{\frac{1}{n-1}}} - \frac{1}{F(R)^{\frac{1}{n-1}}}\right]$$

$$\leq \frac{\lambda^{\frac{1}{n-1}} C(n, K)}{(\lambda D(r))^{\frac{1}{n-1}}} = \frac{C(n, K)}{D(r)^{\frac{1}{n-1}}} .$$

Let λ approach zero and we obtain the desired inequality.

6. Improvement on Hölder exponent.

LEMMA 6. *If*

$$|\nabla f|^2 \leq nKJ^{2/n} ,$$

then

$$|f_{x_1}|^n \leq C^{n/2} |\nabla f|^n ,$$

where

$$C = \frac{K(n-1)^{2/n} - \lambda}{(1-\lambda)^n K(n-1)^{2/n}} ,$$

for any such that $0 < \lambda < 1$.

Proof.

$$J \leq |f_{x_1}| \left(\frac{n-1}{(n-1)^{\frac{n-1}{2}}}\right)\left[\sum_{\substack{i=1,\dots,n \\ j=2,\dots,n}} u_{i,j}^2\right]^{\frac{n-1}{2}} ,$$

since

$$|\det(a_{ij})| \leq \frac{1}{n^{n/2}}\left(\sum_{i,j=1}^{n} a_{ij}^2\right)^{n/2} .$$

Because

$$\frac{n}{(n-1)^{\frac{n}{n-1}}} a^{1/n} b^{\frac{n-1}{n}} \leq \lambda^{n-1} a + \lambda b \quad \text{for} \quad 0 < \lambda < 1 ,$$

$$|\nabla f|^2 \leq nKJ^{2/n} \leq nK |f_{x_1}|^{2/n} (n-1)^{\frac{-1-n}{n}} \left[\sum_{\substack{i=1 \\ j=2}}^{n} u_{i,j}^2\right]^{\frac{n-1}{n}}$$

$$\leq K(n-1)^{2/n}\left[\lambda^{n-1} |f_{x_1}|^2 + \lambda \sum_{\substack{i=1 \\ j=2}}^{n} u_{i,j}^2\right] .$$

Hence

$$|f_{x_1}|^2 \leq \frac{(K(n-1)^{2/n} - \lambda)}{(1-\lambda^n)K(n-1)^{2/n}} |\nabla f|^2 .$$

A simple calculation shows that there is exactly one value of λ between 0 and 1 which will minimize $C(\lambda)$. To find the value of λ, solve the equation

$$(n-1)\lambda^n - nK(n-1)^{2/n}\lambda^{n-1} + 1 = 0 .$$

The Hölder exponent μ of Theorem 1 and Theorem 2 is not the largest exponent that can be obtained. In the proofs of Theorem 1 and Theorem 2 if Lemma 6 were used, the size of μ would be increased.

The constant of Lemma 2 also determines the size of μ. We conjecture that the best constant for this lemma is 1, i.e.,

$$\int_{\dot{S}_r} |u|^n dA \leq r^n \int_{\dot{S}_r} |u_t|^n dA \quad \text{if} \quad \int_{\dot{S}_r} u dA = 0 . \tag{21}$$

This is true if $n = 2$ for then the inequality is Wirtinger's inequality.

If (21) is true, then μ could be defined by the equation

$$\frac{1}{\mu} = \left[\frac{K(n-1)^{2/n} - \lambda}{1-\lambda^n}\right]^{n/2} n^{n/2}(n-1)^{1-1/n-n/2} , \tag{22}$$

where λ is the root between 0 and 1 of the equation

$$(n-1)\lambda^n - nK(n-1)^{2/n}\lambda^{n/1} + 1 = 0 .$$

We further conjecture that this value of μ will be the "best" that can be obtained for given K.

STANFORD UNIVERSITY

NOTE ON ALDER'S POLYNOMIALS

L. CARLITZ

1. Alder's polynomial $G_{M,t}(x)$ may be defined by means of

$$(1)\qquad 1+\sum_{s=1}^{\infty}(-1)^s k^{Ms} x^{\frac{1}{2}s\{(2M+1)s-1\}}(1-kx^{2s})\frac{(kx)_{s-1}}{(x)_s} = \prod_{n=1}^{\infty}(1-kx^n)\sum_{t=0}^{\infty}\frac{k^t G_{M,t}(x)}{(x)_t},$$

where M is a fixed integer ≥ 2 and

$$(a)_t = (1-a)(1-ax)\cdots(1-ax^{t-1}),\ (a)_0 = 1 .$$

Alder [1] obtained the identities

$$(2)\qquad \prod_{n=1}^{\infty}\frac{(1-x^{(2M+1)n-M})(1-x^{(2M+1)n-M-1})(1-x^{(2M+1)n})}{1-x^n} = \sum_{t=1}^{\infty}\frac{G_{M,t}(x)}{(x)_t},$$

$$(3)\qquad \prod_{n=1}^{\infty}\frac{(1-x^{(2M+1)n-1})(1-x^{(2M+1)n-2M})(1-x^{(2M+1)n})}{1-x^n} = \sum_{t=0}^{\infty}\frac{x^t G_{M,t}(x)}{(x)_t}$$

thus generalizing the well-known Rogers-Ramanujan identities. Singh [2, 3] has further generalized (2), (3); he showed that

$$\prod_{n=1}^{\infty}\frac{(1-x^{(2M+1)n-s})(1-x^{(2M+1)n-2M-1+s})(1-x^{(2M+1)n})}{1-x^n} = \sum_{t=0}^{\infty}\frac{A_s(x,\ t)G_{m,t}(x)}{(x)_t},$$

where the $A_s(x,\ t)$ are polynomials in x.

In a recent paper [4] Singh has proved that

$$(4)\qquad G_{M,t}(x) = x^t \qquad (t \leq M-1) .$$

In the present note we give another proof of (4) and indeed obtain the explicit formula

$$(5)\qquad G_{M,t}(x) = \sum_{\substack{Ms\leq t\\ s\geq 0}}(-1)^s\frac{(x)_t}{(x)_s(x)_{t-Ms}}x^{\frac{1}{2}s(s-1)+st}(1-x^s+x^{t-Ms+s})$$

valid for all t.

2. Since

$$(1-kx^{2s})(kx)_{s-1} = (kx)_s + kx^s(1-x^s)(kx)_{s-1} ,$$

the left member of (1) is equal to

Received June 26, 1959.

$$1+\sum_{s=1}^{\infty}(-1)^s k^{Ms}x^{\frac{1}{2}s\{(2M+1)s-1\}}\left\{\frac{(kx)_s}{(x)_s}+kx^s\frac{(kx)_{s-1}}{(x)_{s-1}}\right\}$$

$$=\sum_{s-0}^{\infty}(-1)^s k^{Ms}x^{\frac{1}{2}s\{(2M+1)s=1\}}\frac{(kx)_s}{(x)_s}$$

$$-\sum_{s=0}^{\infty}(-1)^s k^{M(s+1)+1}x^{\frac{1}{2}(s+1)\{(2M+1)(s+1)-1\}+(s+1)}\frac{(kx)_s}{(x)_s}$$

$$=\sum_{s=0}^{\infty}(-1)^s k^{Ms}x^{\frac{1}{2}s\{(2M+1)s-1\}}\frac{(kx)_s}{(x)_s}\{1-k^{M+1}x^{(M+1)(2s+1)}\} .$$

Thus (1) becomes

$$\sum_{t=0}^{\infty}\frac{k^tG_{M,t}(x)}{(x)_t}=\sum_{s=0}^{\infty}(-1)^s k^{Ms}x^{\frac{1}{2}s\{(2M+1)s-1\}}\cdot\frac{1-k^{M1}x^{(M+1)(2s+1)}}{(x)_s}\prod_{j=1}^{\infty}(1-kx^{s+j})^{-1}$$

(6)

$$=\sum_{s=0}^{\infty}(-1)^s k^{Ms}x^{\frac{1}{2}s\{(2M+1)s-1\}}\cdot\frac{1-k^{M+1}x^{(M+1)(2s+1)}}{(x)_s}\sum_{j=0}^{\infty}\frac{k^jx^{sj+j}}{(x)_j} .$$

For $t<M$, it is clear that the coefficient of k^t on the right is simply $x^t/(x)_t$. This proves Singh's result (4).

For $t=M$ we get

$$\frac{G_{M,M}(x)}{(x)_M}=-\frac{x^M}{1-x}+\frac{x^M}{(x)_M} ,$$

so that

$$G_{M,M}(x)=x^M-x^M\frac{(x)_M}{1-x} ,$$

which also was found by Singh.

For $t=M+1$, similarly, we have

$$\frac{G_{M,M+1}(x)}{(x)_{M+1}}=\frac{x^{M+1}}{(x)_{M+1}}-x^{M+1}-\frac{x^{M+2}}{(1-x)^2} ,$$

so that

(7)

$$G_{M,M+1}(x)=x^{M+1}\left\{1-(x)_{M+1}-x\frac{(x)_{M+1}}{(1-x)^2}\right\}$$
$$=x^{M+1}\{1-(1+x^3)(x^3)_{M-1}\} .$$

also due to Singh.

3. For arbitrary $t\geq M+1$, it follows from (6) that

$$G_{M,t}(x)=\sum_{Ms\leq t}(-1)^s\frac{(x)_t}{(x)_s(x)_{t-Ms}}x^{\frac{1}{2}s\{(2M+1)s-1\}(s+1)(t-Ms)}$$
$$-\sum_{M(s+1)\leq t}(-1)^s\frac{(x)_t}{(x)_s(x)_{t-M(s+1)-1}}x^{e_s} ,$$

where

$$e_s = \tfrac{1}{2}s\{(2M+1)s-1\} + (s+1)\{t-M(s+1)-1\}(M+1)(2s+1)\,.$$

This simplifies to

$$(8)\qquad \begin{aligned} G_{M,t}(x) &= x^t \sum_{Ms\le t} (-1)^s \frac{(x)_t}{(x)_s(x)_{t-Ms}} x^{\frac{1}{2}s(s-1)+s(t-M)} \\ &+ \sum_{0<Ms<t} (-1)^s \frac{(x)_t}{(x)_{s-1}(x)_{t-Ms-1}} x^{\frac{1}{2}s(s-1)+st}\,, \end{aligned}$$

or if we prefer

$$(9)\qquad G_{M,t}(x) = \sum_{\substack{Ms\le t\\ s\ge 0}} (-1)^s \frac{(x)_t}{(x)_s(x)_{t-Ms}} x^{\frac{1}{2}s(s-1)+st}(1 - x^s + x^{t-Ms+s})\,.$$

For example (9) reduces to

$$(10)\qquad G_{M,t}(x) = x^t\left\{1 - \frac{(x)_t}{(x)_1(x)_{t-M}}(1 - x + x^{t-M+1})\right\}$$

for $M+1 \le t \le 2M-1$. When $t = M+1$, it is easily verified that (9) reduces to (7). Singh [4] conjectured the truth of (10) for $t \le 2(M-1)$.

REFEFENCES

1. H. L. Alder, *Generalizations of the Rogers-Ramanujan identities*, Pacific J. Math. **4** (1954), 161–168.

2. V. N. Singh, *Certain generalized hypergeometric identities of the Rogers-Ramanjan type*, Pacific J. Math. **7** (1957), 1011–1014.

3. ———, *Certain generalized hypergeometric identities of the Rogers-Ramanujan type* (II), Pacific J. Math. **7** (1957), 1691–1699.

4. ———, *A note on the computation of Alder's polynomials*, Pacific J. Math. **9** (1959), 271–275.

DUKE UNIVERSITY

UNIONS OF CELL PAIRS IN E^3

P. H. DOYLE

In [4] it is shown that there are pairs of cells of all dimensions possible in euclidean 3-space, E^3, which are tame separately, but which have a wild set as their union. Such pairs can be constructed when the individual cells intersect in a single point. The present paper gives conditions that unions of some such pairs be tame sets as well as a number of other results.

LEMMA 1. *Let D_1 be a disk which is polyhedral and which lies on the boundary, ∂T, of a tetrahedron T in E^3. If D_2 is a disk in E^3 which has a polygonal boundary and is locally polyhedral mod ∂D_2 while $D_2 \cap T = D_2 \cap D_1 = \partial D_2 \cap \partial D_1 = J$, an arc, then $D_1 \cup D_2$ is a tame disk.*

Proof. Let P_1 and P_2 be polyhedral disks in ∂T, $P_1 \cap P_2 = \square$ and $(P_1 \cup P_2) \cap D_1 = \square$. Then $\overline{\partial T \backslash (P_1 \cup P_2)}$ is a polyhedral annulus, A_1. If Q is a polyhedral disk in $D_2 \backslash \partial D_2$, then $\overline{D_2 \backslash Q}$ is an annulus A_2 which is locally polyhedral mod ∂D_2. By applying Lemma 5.1 of [8] to A_1 and A_2 one obtains a space homeomorphism h carrying E^3 onto E^3 while $h(D_1 \cup D_2)$ is a polyhedral set. This completes the proof of Lemma 1.

LEMMA 2. *Let D_1 be the disk of Lemma 1 while D_2 is a tame disk in E^3 such that $D_2 \cap T = D_2 \cap D_1 = \partial D_2 \cap \partial D_1 = J$, an arc. Then $\partial T \cup \partial D_2$ is tame.*

Proof. By Theorem 2 of [3] $\partial D_1 \cup \partial D_2$ is locally tame and hence tame by [1] or [8]. Let a be a point of ∂J and J' be an interval of ∂D_1 having a as an end point and $J' \cap \partial D_2 = a$. We choose a polygonal disk M on ∂T with $(J'/\partial J')$ in its interior while $\partial D_1 \cap M = J'$. By a swelling [5] of M toward the component of $E^3 \backslash \partial T$ which meets ∂D_2 we obtain a disk M' which is locally polyhedral mod ∂M and $M' \cap \partial T = \partial M = \partial M'$. The sphere $S = M' \cup (\partial T \backslash M)$ is tame by [8] and S is pierced at a by a tame arc lying on $\partial(D_1 \cup D_2)$. Hence by [7] $\partial D_2 \cup S$ is locally tame at a. We select an arc P in $(S \backslash M') \cup a$ which is locally polyhedral except at the point a. There is an arc A on ∂D_2 which lies in the exterior of S except for its end point a. The arc $A \cup P$ is tame since $S \cup \partial D_2$ is tame. Let the arc P be swollen into a 3-cell C^3 with P in its interior such that C^3 is locally polyhedral mod a, $C^3 \cap S$ is a disk while $C^3 \cap M = a$. Then ∂C^3 is pierced at a by $A \cup P$ and so $A \cup P \cup \partial C^3$ is tame by [7]. Evidently there is an arc P' on ∂C^3 so

Received April 27, 1959. The work on part of this paper was supported by the National Science Foundation Grant G-2793.

that $A \cup P'$ pierces ∂T at a. Again by [7] $\partial D_2 \cup \partial T$ is locally tame at a. A similar argument applies to the other end point of ∂J. Hence $\partial D_2 \cup \partial T$ is tame. This proves Lemma 2.

THEOREM 1. *Let D_1 and D_2 be two tame disks in E^3 such that $D_1 \cap D_2 = \partial D_1 \cap \partial D_2 = J$, an arc. Then $D_1 \cup D_2$ is a tame disk.*

Proof. Since D_1 is tame there is a homeomorphism h_1 of E^3 onto E^3 such that $h_1(D_1)$ is a plane triangle. The disk $h_1(D_1)$ is to be swollen so that a 3-cell e^3 is formed such that

(i) $h_1(D_1) \subset \partial e^3$,

(ii) e^3 is tame,

(iii) and $e^3 \cap h_1(D_2) = h_1(J)$.

That such a cell e^3 exists follows from Lemma 5.1 of [5] and Theorem 9.3 of [8].

There is a homeomorphism h_2 of E^3 onto E^3 which carries ∂e^3 and $h_1(D_1)$ onto the boundary of a tetrahedron and a polyhedral disk, respectively. By Lemma 2 $h_2(e^3) \cup h_2h_1(\partial D_2)$ is a tame set. By Theorem 2 of [6] we can insist that $h_2h_1(D_2)$ be locally polyhedral mod $h_2h_1(\partial D_2)$, while $h_2h_1(\partial D_2)$ is polygonal. Hence by Lemma 1 $h_2h_1(D_1 \cup D_2)$ is tame and so $D_1 \cup D_2$ is tame.

The following result gives a characterization of tame 1-dimensional complexes in E^3. By *a* 1_n-star we mean a homeomorphic image of a 1-dimensional simplicial complex K with a vertex x whose star is K and x is the common end point of the n segments meeting only in x.

THEOREM 2. *If N is a 1_n-star in E^3 such that $(n-1)$ of the branches of N lie on a disk D which meets the remaining branch J at x only and if each arc in N is tame, then N is tame.*

Proof. By [2] we may assume that D is locally polyhedral mod N. An application of the method in Theorem 1 of [3] makes it possible to select a subset D' of D which is a disk consisting of $(n-1)$ tame disks which contain arcs with x as an end point of all branches of N except J. An argument almost identical with that of Theorem 2 of [3] suffices to show that $J \cup D'$ is tame and hence N is tame by [1] or [8].

COROLLARY 1. *Let G be a graph in E^3 such that the star of each vertex of G meets the conditions of Theorem 2, then G is tame. The conditions are evidently necessary as well.*

COROLLARY 2. *Let D be a tame disk and J a tame arc in E^3. If $D \cap J = \partial D \cap J = p$, an end point of J, and if $\partial D \cup J$ is tame, then $D \cup J$ is tame.*

Proof. Since D is tame there is a space homeomorphism h which carries D onto a face of a tetrahedron T, $[h(J)\backslash h(p)] \subset E^3\backslash T$. Let P be a segment on $h(\partial D)$ with $h(p)$ as an end point. We enclose P in a polyhedral disk M in ∂T such that P spans M and $h(\partial D) \cap M = P$. We swell M as in Lemma 2 to obtain a tame disk M' such that $\partial M' = \partial M$, and $M'\backslash \partial M' \subset E^3\backslash T$. Then $h(J) \cup h(\partial D)$ contains a tame arc which pierces the tame sphere [8] $S = M' \cup (\partial T\backslash M)$ at $h(p)$ and so $S \cup h(J)$ is tame by [7]. The construction of an arc P' as in Lemma 2 completes the proof.

In Example 1.4 of [4] an arc A which is the union of two tame arcs is shown. Although A has an open 3-cell complement in compactified E^3, it is nevertheless wild. A similar example can be obtained from Example 1.4 of two tame disks which meet at a point on the boundary of each and which have a wild union. In this connection we give the following result.

THEOREM 3. *Let D_1 and D_2 be disks in E^3 such that each arc in D_1 and D_2 is tame and $D_1 \cap D_2 = \partial D_1 \cap \partial D_2 = J$, an arc. Then $D_1 \cup D_2$ is a disk such that each arc in $D_1 \cup D_2$ is tame.*

Proof. Let J' be an arc in $D_1 \cup D_2$. If $\partial J'$ does not lie in $\partial D_1 \cup \partial D_2$ we extend J' so that this is the case, obtaining $J'' \supset J'$, $\partial J'' \subset \partial D_1 \cup \partial D_2$ and $J'' \subset D_1 \cup D_2$. By [2] there is a disk D such that $\partial D = \partial(D_1 \cup D_2)$, $J \cup J'' \subset D$ and D is locally polyhedral mod $J \cup J'' \cup \partial D$. The arc J in D is the intersection of two disks in D, D_1' and D_2', such that $D_1' \cup D_2' = D$. Consider any point x of J'' in $D_1'\backslash \partial D_1'$. In [3] a method is given for enclosing x in the interior of a tame subdisk of D_1'. Hence D_1' is locally tame at each of its interior points and $\partial D_1'$ is tame. By [8] D_1' is tame. A similar argument can be applied to D_2'. Hence $D_1' \cup D_2'$ is a tame disk by Theorem 2. Then J'' is tame and so J' is tame. Since J' was arbitrarily chosen $D_1 \cup D_2$ is a disk in which each arc is tame.

COROLLARY 1. *Let L_1 and L_2 be tame disks which intersect in a single point on the boundary of each. If $L_1 \cup L_2$ lies on a disk in which each arc is tame, then $L_1 \cup L_2$ is tame.*

Proof. Let $L_1 \cup L_2$ lie on a disk D such that each arc in D is tame. By Theorem 2 $\partial L_1 \cup \partial L_2$ is tame. There is a disk D' in D with a tame boundary such that $D' \cap (L_1 \cup L_2) \subset \partial L_1 \cup \partial L_2$ while $D' \cup L_1 \cup L_2$ is a disk. Then by [2] there is a disk D'' such that $\partial D'' = \partial D'$, D'' is locally polyhedral mod $\partial D''$ and $\partial D'' \cap (L_1 \cup L_2) = \partial D' \cap (L_1 \cup L_2)$. Now D'' is tame by [8] and so $D'' \cup L_1 \cup L_2$ is tame by Theorem 2. It follows that $L_1 \cup L_2$ is tame.

REFERENCES

1. R. H. Bing, *Locally tame sets are tame*, Ann. of Math. **59** (1954), 145–158.

2. R. H. Bing, *Approximating surfaces with polyhedral ones*, Ann. of Math. **65** (1957), 456–483.
3. P. H. Doyle, *Tame triods in E^3*, Proc. Amer. Math. Soc. **10** (1959), 656–658.
4. R. H. Fox and E. Artin, *Some wild cells and spheres in three-dimensional space*, Ann. of Math. **49** (1948), 979–990.
5. O. G. Harrold, H. C. Griffith, and E. E. Posey, *A Characterization of tame curves in three-space*, Trans. Amer. Math. Soc. **79** (1955), 12–34.
6. E. E. Moise, *Affine structures in 3-manifolds, V. The triangulation theorem and Hauptvermutung*, Ann. of Math. **56** (1952), 96–114.
7. ———, *Afflne structures in 3–manifolds, VII. Disk which are pierced by intervals*, Ann. of Math. **58** (1953), 403–408.
8. ———, *Affline structures in 3-manifolds, VIII. Invariance of the knot-types; Local tame imbedding*, Ann. of Math. **59** (1954), 159–170.

MICHIGAN STATE UNIVERSITY
UNIVERSITY OF TENNESSEE

A CLASS OF SMOOTH BUNDLES OVER A MANIFOLD

JAMES EELLS, JR.

1. Introduction. In this paper we illustrate certain constructions of importance in the geometry of smooth manifolds. First of all we prove that *a homogeneous space B of a connected Lie group G can always be represented as a homogeneous space of a contractible Lie group E*, necessarily of infinite dimension in general. In particular, that representation shows that the loop space of B can be replaced effectively by a Lie group of infinite dimension. The construction is a special case of a general theory of differentiable structures in function spaces [4]. Secondly, we examine relations between the Lie algebra of G and that of E (this latter being a Banach-Lie algebra), in case G is compact and semi-simple.

As an application we consider certain differentiable fibre bundles over a smooth (i.e., infinitely differentiable) manifold X having infinite dimensional Lie structure groups. Particular attention is given to the bundles associated with maps of X into a sphere; these bundles are important because they are in natural (Poincaré dual) correspondence with certain equivalence classes of normally framed submanifolds of X. Using a theory of smooth differential forms in function spaces, *we give explicit integral representation formulas for the characteristic classes of these bundles*. These formulas provide examples of a residue theory of differential forms with singularities [1]—and express those forms with singularities as forms without singularities in differentiable bundles over X.

2. The homogeneous spaces. (A) Let G be a connected Lie group (of finite dimension!), and let $L(G)$ denote its Lie algebra, considered as the tangent space to G at its neutral element e. If K is a closed subgroup of G, we let B denote the homogeneous space G/K of left cosets of K. The coset map $\pi : G \to B$ is an analytic fibre bundle map [9, § 7].

We now construct an *acyclic* fibre bundle over B; our construction is a variant of Serre's space of paths over B based at a point [8, Ch. IV]. For this purpose we have chosen a special class of paths on G suitable for our applications in § 5. (These path spaces are also of importance in the calculus of variations.)

(B) Let G be given a left invariant Riemann structure, determined by an inner product on $L(G)$. If $\mathscr{T}(G)$ denotes the tangent vector bundle of G with projection map $q : \mathscr{T}(G) \to G$, then $\mathscr{T}(G)$ has induced Riemann structure. If u, v are tangent vectors at a point

Received June 10, 1959. Research partially supported by the Office of Naval Research.

$m \in G$, we let $(u, v)_m$ denote their inner product, and $|v|_m$ denote the length of v.

DEFINITION. If I is the unit interval $\{t \in I : 0 \leq t \leq 1\}$, we say that a map $x : I \to G$ is an *admissible path on* G if it satisfies the following conditions:

(1) $x(0) = e$, the neutral element of G;

(2) x is absolutely continuous in the metric of G; then its tangent vector $x'(t)$ exists for almost all $t \in I$, and we require that

(3) the tangent map $x' : I \to \mathscr{T}(G)$ is square integrable; i.e., the Lebesgue integral

$$\int_0^1 | x'(t) |_{x(t)}^2 dt \tag{1}$$

is finite. We observe that $x(t) = q \circ x'(t)$ for each $t \in I$ for which $x'(t)$ exists.

Let $E(G)$ denote the totality of admissible paths on G. Using pointwise multiplication and metric defined analogously to (1), it is easily seen that $E(G)$ is a topological (metrizable) group. As in the case of continuous path spaces [8, p. 481], $E(G)$ is a contractible group with contraction $h : E(G) \times I \to E(G)$ given by $h(x, t)s = x(ts)$.

Let $p : E(G) \to G$ be defined by $p(x) = x(1)$. Then p is a continuous epimorphism whose kernel is the subgroup $\Omega(G) = \{x \in E(G) : x(1) = e\}$ of admissible loops on G; thus we have an exact sequence

$$0 \longrightarrow \Omega(G) \longrightarrow E(G) \xrightarrow{\ p\ } G \longrightarrow 0 \tag{2}$$

of topological groups. If $E(G, K) = \{x \in E(G) : x(1) \in K\}$, then $E(G, K)$ is a closed subgroup of $E(G)$, and the composition $\lambda = \pi \circ p : E(G) \to G \to B$ is a representation of B as a homogeneous space of $E(G)$, with $E(G, K)$ as fibre over $b_0 = \pi(K) \in B$.

PROPOSITION. *$\lambda : E(G) \to B$ is a principal $E(G, K)$-bundle.*

To prove that it remains (by [9, p. 30]) to show that there is a local section of $E(G)$ defined in a neighborhood of b_0; because π is a bundle map it suffices to find a neighborhood V of e in G and a section f of $E(G)$ over V. We use the Riemann structure of G to obtain a neighborhood V of e such that for any point $m \in V$ there is a unique geodesic segment $x_m : I \to V$ such that $x_m(0) = e$ and $x_m(1) = m$; then x_m is clearly an admissible path, and $f(m) = x_m$ is a continuous map of V into $E(G)$ such that $p \circ f(m) = m$ for all $m \in V$.

(C) The following result is an application of a general theory of function space manifolds [4].

THEOREM. *Let G be a connected Lie group, and $E(G)$ the space of* *its admissible paths. Then $E(G)$ is an infinite dimensional Lie group modeled on a separable Hilbert space. The map $p: E(G) \to G$ is an analytic bundle epimorphism.*

We recall the principal ideas of that construction. Given $x \in E = E(G)$, *the tangent space to E at x* is the separable Hilbert space $E(x)$ of maps $u: I \to \mathscr{T}(G)$ such that

(1) $q \circ u(t) = x(t)$ for all $t \in I$,

(2) $u(0) = 0$ (the zero in $L(G)$), and

(3) the map u is absolutely continuous with square integrable tangent vector field, and the norm $|u|_x$ induced from the inner product (3) below is finite. Thus $E(x)$ is considered as the space of *admissible variations* of the path x. The algebraic operations in $E(x)$ are defined pointwise; i.e., if $u, v \in E(x)$ and $a, b \in \boldsymbol{R}$, then $(au + bv)t = au(t) + bv(t)$, where the right member is computed in the tangent space $G(x(t))$. A symmetric, bilinear form in $E(x)$ is defined by

$$(u, v)_x = \int_0^1 (u'(t), v'(t))_{x(t)} dt \; ; \tag{3}$$

this is an inner product, for if $(u, u)_x = 0$, then $|u'(t)|_{x(t)} = 0$ for almost all $t \in I$, and the condition that u is admissible then implies $u(t) = 0$ for all $t \in I$. We emphasize that each $E(x)$ is complete (by standard L^2 theory), a property that is used in the theory of differentiation in infinite dimensional linear spaces.

Using the natural correspondence (defined locally) between geodesic segments on G emanating from a point m and tangent vectors in $G(m)$, we can find a neighborhood U_x (called a *coordinate patch*) of x in $E(G)$ which is mapped homeomorphically (by a map ϕ_x called a *coordinate system*) onto a neighborhood of 0 in $E(x)$ [4, § 3]. In overlapping coordinate patches U_x, U_y we have a map

$$\phi_{xy} : \phi_x(U_x \cap U_y) \longrightarrow \phi_y(U_x \cap U_y)$$

defined by $\phi_{xy}(u) = \phi_y \circ \phi_x^{-1}(u)$, and this map is analytic in its domain of definition. (If ϕ is a map of an open subset U of a Hilbert space E into a Hilbert space F, then *ϕ is analytic in U* if every $x \in U$ has a neighborhood in which ϕ can be expressed by the convergent power series

$$\phi(x + v) = \phi(x) + \sum_{k=1}^{\infty} P_\phi^k(x, v)/k! \; ,$$

where $P_\phi^k(x, v)$ denotes the kth iterated directional derivative of ϕ at x in the direction v.) Easy modifications of standard Lie group theory show that the group operation in $E(G)$ is analytic and that $p: E(G) \to G$ is an analytic homomorphism.

COROLLARY. *The fibration* $\lambda : E(G) \to B$ *is an analytic bundle map.*

(D) REMARK. *The inner product* (3) *is easily seen to provide an analytic Riemann structure on* $E(G)$. We note, however, that it is not left invariant on $E(G)$.

Suppose we let G *act* on $E(G)$ by $T_g(x)t = gx(t)g^{-1}$ for all $t \in I$ and $x \in E(G)$. If G is compact and semi-simple and if the inner product (3) is computed using the bi-invariant Riemann metric on G (see our § 3A), then the *Riemann structure on* $E(G)$ *is G-invariant.*

3. The Lie algebra of certain path groups. (A) Suppose that G is connected, compact, and semi-simple. Then its Killing form [7, §§ 6, 11] defines a bi-invariant Riemann structure on G (essentially unique); furthermore, the inner product and the bracket in $L(G)$ are related by

$$([x, y], z) = (x, [y, z]) \tag{1}$$

for all $x, y \in L(G)$. By taking a suitable real multiple of the Killing form we can suppose that the norm induced from the inner product and the bracket in $L(G)$ are related by

$$|[x, y]| \leq |x||y| \tag{2}$$

for all $x, y \in L(G)$.

(B) If e also denotes the neutral element of $E(G)$ (so that $e(t) = e$ for all $t \in I$), then the tangent space $E(e)$ consists of those admissible paths on $L(G)$ starting at 0; we introduce the bracket of u and v in $E(e)$ by

$$[u, v]t = [u(t), v(t)] \qquad \text{for all } t \in I. \tag{3}$$

We will call $E(e)$ *the Lie algebra of* $E(G)$, and henceforth will denote it by $L(E(G))$; note that $L(E(G) = E(L(G))$. Of course the exponential map exp: $L(E(G)) \to E(G)$ is defined by $(\exp u)t = \exp(u(t))$ for all $t \in I$.

If $|u|_e^2 = (u, u)_e$ in the notation of § 2 (3), then the following result shows that the bracket (3) on $L(E(G))$ is continuous.

LEMMA. *For any* $u, v \in L(E(G)$ *we have*

$$|[u, v]|_e \leq 2|u|_e|v|_e . \tag{4}$$

Proof. First of all, we note that if $m_u = \max\{|u(t)| : t \in I\}$, then $m_u \leq |u|_e$. Namely, for any $t \in I$ we apply the Schwarz inequality to obtain

$$|2u(t) - u(1)|^2 = \left|\int_0^1 \operatorname{sgn}(t-s)u'(s)ds\right|^2 \leq \int_0^1 \operatorname{sgn}(t-s)^2 ds \int_0^1 |u'(s)|^2 ds .$$

Thus

$$m_u \leq \max\{|\,2u(t) - u(1)\,| : t \in I\} \leq |\,u\,|_e\ .$$

By (2) and the Schwarz inequality in $L(G)$ we find that $|\,[u, v]\,|_e^2$ is bounded by

$$\begin{aligned}\int_0^1 \{|\,u'(t)\,|^2\,|\,v(t)\,|^2 + 2\,|\,u'(t)\,|\,|\,v(t)\,|\,|\,u(t)\,|\,|\,v'(t)\,| + |\,u(t)\,|^2\,|\,v'(t)\,|^2\}dt \\ \leq m_v^2 \int_0^1 |\,u'(t)\,|^2 dt + 2m_u m_v \int_0^1 |\,u'(t)\,|\,|\,v'(t)\,|\,dt + m_u^2 \int_0^1 |\,v'(t)\,|^2 dt \\ \leq 4\,|\,u\,|_e^2\,|\,v\,|_e^2\ .\end{aligned}$$

The inequality (4) follows.

REMARK. Unlike the finite dimensional Hilbert-Lie algebra $L(G)$, $L(E(G))$ does not satisfy a relation of the form (1). Thus the bracket in $L(E(G))$ respects its Banach space structure—i.e., *$L(E(G)$ is a Banach-Lie algebra*—rather than its structure as a Hilbert space.

(C) Let $p_* : L(E(G)) \to L(G)$ be defined by $p_*(u) = u(1)$; clearly p_* is a Lie algebra epimorphism, and the inequality

$$|\,u(t_2) - u(t_1)\,| \leq |\,t_1 - t_2\,|^{1/2}\,|\,u\,|_e \qquad \text{for any } t_1, t_2 \in I$$

shows that $|\,p_*(u)\,| \leq |\,u\,|_e$ for all $u \in L(E(G))$.

Our next result establishes an infinitesimal analogue of the analytic bundle over G given by Theorem 2C.

THEOREM. *If G is a connected, compact, semi-simple Lie group, then p_* is a continuous Lie epimorphism with kernel $L(\Omega(G)) = \Omega(L(G))$, the closed ideal of admissible loops on $L(G)$; i.e.,*

$$(5) \qquad 0 \longrightarrow L(\Omega(G)) \longrightarrow L(E(G)) \xrightarrow{p_*} L(G) \longrightarrow 0$$

is an exact sequence of Banach-Lie algebras. Furthermore, as Hilbert spaces (but not as Lie algebras), *p_* induces an orthogonal direct decomposition $L(E(G)) \approx L(\Omega(G)) \oplus M$, where M is a vector space isomorphic to $L(G)$.*

Proof. The first statement follows from the algebraic properties of p_* and the fact that p_* is bounded, and therefore continuous. To prove the second, we define a map j: $L(G) \to L(E(G))$ by letting $j(x)$ be the linear path $j(x)t = tx$ for each $x \in L(G)$; then j is a linear map of $L(G)$ onto a subspace M of $L(E(G))$, and $p_* \circ j$ is the identity; moreover, i is an isometry, because for any $x, y \in L(G)$,

$$(j(x), j(y))_e = \int_0^1 (x, y)dt = (x, y)\ .$$

Note, however, that M is not a subalgebra of $L(E(G))$.

The subspaces $L(\Omega(G))$ and M are orthogonal complements in $L(E(G))$, for if $x \in L(G)$ and $v \in L(\Omega(G))$, then

$$(j(x), v)_e = \int_0^1 (x, v'(t))dt = (x, v(1)) - (x, v(0)) = 0 \ .$$

COROLLARY. *The group $\Omega(G)$ of admissible loops on G forms a subgroup of $E(G)$ whose codimension (as a submanifold of $E(G)$) equals the dimension of G.*

REMARK. If K is a closed subgroup of G and if we set $\lambda_* = \pi_* \circ p_* : L(E(G)) \rightarrow L(G) \rightarrow L(G)/L(K)$, then we have an exact sequence of vector spaces

$$0 \longrightarrow L(E(G, K)) \longrightarrow L(E(G)) \xrightarrow{\lambda_*} L(G)/L(K) \longrightarrow 0 \ .$$

(D) PROBLEM. Consider $L(E(G))$ as a Hilbert space, and form its *topological exterior algebra* $C^*(L(E(G)))$, using the natural inner product on its pth exterior power. The inequality (4) implies that we can construct the Lie algebra cochain complex as in [7, § 3] and that the differential operator in $C^*(L(E(G)))$ is continuous. The elements $\omega \in C^p(L(E(G)))$ determine left invariant differential p-forms on $E(G)$—an important property because a version of de Rham's Theorem is valid for $E(G)$ (see § 5A). What are the relations between the derived cohomology algebras $H^*(L(E(G)))$, $H^*(L(\Omega(G)))$, and $H^*(L(G)) \approx H^*(G; \boldsymbol{R})$?

As a first step, because $L(\Omega(G))$ is a closed ideal in $L(E(G))$ we can appeal to our Theorem 3C and Theorem 4 of *Cohomology of Lie algebras*, G. Hochschild and J-P. Serre, Annals of Math. **57** (1953), 591–603, to obtain the

PROPOSITION. *The filtration of $C^*(L(E(G)))$ by the ideal $L(\Omega(G))$ determines a spectral sequence such that*

$$E_2^{p,q} = H^p(L(G); H^q(L(\Omega(G))) \ ,$$

and whose terminal algebra E_∞ is the graded algebra associated with $H^*(L(E(G)))$, suitably filtered.

4. The bundles over a manifold. (A) Let $B = G/K$ be the homogeneous space of § 2A. Since $E(G)$ is contractible, the fibre bundle $\lambda : E(G) \rightarrow B$ can be interpreted as a universal bundle [9, § 19] for the infinite dimensional Lie group $E(G, K)$. In particular, by the Classification Theorem for principal bundles we have the

PROPOSITION. *If X is a paracompact smooth manifold of finite* *dimension, then the isomorphism classes of smooth principal $E(G, K)$-bundles over X are in natural one-to-one correspondence with the smooth homotopy classes of maps of X into B.*

In that statement we have made use of the fact that for maps of X into B their classification by homotopy equivalence coincides with classfication by smooth homotopy equivalence.

REMARK. There is a certain uniqueness theorem for universal bundles over B, which implies that for any other contractible bundle over B with group Γ, the homotopy groups of Γ are isomorphic to those of $E(G, K)$; see [6, p. 284]. Of course, it follows directly from the homotopy sequence of a bundle and the 5-lemma that the homotopy groups of $E(G, K)$ are isomorphic to those of the loop space of B.

(B) Suppose that B is $(n-1)$-connected and that the nth homotopy group $\pi_n(B)$ is infinite cyclic $(n > 1)$; then the group $E(G, K)$ is $(n-2)$-connected, and the connecting homomorphism of the homotopy sequence of the universal bundle of B is an isomorphism of $\pi_n(B)$ onto $\pi_{n-1}(E(G, K))$.

Let $\mu: W \rightarrow X$ be an $E(G, K)$-bundle over X. Its *characteristic class* [9, p. 178] is the primary obstruction to the construction of a section of the bundle. The condition $n > 1$ insures that its structural group is 0-connected, whence the bundle $\mathscr{B}$ of local coefficients (used in defining characteristic classes in general) is simple [9, p. 153]. *To orient the bundle* is to choose one of the two isomorphism of $\mathscr{B}$ onto the product bundle $X \times Z$. Thus the characteristic class of an oriented $E(G, K)$-bundle over X is a cohomology class $w \in H^n(X, \boldsymbol{Z})$.

It is well known that such a characteristic class can be represented by a transgressive pair of cochains (a^n, c^{n-1}). (A *transgressive pair* in a bundle consists of a cochain of some sort c on W whose restriction to a fibre is a cocycle of $E(G, K)$, and such that its coboundary $dc = \mu^* a$ for some cocycle a of X.) Furthermore, the restriction of c^{n-1} to a fibre defines the generator of $H^{n-1}(E(G, K); \boldsymbol{Z}) \approx \boldsymbol{Z}$ which is the negative of that determined by the orientation of the bundle.

Let w_0 be the characteristic class of the universal oriented bundle $\lambda: E(G) \rightarrow B$. Suppose that $\mu: W \rightarrow X$ is induced by the smooth map $f: X \rightarrow B$, and let $g: W \rightarrow E(G)$ be a smooth bundle map covering f [9, § 19]. If (a_0, c_0) is a transgressive pair representing w_0, then $a = f^* a_0$, $c = g^* c_0$ is known to be a transgressive pair representing the characteristic class w of $\mu: W \rightarrow X$ [2, § 18].

5. Representations of the characteristic classes. (A) Let Y be any paracompact smooth manifold modeled on a Hilbert space E. A *differential r-form η on Y* assigns to each point $y \in Y$ an alternating r-linear functional (with real values) on the tangent space $Y(y)$, which is continuous simultaneously in the r variables, using the Hilbert space topology in $Y(y)$. In terms of the differentiable structure on Y we can define the exterior algebra $\mathscr{E}^*(Y)$ of smooth differential forms on Y and its derived cohomoly algebra $H^*(\mathscr{E}^*(Y))$. It is known (an extension of de Rham's Theorem [4, § 4]) *that there is a canonical isomorphism of* $H^*(\mathscr{E}^*(Y))$ *onto* $H^*(Y; \boldsymbol{R})$, *the singular real cohomology algebra of* Y.

We remark that this result uses the local Hilbert space structure of Y in two ways:

(1) the square of the norm in E is an analytic function on E, which implies that there are sufficiently many smooth functions on Y;

(2) there is a natural Hilbert space structure on the rth exterior power of E; its completeness is used essentially in the differentiability of differential forms.

We will now give examples of such forms which are transgressive pairs on $E(G, K)$-bundles over X.

(B) We have seen in Theorem 2C that the group $E(G)$ of admissible paths on a connected Lie group G is itself a Lie group modeled on a Hilbert space. Since $E(G)$ is contractible, the general existence theorem quoted in (A) insures that any smooth closed r-form ω on $E(G)$ is the exterior differential of a smooth $(r-1)$-form ξ (for $r > 0$). The following result uses a standard homotopy construction to give an explicit formula for ξ in case ω is the p^*-image of a form on G.

PROPOSITION. *Given any smooth closed* r*-form* ω *on* G $(r > 0)$, *consider the* $(r-1)$*-form on* $E(G)$ *defined as follows: For any* $x \in E(G)$ *and* $r-1$ *vectors* $u_1, \cdots, u_{r-1}$ *in the tangent space at* x, *set*

$$(1) \qquad \xi(x) \cdot u_1 \vee \cdots \vee u_{r-1} = \int_0^1 \{\omega(x(t)) \cdot x'(t) \vee u_1(t) \vee \cdots \vee u_{r-1}(t)\} dt ,$$

where $x'(t)$ *denotes the tangent vector to* x *at* $x(t)$, *and the bracket in the right member* (*involving the exterior product* $\vee$) *is computed in the tangent space* $G(x(t))$. *Then* ξ *is a smooth* $(r-1)$*-form on* $E(G)$ *and* $d\xi = p^*\omega$.

Proof. The contraction $h : I \times E(G) \to E(G)$ given by $h(t, x)s = x(ts)$ is simultaneously continuous in the arguments (t, x), and is a smooth function of x for each $t \in I$. Furthermore, for each $x \in E(G)$ the differential $h_*(t, x)$ is a square integrable function of t; in particular, if e_1 denotes the unit vector of I, then $(h_*(t, x) \cdot e_1)s = sx'(ts)$ for almost all $x \in I$.

Because the homomorphism p is analytic, the induced form $\omega^* = p^*\omega$ is a smooth closed r-form on $E(G)$ for which

$$(2) \qquad \xi(x) = (k\omega^*)x = \int_0^1 h^*\omega^*(t, x) \wedge e_1 dt$$

exists (as a Lebesgue integral, where the integrand in the right member involves the interior product with e_1). The explicit formula (3) for $\xi(x)$ below shows that $\xi(x)$ is actually an $(r-1)$-covector and that ξ is smooth. Standard reasoning about homotopy operators for differential forms leads to the identity $\omega^* = dk\omega^* + kd\omega^*$, and because $d\omega = 0$, we have $d\xi = \omega^*$.

Consider the composite map $q = p \circ h : I \times E(G) \to B$. It is easily checked that $q_*(t, x)e_1 = x'(t)$ for almost all $t \in I$, and for any u in the tangent space at x (interpreted as the vector $0 \oplus u$ in the tangent space of $I \times E(G)$ at (t, x)) we have $q_*(t, x)u = u(t)$. If we take vectors $u_1, \cdots, u_{r-1}$ as in the hypotheses,

$$
\begin{aligned}
\xi(x) \cdot u_1 \vee \cdots \vee u_{r-1} &= \int_0^1 h^* \circ p^*\omega(t, x) \cdot e_1 \vee u_1 \vee \cdots \vee u_{r-1} dt \\
&= \int_0^1 \{\omega(x(t)) \cdot q_*(t, x)e_1 \vee q_*(t, x)u_1 \vee \cdots \vee q_*(t, x)u_{r-1}\} dt \qquad (3) \\
&= \int_0^1 \{\omega(x(t)) \cdot x'(t) \vee u_1(t) \vee \cdots \vee u_{r-1}(t)\} dt .
\end{aligned}
$$

COROLLARY. *Let $\lambda : E(G) \to B$ be the universal $E(G, K)$-bundle of* § 2B. *Then for any smooth closed r-form ω_0 on B, the formula* (1) *with ω replaced by $\pi^*\omega_0$ defines a smooth $(r-1)$-form ξ_0 on $E(G)$ such that $d\xi_0 = \lambda^*\omega_0$.*

If $i : E(G, K) \to E(G)$ is the inclusion homomorphism, then we remark that $\eta_0 = i^*\xi_0$ *is the suspension of* ω_0 in the sense of [8, p. 453]. Applying [8, Cor. 2, p. 469], we obtain the

COROLLARY. *If B is $(n-1)$-connected and $\pi_n(B)$ is infinite cyclic $(n > 1)$ and if ω_0 is a closed n-form representing a generator v of $H^n(B; \boldsymbol{Z})$, then (ω_0, ξ_0) is a transgressive pair representing v.*

REMARK. Suppose that $\mathscr{G}$ is connected, compact, and semi-simple. Then the bi-invariant Riemann structure on G induces an analytic G-invariant Riemann structure on B. In the preceding corollary a generator v is then represented by a unique harmonic n-form ω_0; furthermore, ω_0 is G-invariant, and $\pi^*\omega_0$ can be expressed as an exterior polynomial in (left invariant) Maurer-Cartan forms on G. Thus *the generator v is uniquely represented by a transgressive pair (ω_0, ξ_0) where ω_0 is harmonic and where ξ_0 is defined by* (1); see § 6A.

(C) We return to the oriented universal bundle $\lambda : E(G) \to B$, where B is $(n-1)$-connected and $\pi_n(B)$ is infinite cyclic $(n > 1)$. (These assumptions can be relaxed at the expense of simplicity of exposition.)

Let X be a smooth manifold of finite dimension, and let $\mu : W \to X$ be a smooth oriented $E(G, K)$-bundle over X with characteristic class w.

Suppose that bundle is induced by a smooth map f of X into B, and let g be a smooth bundle map covering f:

$$\begin{array}{ccc} W & \xrightarrow{g} & E(G) \\ \downarrow \mu & & \downarrow \lambda \\ X & \xrightarrow{f} & B \end{array}$$

If (ω_0, ξ_0) is a transgressive pair of forms representing the characteristic class w_0 of $\lambda : E(G) \to B$ as in (B), then $\omega = f^*\omega_0$, $\xi = g^*\xi_0$ is a transgressive pair representing w (§ 4B).

DEFINITION. An *admissible partial section* of the bundle $\mu : W \to X$ is a smooth section ϕ defined over $X - e(\phi)$, where $e(\phi)$ is a smooth polyhedral subset of X with $\dim e(\phi) \leq \dim X - n$. Admissible partial sections exist because $E(G, K)$ is $(n-2)$-connected. (For example, we can take a smooth locally finite simplicial subdivision L of X and let L_* be a dual subdivision; then standard obstruction theory provides a smooth section over a neighborhood of the $(n-1)$-skeleton $L^{(n-1)}$ of L which can be smoothly extended over $X - L_*^{(m-n)}$, where $m = \dim X$.)

The following result is an example of the general representation theorem of [1, § 4]; note that the present pair $(\omega, \phi^*\xi)$ satisfies the conditions of Corollary 5B of [1]. We will use freely the concepts and results of that paper. As usual in constructing integral formulas for characteristic classes, our method of proof follows that of the Gauss-Bonnet Theorem as given by Chern [3, § 2]: We first obtain a transgressive pair of forms representing the class; we then appeal to Stokes' Formula to localize and interpret the residue (i.e., the right member of (4) below.

THEOREM. *In the above notation, the characteristic class w of the oriented bundle $\mu : W \to X$ is represented by*

$$w \cdot c = \int_c \omega - \int_{\partial c} \phi^*\xi \tag{4}$$

for any admissible partial section ϕ, where c is any smooth integral n-chain on X whose boundary does not intersect $e(\phi)$.

Proof. First of all, $(\omega, \phi^*\xi)$ is an $(\boldsymbol{R}, n)$-pair on X because ϕ is admissible, and in $X - e(\phi)$ we have $d(\phi^*\xi) = \phi^* d\xi = (\mu \circ \phi)^*\omega = \omega$. Secondary, to verify (4) it suffices to do so for the n-simplexes of a simplical subdivision L of X (by Corollary 5A of [1]), provided that $e(\phi)$ lies on the $(m-n)$-skeleton of the dual L_*. Furthermore, in considering its obstruction cocycle we will suppose that ϕ is defined only on $L^{(n-1)}$, and then make below a (piecewise smooth) extension to $L^{(n)} - e$, where e is a discrete set of points; such an alteration will not change the obstruction class.

Let b_σ be the barycenter of the oriented n-simplex σ, and let σ_t be that simplex radially contracted toward b_σ by the ratio $1:(1-t)$, using an admissible coordinate system on X containing σ. Let h be a smooth covering homotopy of that contraction. For any $t<1$ and x in $\partial\sigma_t$ let $r(x)$ be the radial projection x on $\partial 6$; setting $\phi(x) = h(t,\phi(r(x)))$ defines an extension of ϕ over $\sigma - b_\sigma$.

Applying Stokes' Formula to the chain $\tau_t = \sigma - \sigma_t$ we obtain

$$(5) \qquad -\int_{\partial\sigma_t}\phi^*\xi = \int_{\tau_t}\omega - \int_{\partial\sigma}\phi^*\xi .$$

As $t\to 1$ the right member approaches the right member of (4) with $c=\sigma$, because ω is defined on all σ. To complete the proof of the theorem we will show that as $t \to 1$ the left member determines the obstruction cocycle.

Since $-\xi$ defines the generator of $\mu^{-1}(b_\sigma)$ by § 4B, we see that (writing w for the obstruction cocycle)

$$w\cdot\sigma = -\int_{\partial\sigma}\phi^*\xi .$$

On the other hand, the homotopy h satisfies a Lipschitz condition locally on $\mu^{-1}(\sigma)$ (relative to any metric on W), whence there is a number M independent of t such that $t<1$ implies

$$\left|\int_{\phi(\partial\sigma)}\xi - \int_{\phi(\partial\sigma_t)}\xi\right| \le M|1-t| .$$

Using the transformation of integral formula, we find that

$$\left|w\cdot\sigma + \int_{\partial\sigma_t}\phi^*\xi\right| = \left|\int_{\partial\sigma}\phi^*\xi - \int_{\partial\sigma_t}\phi^*\xi\right| \le M|1-t| .$$

This shows that as $t\to 1$ the left member of (5) approaches $w\cdot\sigma$, and formula (4) follows.

6. Spherical maps of a manifold. (A) As an example of the proceding constructions let $G = SO(n+1)$, the rotation group in its usual matrix representation in numerical space $\boldsymbol{R}_{n+1}$. Let $K=SO(n)$, considered as the subgroup of G which acts trivially on the $(n+1)$th axis of $\boldsymbol{R}_{n+1}$. The unit sphere S_n in $\boldsymbol{R}_{n+1}$ is then naturally identified with the homogeneous space G/K, and the coset map $\pi: SO(n+1)\to S_n$ represents $SO(n+1)$ as the principal $SO(n)$-bundle of orthonormal n-frames on S_n]**9**, § 7]. We will suppose that S_n has its usual Riemann structure and is oriented by the coordinate axes in $\boldsymbol{R}_{n+1}$. *Henceforth we denote the infinite dimensional Lie group* $E(SO(n+1), SO(n))$ *by* Λ_n.

Let ω_{ij} $(1 \leq i < j \leq n+1)$ be a base of Maurer-Cartan forms for the conjugate space of $L(SO(n+1))$; if we let $k(n)$ denote the reciprocal of the volume of S_n, then the exterior polynomial (the *Kronecker Index form*) on $SO(n+1)$ given by

$$(1) \qquad \omega_0^* = k(n)\omega_{1,n+1} \vee \cdots \vee \omega_{n,n+1}$$

is known to be S_n-basic (i.e., there is a unique $SO(n+1)$-invariant n-form ω_0 on S_n such that $\pi^*\omega_0 = \omega_0^*$), and thereby represents the harmonic generator of $H^n(S_n; \boldsymbol{Z})$.

Suppose n is even; then a crucial step in the derivation of the Gauss-Bonnet Theorem [3] for S_n establishes that ω_0 is part of a transgressive pair in the principal frame bundle of S_n. If n is odd, then ω_0 does not generally have that property. However, *for all* $n > 1$ *Proposition* 5B *gives an explicit transgressive pair in the oriented universal bundle of* S_n, *determined entirely by the Kronecker Index form.*

(B) If X is a compact, oriented, smooth Riemann manifold of dimension $n+m$, then the isomorphism classes of smooth principal Λ_n-bundles over X play an important role in its geometry, primarily because of the following construction: Let V be a closed, oriented, m-dimensional regularly imbedded submanifold of X; suppose that V admits a smooth normal n-frame in X, and let ϕ be such a frame field; we will call the pair (V, ϕ) *a normally framed submanifold of* X. These have been studied by Kervaire [5, § 1] and Thom [10, Ch. II, 4]. It is known that certain equivalence classes of normally framed m-submanifolds of X are in natural one-to-one correspondence with the homotopy classes of maps of X into S_n [5, § 1]. Combining with the Classification Theorem for Λ_n-bundles, we have the

PROPOSITION. *If X is a compact, oriented, smooth Riemann $(n+m)$-manifold, then there is a natural one-to-one correspondence between equivalence classes of normally framed m-submanifolds of X and isomorphism classes of smooth Λ_n-bundles over X.*

Let (V, ϕ) be a normally framed m-submanifold, and let $i : V \to X$ be the inclusion map; then since V is closed and oriented (the orientation on X and the frame field ϕ determine an orientation of V) we have a distinguished generator $v_0 \in H_m(V, \boldsymbol{Z})$, which determines a definite homology class $i_*(v_0) = v \in H_m(X, \boldsymbol{Z})$; Furthermore, v depends only on the equivalence class of (V, ϕ). On the other hand, applying a theorem of Thom [10, Théorème II.2], we obtain the

PROPOSITION. *In the correspondence of the above proposition, the homology class of a normally framed submanifold is the Poincaré dual of the characteristic class of the oriented Λ_n-bundle associated with it.*

(C) Let X be a smooth manifold of finite dimension. In the study of differential forms with singularities [1] it is important (e.g., in working with exterior products of such forms) to know when a closed $(\boldsymbol{Z}, r)$-pair is cohomologous to a pair defined in terms of a transgressive pair (as in Theorem 5C). For example, it is well known that the isomorphism classes of $SO(2)$-bundles over X are (by their characteristic classes) in natural one-to-one correspondence with the elements of $H^2(X; \boldsymbol{Z})$. An easy construction shows that *every 2-dimensional integral cohomology class of X can be represented by a transgressive pair in a canonically defined $SO(2)$-bundle over X.*

A cohomology class $u \in H^n(X; \boldsymbol{Z})$ is said to be *spherical* if there is a map $f : X \to S_n$ such that $u = f^*(s)$ for some $s \in H^n(S_n; \boldsymbol{Z})$. The representation theorem [1, § 4] of cohomology classes by forms with singularities together with our *Theorem* 5C *gives a transgressive integral representation formula for every spherical class of X in a Λ_n-bundle.* That bundle is uniquely defined by the homotopy class of $f : X \to S_n$, but is not generally determined by u.

Example. Suppose that X has dimension n. The Hopf Classification Theorem then implies that the isomorphism classes of smooth Λ_n-bundles over X are in natural one-to-one correspondence with the elements of $H^n(X; \boldsymbol{Z})$, the correspondence assigning to each isomorphism class its characteristic class. *Theorem* 5C *gives a transgressive integral representation formula for each element v of $H^n(X; \boldsymbol{Z})$ in a bundle canonically associated with v.* Of course that fact is significant only for compact manifolds, because $H^n(X; \boldsymbol{Z}) = 0$ if X is open. On the other hand, it is particularly useful for non-orientable compact manifolds, because then $H^n(X; \boldsymbol{Z})$ has torsion, in which case the singularity of a $(\boldsymbol{Z}, n)$-pair representing v plays an essential role.

If X is orientable and if its Euler characteristic $\chi(X) \neq 0$, then the Gauss-Bonnet Theorem provides a transgressive integral formula for the elements of $H^n(X; \boldsymbol{Z})$ in a finite dimensional bundle over X. In general (and for lower dimensional spherical classes) it appears necessary to use infinite dimensional smooth bundles to obtain such a formula.

Bibliography

1. C. B. Allendoerfer and J. Eells, *On the cohomology of a manifold*, Comm. Math. Helv. **32** (1958), 165–179.
2. A. Borel, Selected topics in the homology theory of fibre bundles. Mimeographed Notes. Univ. of Chicago (1954).
3. S. S. Chern, *On the curvatura integra in a Riemannian manifold*, Annals of Math. **46** (1945), 674–684.
4. J. Eells, *On the geometry of function spaces*, Sym. Inter. de Top. Alg. Mexico (1958), 303–308.

5. M. Kervaire, *Courbure intégrale généralisée et homotopie*, Math. Ann. **131** (1956), 219-252.

6. J. Milnor, *Construction of universal bundles I*, Ann. of Math. **63** (1956), 272–284.

7. Séminaire S. Lie; E.N.S. 1954–5.

8. J-P. Serre, *Homologie singulière des espaces fibrés*, Ann. of Math. **54** (1951), 425–505.

9. N. E. Steenrod, The Topology of Fibre Bundles, Princeton, 1951.

10. R. Thom, *Quelques propriétés globales des variétés différentiables*, Comm. Math. Helv **28** (1954), 17–86.

COLUMBIA UNIVERSITY

COMPUTATIONS OF THE MULTIPLICITY FUNCTION

S. R. FOGUEL

1. Introduction. Let H be a separable Hilbert space. The following two problems will be studied:

1. Given a bounded normal operator A, of multiplicity m, what are the conditions, on the bounded measurable function f, so that the multiplicity of $S = f(A)$ is $n, n < \infty$?

2. How to compute the multiplicity of a normal operator that commutes with a given normal operator, of finite multiplicity?

NOTATION. Let S be a normal operator of multiplicity n, $n < \infty$. There exist a Borel measure μ and n Borel sets in the complex plane $e_1 \supset e_2 \supset \cdots \supset e_n$, such that, up to unitary equivalence,

$$H = \sum_{i=1}^{n} L_2(\mu, e_i) \tag{1.1}$$

$$S\begin{pmatrix} f_1(\lambda) \\ \vdots \\ f_n(\lambda) \end{pmatrix} = \begin{pmatrix} \lambda f_1(\lambda) \\ \vdots \\ \lambda f_n(\lambda) \end{pmatrix}$$

This is the Multiplicity Theorem. (See Theorem X. 5.10) of [1]. The operator S has uniform multiplicity if $e_1 = e_2 = \cdots = e_n$.

The resolution of the identity, of a normal operator A, will be denoted by $E(A; \alpha)$. The Boolean algebra of projections, generated by $E(A; \alpha)$ will be denoted by $\mathfrak{E}_A$. Let $E(\alpha)$ stand for $E(S; \alpha)$ and $\mathfrak{E}$ for $\mathfrak{E}_S$. Throughout this note all operators are assumed to be bounded.

We shall use the following results from [2]:

Let S be a normal operator of multiplicity n, and B a normal operator that commutes with S. Let H and S be represented by 1.1.

THEOREM A. *There exist k Borel measurable bounded complex functions $y_1(\lambda), \cdots, y_k(\lambda)$ and k matrices of Borel measurable bounded complex functions $\varepsilon_1(\lambda), \cdots, \varepsilon_k(\lambda)$ such that:*

For a fixed λ the matrices $\varepsilon_i(\lambda)$ are disjoint self adjoint projections whose sum is the identity and

$$B\begin{pmatrix} f_1(\lambda) \\ \vdots \\ f_n(\lambda) \end{pmatrix} = \left(\sum_{i=1}^{k} y_i \varepsilon^i(\lambda)\right)\begin{pmatrix} f_1(\lambda) \\ \vdots \\ f_n(\lambda) \end{pmatrix}. \tag{1.2}$$

Received October 21, 1958, and in revised form April 24, 1959. This work has been partially supported by the National Science Foundation.

Equivalently, if the self adjoint projections E_i, *are defined by*

$$E_i\begin{pmatrix} f_1(\lambda) \\ \vdots \\ f_n(\lambda) \end{pmatrix} = \varepsilon_i(\lambda)\begin{pmatrix} f_1(\lambda) \\ \vdots \\ f_n(\lambda) \end{pmatrix}$$

then

$$(1.3) \qquad \begin{cases} B = \sum_{i=1}^{k} y_i(S)E_i \\ E(B;\alpha) = \sum_{i=1}^{k} E(y_i^{-1}(\alpha))E_i \ . \end{cases}$$

REMARK. In the above decomposition the numbers $y_i(\lambda)$ for a fixed λ are different eigenvalues of a certain matrix. Thus for each λ there is an integer $k' \leq k$ such that

$$y_i(\lambda) \neq y_j(\lambda) \qquad i \neq j \qquad i, j \leq k', \qquad \varepsilon_i(\lambda) \neq 0 \qquad i \leq k' \ ,$$

and

$$y_{k'+1}(\lambda) = \cdots = y_k(\lambda) = 0 \ ,$$
$$\varepsilon_{k+1}(\lambda) = \cdots = \varepsilon_{k+1}(\lambda) = 0 \ .$$

This is essential for the proof of Lemma 2.1. Also the matrices $\varepsilon_i(\lambda)$ are $n \times n$ matrices.

THEOREM B. *The number* n *is the largest integer such that there exists a nilpotent operator, commuting with* S, *of order* n. *See* [2] *Theorem* 3.1 *and its corollary.*

2. The multiplicity of a function of an operator. The main result in this section is:

THEOREM 2.1. *Let* A *be a normal operator of multiplicity* m, $m < \infty$, *and* f *a bounded measurable function. The operator* $S = f(A)$ *has finite multiplicity, if and only if, there exist* k *disjoint Borel sets* $\beta_1, \cdots, \beta_k$ *and* k *bounded measurable functions* $z_1(\lambda), \cdots, z_k(\lambda)$ *such that:*

a. $\sigma(A) = \bigcup_{i=1}^{k} \beta_i$.

b. *if* $\lambda \in \beta_i$ *then* $z_i(f(\lambda)) = \lambda$ *almost*

everywhere, with respect to $E(A;\alpha)$.

Proof of sufficiency of conditions a and b. Let S_i and A_i be the restrictions of S and A to $E(A;\beta_i)H$. Then

$$S_i = \int_{\beta i} f(\lambda)E(A; d\lambda)$$

hence

$$z_i(S_i) = A_i .$$

Now, it follows from Theorem B that

$$muA_i \geq muS_i \qquad (muT = \text{multiplicity of } T)$$

But the multiplicity function is subadditive:

$$muS \leq \sum_{i=1}^{k} muS_i .$$

To see this we have to observe that muS is the smallest number n such that there exists a set of n elements, $\{x_1, \cdots x_n\}$, $x_i \in H$ and span $\{E(\alpha)x_i, \alpha \text{ a Borel set}\} = H$. ($n$ generating elements.)

Thus

$$muA \leq \sum_{i=1}^{k} muS_i \leq \sum_{i=1}^{k} muA_i \leq mk < \infty .$$

In order to prove necessity we need the following:

LEMMA 2.1. *Let* $S = f(A)$ *have finite multiplicity* n *and let*

$$A = \sum_{i=1}^{k} z_i(S)E_i$$

be the representation 1.3 *then* $E_i \in \mathfrak{E}_A$.

Proof. For every Borel set α $E(\alpha) \in \mathfrak{E}_A$ because $S = f(A)$. Let $E(\alpha)$ be maximal with respect to the property that $E(\alpha)E_1 \in \mathfrak{E}_A$. Such a maximal projection exists by Zorn's Lemma. Now if $E(\sigma(S) - \alpha) \neq 0$ there exists, by the proof of 3.2 in [2] a set β such that:

$$\beta \subseteq \sigma(S) - \alpha \qquad E(\beta) \neq 0$$

and for some Borel set γ

$$E(\beta)E_1 = E(\beta)E(A; \gamma) \in \mathfrak{E}_A .$$

This contradicts the maximality of α, hence $E(\alpha) = I$.

Proof of necessity of conditions a and b. Let S hsve finite multiplicity n. By Lemma 2.1 there exist n sets β_i such that $E(A; \beta_i) = E_i$. Thus

$$E(A; \beta_i)E(A; \beta_j) = 0 \text{ if } i \neq j$$

and

$$\sum_{i=1}^{k} E(A; \beta_i) = I .$$

Therefore the sets β_i can be chosen to be disjoint and satisfy condition a. Also

$$A = \sum_{i=1}^{k} z_i(S)E_i = \sum_{i=1}^{k} z_i(f(A))E(A; \beta_i) = \sum_{i=1}^{k} \int_{\beta} z_i(f(\lambda)E(A; d\lambda) .$$

Hence, if $\beta \subset \beta_i$ then

$$E(A; \beta)A = \int_{\beta} \lambda E(A; d\lambda) = \int_{\beta} z_i(f(\lambda))E(A; d\lambda)$$

or: on the set $\beta_i \lambda = z_i(f(\lambda))$ almost everywhere with respect to the measure $E(A; \alpha)$.

DEFINITION. The function f will be said to have k repetitions, with respect to the measure $E(A; \alpha)$, if conditions a and b of Theorem 2.1 are satisfied.

In the rest of this section we compute muS. It is enough to consider the case where the operator A has uniform multiplicity m: otherwise A can be written as direct sum of operators of uniform multiplicity and one has to study each component of A separately.

The following Theorem is needed:

THEOREM 2.2 *Let H be the direct sum of the orthogonal subspaces $H_1, \cdots, H_k$. Let S_i be a normal operator, on H_i, of uniform multiplicity m_i and S be the direct sum of S_i.*

If

$$E(S; \alpha) = 0 \text{ whenever } E(S_i; \alpha) = 0 \text{ for some } i$$

then

$$muS = \sum_{i=1}^{k} m_i .$$

Proof. It is enough to prove that $muS \geq \sum_{i=1}^{k} m_i$. Let $\sigma = \sigma(S_1) = \cdots = \sigma(S_k) = \sigma(S)$. By the Spectral Multiplicity Theorem each operator S_i can be described as follows: There exists a measure μ_i on σ and H_i is the direct sum of m_i spaces $L_2(\mu_i)$. The operator S_i is given by

$$S_i \begin{pmatrix} f_1(\lambda) \\ \vdots \\ f_{m_i}(\lambda) \end{pmatrix} = \begin{pmatrix} \lambda f_1(\lambda) \\ \vdots \\ \lambda f_{m_i}(\lambda) \end{pmatrix} .$$

Now, the measures μ_i are equivalent, by the condition of the Theorem. Thus there exist functions φ_i, $\varphi_i \in L(\mu_{i+1})$ $1 \leq i \leq k-1$ such that

$$\mu_i(e) = \int_e \varphi_i(\lambda) d\mu_{i+1}$$

for every Borel set e. (Radon Nikodym Theorem, see [3], p. 128). Let us define an operator on H:

If $x \in H_i$,

$$x = \begin{pmatrix} f_1(\lambda) \\ \vdots \\ f_{m_i-1}(\lambda) \\ 0 \end{pmatrix}$$

then

$$Mx \in H_i, \quad Mx = \begin{pmatrix} 0 \\ f_1(\lambda) \\ \vdots \\ f_{m_i-1}(\lambda) \end{pmatrix}.$$

If

$$x \in H_i, \quad x = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ f_{m_i}(\lambda) \end{pmatrix}$$

then

$$Mx \in H_{i+1}, \quad Mx = \begin{pmatrix} \sqrt{\varphi_i(\lambda)}\, f_{m_i}(\lambda) \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Where H_{k+1} is the zero space.

It is easy to see that M is a bounded operator and

$$M^{\sum_{i=1}^{k} m_i} = 0$$

but

$$M^{\sum_{i=1}^{k} m_i - 1} \neq 0 .$$

Also $MS = SM$, hence $muS \geq \sum_{i=1}^{k} m_i$.

REMARK. It was proved in Theorem 2.1 that if a function f has k repetitions then

$$muf(A) \leq kmuA \ .$$

However the number of repetitions of a function is not uniquely defined. In order to compute $muf(A)$ we have to find the minimal number of repetitions. This is what the next Theorem does.

THEOREM 2.3. *Let A be a normal operator of uniform multiplicity m. Let f be a bounded measurable function which has k repetitions with respect to the measure $E(A;\alpha)$. A necessary and sufficient condition that $muS = mk$, where $S = f(A)$, is:*

There exists a Borel set α_0

$$E(A; f^{-1}(\alpha_0)) \neq 0 \tag{2.1}$$

and

$E(A; f^{-1}(\alpha)) = 0$ whenever $E(A; f^{-1}(\alpha) \cap \beta_i) = 0$ for some i and $\alpha \subset \alpha_0$.

Proof. Assume condition 2.1. We may restrict A and S to $E(A; f^{-1}(\alpha_0))H$. Let

$$H_i = E(A; f^{-1}(\alpha_0) \cap \beta_i)H \ ,$$

and A_i, S_i the restriction of A, S to H_i. Now

$$f(A_i) = S_i \qquad\qquad z_i(S_i) = A_i$$

(See Theorem 2.1.). Thus the operators S_i have uniform multiplicity m because the operators A_i do. It follows from Theorem 2.2 that the multiplicity of S restricted to $E(A; f^{-1}(\alpha_0))H$ is mk. But $muS \leq mk$, hence $muS = mk$.

(Note that on α_0 the operator S has uniform multiplicity mk). Conversely, let us assume that for each Borel set α_0 with $E(A; f^{-1}(\alpha_0)) \neq 0$, there exists a subset α such that $E(A; f^{-1}(\alpha)) \neq 0$ but $E(A; f^{-1}(\alpha) \cap \beta_i) = 0$ for some i. Let $E(A; f^{-1}(\alpha_1))$ be maximal with respect to the property

$$E(A; f^{-1}(\alpha_1))E(A; \beta_1) = 0$$

Let $E(A; f^{-1}(\alpha_2))$ be maximal, with respect to the property

$$\alpha_2 \cap \alpha_1 = \varphi \text{ and } E(A; f^{-1}(\alpha_2))E(A; \beta_2) = 0$$

and choose inductively $\alpha_3 \cdots \alpha_n$, $\alpha_i \cap \alpha_j = \varphi$

$$E(A; f^{-1}(\alpha_j))E(A: \beta_j) = 0$$

There exist such maximal projections by Zorn's Lemma. Now if $E(A; \bigcup_{i=1}^{k} f^{-1}(\alpha_i)) \neq I$ there will be a set α and an integer j such that

$$\alpha \cap \left(\bigcup_{i=1}^{k} \alpha_i\right) = 0; \qquad E(A; f^{-1}(\alpha) \cap \beta_j) = 0$$

Thus α_j will not be maximal. Let

$$\bar{\beta}_j = \beta_j \cup (f^{-1}(\alpha_j) \cap \beta_1), \quad j \geq 2 .$$

Then $\bigcup_{j=2}^{u} \bar{\beta}_j = \sigma(A)$ and on $\bar{\beta}_j$ the function f possesses a bounded measurable inverse. Thus f has $k-1$ repetitions and $muS \leq m(k-1)$.

3. The multiplicity of a matrix of functions. Let S be a normal operator of uniform multiplicity n. Let B be a normal operator and $BS = SB$. The operator B is represented as the matrix of functions $\sum_{i=1}^{k} y_i(\lambda)\varepsilon_i(\lambda)$ and also $B = \sum_{i=1}^{k} y_i(S)E_i$ (Equation 1.2 and 1.3). Let us denote by B_i and S_i the restrictions of B and S, respectively, to $E_iH = H_i$.

THEOREM 3.1. *The operator B has finite multiplicity, if and only if, the functions y_i have $j_i(j_i < \infty)$ repetitions with respect to the spectral measure of S_i.*

Also

$$\max_i muB_i \leq \sum_{i=1}^{k} mu\, B_i \leq \sum_{i=1}^{k} j_i muS_i .$$

Proof. From the definition of multiplicity, as the smallest number of generating elements, it follows that

$$\max_i muB_i \leq muB \leq \sum_{i=1}^{k} muB_i .$$

Now, $B_i = y_i(S_i)$, hence the rest of the Theorem follows from Theorem 2.1. The problem of this section is reduced to the following

$$H = \sum_{i=1}^{k} E_iH \text{ where } E_iE_j = 0 \text{ if } i \neq j$$

and B_i = restriction B to E_iH, where the multiplicity of B_i is known. Now by decomposing each operator B_i into sum of operators of uniform multiplicity we will have $H = \sum_{i=1}^{m} H_i$, where the spaces H_i are mutually orthogonal, and C_i = restriction of B to H_i is an operator of uniform multiplicity. We shall show how to compute muB from muC_i by reducing this case to the one studied in Theorem 2.2.

Denote the projection on H_i by F_i. Let $E(B; \alpha_i)$ be the maximal projection such that

$$E(C_i; \alpha_i) = E(B; \alpha_i)F_i = 0 .$$

Such a projection exists by Zorn's Lemma. Finally let $\beta_i = \sigma(B) - \alpha_i$. On β_i the spectral measure of C_i can vanish only when the spectral measure of B vanishes. Now $E(B; \bigcup_{i=1}^m \beta_i) = I$ because $\sum_{i=1}^m F_i = I$.

The set $\sigma(B)$ can be decomposed into disjoint sets γ_j such that

a. Each γ_j is a subset of one of the sets β_{j_0}.

b. If $\gamma_j \cap \beta_i \neq \varphi$ then $\gamma_j \subset \beta_i$.

Assuming, for a moment, that this decomposition is given then

$$muB = \max_j mu(B \text{ restricted to } E(B; \gamma_j)H) .$$

But the multiplicity of B restricted to $E(B; \gamma_j)H$ is

$$\sum_{i|\gamma_j \subset \beta_i} mu(C_i \text{ restricted to } E(B; \gamma_j)H_i)$$

by Theorem 2.2.

We shall show how to choose the sets γ_i by an induction argument on the number m. Let $\gamma_1 = \beta_1 - \bigcup_{i \geq 2} \beta_1\beta_i$. This set (which might be void) satisfies conditions a and b. The rest of $\sigma(B)$ is

$$\Big(\bigcup_{i \geq 2} \beta_1\beta_i\Big) \cup \Big(\bigcup_{i \geq 2} (\beta_i - \beta_1)\Big)$$

In both sets there are only $m - 1$ subsets and by induction there exists a decomposition.

BIBLIOGRAPHY

1. N. Dunford, and J. Schwartz, *Linear Operators*, Vol. II. to appear.

2. S. R. Foguel, *Normal Operators of Finite Multiplicity*. Communications on Pure and Applied Mathematics, Vol. XI, (1958), p. 297.

3. P. R. Halmos, *Measure Theory*. D. Van Nostrand, New York, 1950.

UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA

UNITARY OPERATORS IN C^*-ALGEBRAS

JAMES G. GLIMM AND RICHARD V. KADISON

1. Introduction. We present several results concerning unitary operators in uniformly closed self-adjoint algebras of operators on a Hilbert space (C^*-algebras). Section 2 contains these results the key one of which (Theorem 1) asserts a form of transitivity for unitary operators in an irreducible C^*-algebra (an application of [2, Theorem 1]). Section 3 consists of some applications. The first (Corollary 8) is a clarification of the relation between unitary equivalence of pure states of a C-*algebra and of the representations they induce. The most desirable situation prevails: two pure states of a C^*-algebra are unitarily equivalent (i. e. conjugate via a unitary operator in the algebra) if and only if the representations they induce are unitarily equivalent. The second application (Corollary 9) provides a sufficient condition for two pure states ρ and τ to be unitarily equivalent: viz. $\|\rho - \tau\| < 2$. The final application (Theorem 11) is to the affirmative solution of the conjecture that the * operation is isometric in B^*-algebras [1].

We use the notation $\sigma(A)$ for the spectrum of A; C for the set of complex numbers of modulus 1; $\mathscr{S}^-$ for the strong closure of the set of operators $\mathscr{S}$; and ω_x for the state, $A \rightarrow (Ax, x)$, due to the unit vector x. Our C^*-algebras all contain the identity operator I.

2. Unitary operators. The theorem which follows establishes an n-fold transitivity property for the unitary operators in an irreducible C^*-algebra. Its relation to [2, Theorem 1] is clear—it is, in fact, the multiplicative version of the self-adjoint portion of that theorem.

THEOREM 1. *If $\mathfrak{A}$ is a C^*-algebra acting irreducibly on $\mathscr{H}$ and V is a unitary operator on $\mathscr{H}$ such that $Vx_k = y_k$, $k = 1, \cdots, n$, then there is a unitary operator U in $\mathfrak{A}$ such that $Ux_k = y_k$ and $\sigma(U) \neq C$.*

Proof. Passing to an orthonormal basis for the finite-dimensional space generated by $\{x_1, \cdots, x_n\}$, we see that there is no loss in generality if we assume that $\{x_1, \cdots, x_n\}$, and hence $\{y_1, \cdots, y_n\}$, are orthonormal sets. Moreover, employing a unitary extension of the mapping carrying x_j onto y_j, $j = 1, \cdots, n$ to the space generated by $\{x_1, y_1, \cdots, x_n, y_n\}$ and a diagonalizing basis for this unitary operator; we see that it suffices to consider the case in which $Vx_j = \beta_j x_j$, $|\beta_j| = 1, j = 1, \cdots, n$.

Choose real α_j in the half-open interval $(-\pi, \pi]$ such that $\exp i\alpha_j = \beta_j$, and let A be a self-adjoint operator in $\mathfrak{A}$ such that $Ax_j = \alpha_j x_j$ (such

Received April 27, 1959. The second author is an Alfred P. Sloan Fellow.

an operator exists by [2, Theorem 1]). Define $g(\alpha)$ as α for α in $[\min\{\alpha_j\}, \max\{\alpha_j\}]$, as $\min\{\alpha_j\}$ and $\max\{\alpha_j\}$, for $\alpha \leq \min\{\alpha_j\}$ and $\alpha \geq \max\{\alpha_j\}$, respectively. Then $g(A)$ is a self-adjoint operator in $\mathfrak{A}$ with spectrum in $[\min\{\alpha_j\}, \max\{\alpha_j\}]$, and $g(A)x_j = \alpha_j x_j$. It follows that $\exp ig(A)$ is a unitary operator U in $\mathfrak{A}$, $\sigma(U) \neq C$, and $Ux_j = \beta_j x_j$.

Another unitary analogue of a known result which seems of some value is the following variant of Kaplansky's Density Theorem [3, Theorem 1]. It is a consequence of Kaplansky's theorem and some commutative spectral theory.

THEOREM 2. *If $\mathscr{U}(\mathfrak{A}, k)$ is the set of unitary operators in the C^*-algebra $\mathfrak{A}$ whose distance from I does not exceed k, then $\mathscr{U}(\mathfrak{A}, k)^-$ contains $\mathscr{U}(\mathfrak{A}^-, k)$.*

Proof. Note that $\| U - I \| \leq k$, for a unitary operator U, if and only if $\sigma(U)$ is contained in $\{z\colon |z - 1| \leq k, |z| = 1\}$, a closed subset S_k of the unit circle C. From spectral theory, each unitary operator is a uniform limit of unitary operators which are finite linear combinations of orthogonal spectral projections for it, and which do not have -1 in their spectra (i. e. whose distance from I is less than 2). Thus, it suffices to consider the case where $k < 2$.

Assuming $k < 2$, let $\arg z$ be that number in the open interval $(-\pi, \pi)$ such that $z = \exp[i \arg z]$, for z in S_k; and let f be a continuous extension of arg to C. If $a = 2\sin^{-1}(k/2)$ and $\mathscr{S}(\mathfrak{A}, a)$ denotes the set of self-adjoint operators in $\mathfrak{A}$ with norm not exceeding a, then f maps $\mathscr{U}(\mathfrak{A}, k)$ into $\mathscr{S}(\mathfrak{A}, a)$ continuously in the strong topology, $U = \exp[if(U)]$, and exp maps $i\mathscr{S}(\mathfrak{A}, a)$ into $\mathscr{U}(\mathfrak{A}, k)$ continuously in the strong topology, from spectral theory, [3, Lemma 3], and [3, Lemma 2]. Thus, if U lies in $\mathscr{U}(\mathfrak{A}^-, k)$, $f(U)$ lies in $\mathscr{S}(\mathfrak{A}^-, a)$ and is a strong limit point of $\mathscr{S}(\mathfrak{A}, a)$, from [3, Theorem 1]; so that $U(= \exp[if(U)])$ is a strong limit point of $\mathscr{U}(\mathfrak{A}, k)$.

In the next lemma, we make use of Mackey's concept of disjoint representations [5]. These are *-representations of self-adjoint operator algebras which have no unitarily equivalent non-zero subrepresentations (the restriction of the representation to an invariant subspace). The application in [5] is to unitary representations of groups and ours is to *-representations of algebras—the difference is slight, however; and our lemma and proof are valid for groups.

LEMMA 3. *If $\{\phi_\alpha\}$ are *-representations of the self-adjoint operator algebra $\mathfrak{A}$, then $\{\phi_\alpha\}$ consists of mutually disjoint representations if and only if $\phi(\mathfrak{A})^- = \sum \oplus (\phi_\alpha(\mathfrak{A})^-)$, where $\phi = \sum \oplus \phi_\alpha$.*

Proof. Suppose $\phi_\alpha(\mathfrak{A})$ acts on $\mathscr{H}_\alpha$, $\mathscr{H} = \sum \oplus \mathscr{H}_\alpha$, and P_α is the orthogonal projection of $\mathscr{H}$ upon $\mathscr{H}_\alpha$. If $\phi(\mathfrak{A})^- = \sum \oplus (\phi_\alpha(\mathfrak{A})^-)$, and U is a partial isometry [6] of $E_\alpha(\mathscr{H}_\alpha)$ onto $E_{\alpha'}(\mathscr{H}_{\alpha'})$, where $\alpha \neq \alpha'$ and E_α, $E_{\alpha'}$ are projections commuting with $\phi_\alpha(\mathfrak{A})$ and $\phi_{\alpha'}(\mathfrak{A})$, respectively, such that $U\phi_\alpha(A)U^* = \phi_{\alpha'}(A)E_{\alpha'}$ for each A in $\mathfrak{A}$, then U commutes with $\phi(\mathfrak{A})$. In fact, $U\phi(A) = U\phi_\alpha(A) = \phi_{\alpha'}(A)U = \phi(A)U$. Thus U commutes with $\sum \oplus \phi_\alpha(\mathfrak{A})$ and, in particular, with each P_α. But $UP_\alpha = U = P_\alpha U = 0$; so that $0 = E_\alpha = E_{\alpha'}$ and $\{\phi_\alpha\}$ consists of mutually disjoint representations.

If the ϕ_α are mutually disjoint and V is a partial isometry in the commutant of $\phi(\mathfrak{A})$ with the initial space E_α in P_α and final space $E_{\alpha'}$ in $P_{\alpha'}$ (cf. [6]); then $VE_\alpha\phi_\alpha(A)E_\alpha V^* = VV^*V\phi(A)V^*VV^* = E_{\alpha'}\phi(A)VV^*E_{\alpha'} = \phi_{\alpha'}(A)E_{\alpha'}$, for each A in $\mathfrak{A}$. Thus, by disjointness, E_α and $E_{\alpha'}$ are 0. It follows that the central carrier of P_α is orthogonal to that of each $P_{\alpha'}$, and hence to each $P_{\alpha'}$, with $\alpha' \neq \alpha$ (see [4], for example). Since $\sum_\alpha P_\alpha = I$, and the central carrier of P_α contains P_α, P_α is its own central carrier. In particular, P_α lies in the center of the commutant and therefore in $\phi(\mathfrak{A})^-$. It is immediate from this that $\phi(\mathfrak{A})^- = \sum \oplus (\phi_\alpha(\mathfrak{A})^-)$.

Since the commutant of an irreducible representation consists of scalars, two such are either unitarily equivalent or disjoint. From this and Lemma 3, we have as an immediate consequence:

COROLLARY 4. *If $\{\phi_\alpha\}$ is a family of irreducible *-representations of a self-adjoint operator algebra $\mathfrak{A}$, no two of which are unitarily equivalent, then $\phi(\mathfrak{A})^- = \sum \oplus \mathscr{B}_\alpha$, where $\phi = \sum \oplus \phi_\alpha$ and $\mathscr{B}_\alpha$ is the algebra of all bounded operators on the representation space of ϕ_α.*

We shall need a result asserting the possibility of "lifting" unitary operators from a representing algebra to the original algebra under certain circumstances.

LEMMA 5. *If ϕ is a *-representation of the C^*-algebra $\mathfrak{A}$ and U is a unitary operator in $\phi(\mathfrak{A})$ with $\sigma(U) \neq C$, there is a unitary operator U_0 in $\mathfrak{A}$ such that $\phi(U_0) = U$.*

Proof. As in Theorem 2, we can find a continuous function f on C such that $f(U)$ is self-adjoint and $\exp\,[if(U)] = U$. Let A be a self-adjoint operator in $\mathfrak{A}$ such that $\phi(A) = f(U)$. (Recall that $\phi(\mathfrak{A})$ is a C^*-algebra, so that $f(U)$ lies in $\phi(\mathfrak{A})$. If $\phi(B) = f(U)$ then A may be chosen as $(B^* + B)/2$.) If $U_0 = \exp iA$ then $\phi(U_0) = \exp\,[if(U)] = U$, by uniform continuity of ϕ.

REMARK 6. It may not be possible to lift a given unitary operator (as indicated by the condition $\sigma(U) \neq C$ in Lemma 5). In fact, illustrating this with commutative C^*-algebras, we may deal with the algebras of continuous complex-valued functions on compact Hausdorff spaces and unitary functions on them (functions with modulus 1). View C as the equator of a two-sphere S, and let α be the inclusion mapping of C into S. Then α induces a homomorphism of the function algebra of S onto that of C ("onto" by the Tietze Extension Theorem) which is, of course, the mapping that restricts a function on S to C. The identity mapping of C onto C is a unitary function on C which does not have a continuous unitary extension to S; for such an extension restricted to one hemisphere would amount to a retraction of the disk onto its boundary.

As a corollary to the foregoing considerations, we have the following extension of Theorem 1:

COROLLARY 7. *If $\{\phi_\alpha\}$ is a family of unitarily inequivalent irreducible *-representations of the C^*-algebra $\mathfrak{A}$ on Hilbert spaces $\{\mathscr{H}_\alpha\}$, ϕ is the direct sum of $\{\phi_\alpha\}$, $\mathscr{H}$ of $\{\mathscr{H}_\alpha\}$, $\{x_1, \cdots, x_n\}$ and $\{y_1, \cdots, y_n\}$ are two finite sets of vectors with $\{x_1, \cdots, x_n\}$ linearly independent and each x_j and corresponding y_j in some $\mathscr{H}_\alpha$; then there is an A in $\mathfrak{A}$ such that $\phi(A)x_j = y_j$. If $Bx_j = y_j$ for some self-adjoint or unitary operator B on $\mathscr{H}$ then A may be chosen self-adjoint or unitary, respectively.*

Proof. The argument of [2, Theorem 1] applies directly to the first assertion once we note that the general constructions and norm estimates of that theorem can be performed on each $\mathscr{H}_\alpha$, since each x_j, y_j lie in some $\mathscr{H}_\alpha$; and the strong approximations are valid by virtue of Corollary 4. With B self-adjoint, each $P_\alpha B P_\alpha$ is self-adjoint and $P_\alpha B P_\alpha x_j = y_j$ (for x_j, y_j in $\mathscr{H}_\alpha$), where P_α is the projection of $\mathscr{H}$ onto $\mathscr{H}_\alpha$; so that the argument of [2, Theorem 1], in the self-adjoint case, applies to give a self-adjoint operator $\phi(A)$ such that $\phi(A)x_j = y_j$, $j = 1, \cdots, n$. Of course, A may be chosen self-adjoint in this case. If B is unitary it can be replaced by one which maps each $\mathscr{H}_\alpha$ onto itself and acts in the same way on $\{x_1, \cdots, x_n\}$ (extend the mappings of x_j onto y_j on each $\mathscr{H}_\alpha$). Having the self-adjoint result, in this case, the argument of Theorem 1 now applies to give a unitary operator $\phi(U)$ such that $\phi(U)x_j = y_j$, $j=1, \cdots, n$, and $\sigma[\phi(U)] \neq C$. From Lemma 5, U may be chosen as a unitary operator in $\mathfrak{A}$.

3. Some applications. The next result indicates that the most favorable situation obtains with regard to the relation between pure states which give rise to unitarily equivalent representations.

COROLLARY 8. *If ρ and τ are pure states of the C^*-algebra $\mathfrak{A}$, then ρ and τ induce unitarily equivalent representations of $\mathfrak{A}$ if and only if there is a unitary operator U in $\mathfrak{A}$ such that $\rho(A) = \tau(U^*AU)$ for each A in $\mathfrak{A}$.*

Proof. If such a U exists, and ϕ_ρ and ϕ_τ are the representations due to ρ and τ with left kernels $\mathscr{J}$ and $\mathscr{K}$, respectively; then the mapping V of $\mathfrak{A}/\mathscr{J}$ onto $\mathfrak{A}/\mathscr{K}$ defined by, $V(A + \mathscr{J}) = AU + \mathscr{K}$, is an isometric mapping of a dense subset of the representation space for ρ onto a dense subset of the representation space for τ, since $\rho(A^*A) = \tau(U^*A^*AU)$. Thus V has a unitary extension mapping one representation space onto the other. Moreover, $V^{-1}\phi_\tau(B)V(A+\mathscr{J}) = V^{-1}(BAU+\mathscr{K}) = BA + \mathscr{J} = \phi_\rho(B)(A + \mathscr{J})$, whence the unitary extension of V implements a unitary equivalence between ϕ_τ and ϕ_ρ.

Suppose, now, that V implements a unitary equivalence between ϕ_ρ and ϕ_τ, that x and y are unit vectors in the representation spaces for ϕ_ρ and ϕ_τ, respectively, such that $\omega_x\phi_\rho = \rho$ and $\omega_y\phi_\tau = \tau$, and that $U_0Vy = x$, with U_0 a unitary operator in $\phi_\rho(\mathfrak{A})$ such that $\sigma(U_0) \neq C$ (cf. Theorem 1). Let U be a unitary operator in $\mathfrak{A}$ such that $\phi_\rho(U) = U_0$ (cf. Lemma 5). Then

$$\begin{aligned}\rho(A) = \omega_x\phi_\rho(A) = (\phi_\rho(A)U_0Vy, U_0Vy) = (VV^{-1}\phi_\rho(U^*AU)Vy, Vy) \\ = \omega_y\phi_\tau(U^*AU) = \tau(U^*AU) ,\end{aligned}$$

for each A in $\mathfrak{A}$.

COROLLARY 9. *If ρ and τ are pure states of a C^*-algebra $\mathfrak{A}$ such that $\|\rho - \tau\| < 2$, then ρ and τ give rise to unitarily equivalent representations of $\mathfrak{A}$.*

Proof. If ϕ_ρ and ϕ_τ are unitarily inequivalent and ϕ, their direct sum, represents $\mathfrak{A}$ on the direct sum $\mathscr{H}$ of $\mathscr{H}_\rho$ and $\mathscr{H}_\tau$, then there are unit vectors x and y in $\mathscr{H}_\rho$ and $\mathscr{H}_\tau$, respectively, such that $\rho = \omega_x\phi$ and $\tau = \omega_y\phi$. According to Corollary 7, we can find U in $\mathfrak{A}$ such that $\phi(U)x = x$ and $\phi(U)y = -y$ (approximation using Theorem 2 would do). Then $|(\rho - \tau)U| = |(\phi(U)x, x) - (\phi(U)y, y)| = 2$; so that $\|\rho - \tau\| = 2$ (recall that $\|\rho\| = \|\tau\| = 1$, since ρ and τ are states).

REMARK 10. The condition $\|\rho - \tau\| < 2$ noted above is not necessary for unitary equivalence. Indeed, if x and y are orthogonal unit vectors in a Hilbert space $\mathscr{H}$, E is a projection with x in its range and y orthogonal to its range, then $(\omega_x - \omega_y)(2E - I) = 2$, so that $\|\omega_x - \omega_y\| = 2$; while ω_x and ω_y give rise to unitarily equivalent representations of the algebra of all bounded operators on $\mathscr{H}$ (both, in fact, equivalent to the given representation on $\mathscr{H}$). On the other hand, $|(\omega_x - \omega_y)(A)| \leq |(Ax, x - y)| + |(A(x - y), y)| \leq 2\|x - y\|$, when $\|A\| \leq 1$; so that there are pure states giving rise to unitarily equivalent representations the norms of whose differences are as small as we please.

Our next application is to the solution of a minor problem raised by

Gelfand and Neumark in connection with their conditions for a Banach algebra to be isomorphic (and isometric) to a C^*-algebra [1]. In [1], six conditions are listed for this to be the case—the first three being the standard algebraic conditions for a $*$ operation defined on a Banach algebra, viz. $(\alpha a + b)^* = \bar{\alpha}a^* + b^*$, $(ab)^* = b^*a^*$, and $(a^*)^* = a$; the fourth, $\| a^*a \| = \| a^* \| \cdot \| a \|$, is the critical condition relating the metric structure to the $*$ operation; the fifth, $\| a^* \| = \| a \|$, asserts the isometric character of the $*$ operation; and the sixth, the so-called "symmetry" condition, assumes that $a^*a + e$ has an inverse. Gelfand and Neumark conjectured that both the fifth and sixth conditions are consequences of the first four. After much preliminary work (notably by I. Kaplansky), the symmetry question was reduced to showing that the sum of two self-adjoint elements with non-negative spectrum is again such an element. This was done independently by Kelley-Vought and Fukamya (though not recognized as the missing information—Kaplansky pointed this out). We noted that this had been effected without assuming the $*$ operation is isometric, and went on to prove that it was, accordingly, isometric on regular elements. From this, its continuity followed; and one could derive all but the isometric character of the isomorphism in the Gelfand-Neumark theorem, with a little care. During some seminar lectures, we noted, some years ago that the symmetry condition could be derived in a quite natural way in the course of the imbedding proof. The last loose end, establishing the fully isometric character of the $*$ operation, can be tied by the results of this paper. The closing of this last gap would seem to be an appropriate occasion for presenting the finished result in its entirety. From another viewpoint, the supression of the fifth condition introduces subtle traps into these considerations —statements which are made in complete safety with operators require delicate proof in the present circumstances (e. g. despite the Gelfand-Neumark commutative result, we cannot take the commutative case as settled; for the uniform closure of the real algebra generated by a single self-adjoint element is not known *a priori* to consist entirely of self-adjoint elements, since continuity of the $*$ operation is missing—again, the Schwarz inequality for states will not yield the fact that they have norm 1, under these circumstances).

By a B^*-algebra, we shall mean a Banach algebra with unit element e and normalized norm ($\| e \| = 1$, $\| ab \| \leqq \| a \| \cdot \| b \|$) which has a $*$ operation satisfying the first four conditions noted above. An element a is self-adjoint, unitary, positive, or regular, when $a = a^*$, $a^*a = aa^* = e$, $a = a^*$ and the spectrum $\sigma(a)$ of a consists of non-negative real numbers, or a has an inverse, respectively. A state of a B^*-algebra is a linear functional which is 1 at e and real, non-negative on positive elements. We make use of the Hahn-Banach theorem from normed space theory and the following standard facts about complex Banach algebras with a unit and a normalized norm: the spectral radius $r(a)$ of an element a (i.e. $\sup \{|\alpha|: \alpha \in \sigma(a)\}$) is $\lim_{n\to\infty} \| a^n \|^{1/n}$ (and does not exceed $\| a \|$); $e+a$ is regular if $\| a \| < 1$; $\sigma(p(a)) = p(\sigma(a)) (= \{p(\alpha): \alpha \in \varepsilon(a)\})$ for each polynomial p; and the quotient modulo a maximal ideal is the complex numbers, for a commutative algebra.

THEOREM 11 (Gelfand-Neumark). *A B^*-algebra $\mathfrak{A}$ is isometric and $*$-isomorphic with a C^*-algebra.*

Proof. If $a^* = a$ then $\| a^2 \| = \| a \|^2$ so that $\| a^{2^n} \| = \| a \|^{2^n}$ and $\| a \| = r(a)$. Since $p(a)$ is self-adjoint for each real polynomial p, $\| p(a) \| = r(p(a)) = \sup \{| p(\alpha) |: \alpha \text{ in } \sigma(a)\}$. If p is complex then $p = p_1 + ip_2$, with p_1, p_2 real, and

$$r((p_1^2 + p_2^2)(a)) \le r((p_1 - ip_2)(a)) \cdot r((p_1 + p_2)(a)) \le \| [(p_1 + ip_2)(a)]^* \| \cdot \| (p_1 + ip_2)(a) \| = \| (p_1^2 + p_2^2)(a) \| = r((p_1^2 + p_2^2)(a)) .$$

Thus, equality holds throughout; and since

$$r((p_1 - ip_2)(a)) \le \| [(p_1 + ip_2)(a)]^* \| ,$$
$$r((p_1 + ip_2)(a)) \le \| (p_1 + ip_2)(a) \| ,$$

equality must hold in each. Hence $\| p(a) \| = r(p(a)) = \sup \{| p(\alpha) |: \alpha \in \sigma(a)\}$, for complex polynomials p and self-adjoint elements a. The mapping carrying an element $p(a)$ onto the polynomial p on $\sigma(a)$ is an isomorphism of the algebra of (complex) polynomials in a into $C(\sigma(a))$ and has an isometric isomorphism extension mapping the closure $\mathfrak{A}(a)$ onto the closure P of the polynomials in $C(\sigma(a))$.

If $\alpha \in \sigma(a)$, then the mapping $g \to g(\alpha)$ (g in P) is a linear functional of norm 1 on P which assigns 1 to the image in P of e and α to that of a. Via the isometry, this gives rise to a linear functional f_0 of norm 1 on $\mathfrak{A}(a)$, such that $f_0(e) = 1, f_0(a) = \alpha$. Let f be a norm 1 extension of f_0 to $\mathfrak{A}$. If b is self-adjoint and $f(b)$ is not real, by adding a suitable real multiple of e to b we arrive at a self-adjoint element on which f takes a non-zero imaginary value. Suppose $f(b) = i\beta$, with $\beta > 0$ (if $\beta < 0$, use $-b$). Then

$$| f(b + ine) |^2 = \beta^2 + 2\beta n + n^2 \le \| b + ine \|^2 = [r(b + ine)]^2$$
$$= (r(b + ine))(r([b + ine]^*)) = \| b + ine \| \cdot \| b - ine \| = \| b^2 + n^2 e \|$$
$$\le \| b^2 \| + n^2 ,$$

which is absurd for $n > (\| b^2 \| - \beta^2)/2\beta$ (note that $r(c) = r(c^*)$, for each c, since $\sigma(c^*) = \sigma(c)$). Thus f is real on each self-adjoint element. In particular, $f(a) = \alpha$ is real, and $\sigma(a)$ consists of real numbers. Hence, the algebra of complex polynomials on $\sigma(a)$ is invariant under complex conjugation, the Stone-Weierstrass theorem applies, and P is $C(\sigma(a))$. If $b \geq 0$ and $f(b) < 0$, then $\sigma(b - \|b\| e) = \sigma(b) - \|b\|$. Since $\sigma(b) \geq 0$, $r(b - \|b\| e) = \|b - \|b\| e\| \leq \|b\|$. But $|f(b - \|b\| e)| = |f(b) - \|b\|| > \|b\| \geq \|b - \|b\| e\|$, contradicting $\|f\| = 1$. Thus f is a state of $\mathfrak{A}$, $f(a) = \alpha$, and f has norm 1.

If $a_1, \cdots, a_n$ are positive and $\alpha \in \sigma(a_1 + \cdots + a_n)$ there is a state f of $\mathfrak{A}$ such that $\alpha = f(a_1 + \cdots + a_n) = f(a_1) + \cdots + f(a_n) \geq 0$, so that $a_1 + \cdots + a_n$ is positive. If b is self-adjoint and has an inverse, then 0 is not in $\sigma(b)$; so that the image of b in $C(\sigma(b))$ has an inverse; and the inverse of b lies in $\mathfrak{A}(b)$. If b is in $\mathfrak{A}(a)$, with a self-adjoint, then $\mathfrak{A}(b)$ is contained in $\mathfrak{A}(a)$ and the inverse of b lies in $\mathfrak{A}(a)$. Thus the spectrum of a self-adjoint element in $\mathfrak{A}(a)$ is the same relative to $\mathfrak{A}(a)$ and to $\mathfrak{A}$. In view of the isomorphism between $\mathfrak{A}(a)$ and $C(\sigma(a))$, this spectrum is the range of its representing function in $C(\sigma(a))$. Thus $a^2 \geq 0$ for a self-adjoint. With a and b positive, choose a state f of norm 1 such that $f(a) = r(a) = \|a\|$, then $\|a + b\| \geq f(a + b) \geq f(a) = \|a\|$.

Suppose next that $\mathfrak{A}_0$ is a subalgebra of $\mathfrak{A}$ which is maximal with respect to the properties of being abelian and self-adjoint (i.e. $\mathfrak{A}_0^* = \mathfrak{A}_0$). If b commutes with $\mathfrak{A}_0$ then $ba^* = a^*b$, for each a in $\mathfrak{A}_0$; so that $b^*a = ab^*$. Thus, the self-adjoint elements $b + b^*$ and $(b - b^*)/i$ commute with $\mathfrak{A}_0$, and, by maximality, lie in $\mathfrak{A}_0$. Hence $b (= (b + b^*)/2 + i(b - b^*)/2i)$ lies in $\mathfrak{A}_0$; and $\mathfrak{A}_0$ is maximal with respect to the property of being abelian. It follows that $\mathfrak{A}_0$ is closed. If b is the limit of self-adjoint elements in $\mathfrak{A}_0$ and $b = b_1 + ib_2$ with b_1 and b_2 self-adjoint (in $\mathfrak{A}_0$ — the decomposition just noted), then

$$\begin{aligned}\|b_2^2\| &\leq \|(b_1 - a)^2 + b_2^2\| = \|b_1 + ib_2 - a\| \cdot \|b_1 - ib_2 - a\| \\ &\leq \|b - a\| \cdot (\|b^*\| + \|a\|),\end{aligned}$$

with a self-adjoint in $\mathfrak{A}_0$. Choosing a near b, we see that $\|b_2^2\| (= \|b_2\|^2)$ is dominated by an arbitrarily small quantity, so that $b_2 = 0$. Thus b is self-adjoint, and the self-adjoint elements in $\mathfrak{A}_0$ are closed. If a is self-adjoint, the polynomials in a from a commutative self-adjoint algebra which can be imbeded in a maximal one $\mathfrak{A}_0$ (Zorn's Lemma). Since $\mathfrak{A}_0$ is closed, $\mathfrak{A}(a)$ is contained in it. Thus, the closure of the real polynomials in a (which maps onto the algebra of real functions in $C(\sigma(a))$) consists of self adjoint elements.

The isomorphism of $\mathfrak{A}(a^*a)$ with $C(\sigma(a^*a))$ establishes the existence of positive elements b and c in $\mathfrak{A}(a^*a)$ such that $a^*a = b - c$, and $bc = 0$. Thus $(ac)^*(ac) = -c^3$, which is negative, so that $(ac)(ac)^*$ is negative. (In an arbitrary ring with a unit, if c is the inverse to $e - ab$ then $e + bca$ is the inverse to $e - ba$; so that, in a Banach algebra, the spectra of ab and ba with 0 adjoined is the same set.) But with $ac = a_1 + ia_2$, a_1 and a_2 self-adjoint,

$$0 \geq (ac)(ac)^* + (ac)^*(ac) = 2(a_1^2 + a_2^2) \geq 0.$$

Thus,

$$0 = \| a_1^2 + a_2^2 \| \geq \| a_1 \|^2, \| a_2 \|^2;$$

so that $a_1 = a_2 = ac = c^3 = c = 0$, and $a^*a \geq 0$. The function representing a^*a in $C(\sigma(a^*a))$ is real and non-negative and therefore has a continuous non-negative square root. This square root corresponds to an element $(a^*a)^{1/2}$ which is a positive square root of a^*a in $\mathfrak{A}(a^*a)$. If a is regular so is a^* and $(a^*a)^{1/2}$. The element $a(a^*a)^{-1/2}(=u)$ is unitary, since $uu^* = a(a^*a)^{-1}a^* = aa^{-1}a^{*-1}a^* = e$ and $u^*u = (a^*a)^{-1/2}a^*a(a^*a)^{-1/2} = (a^*a)(a^*a)^{-1} = e$. Extend the self-adjoint abelian algebra of polynomials in u and u^* to a maximal one $\mathfrak{A}_0$; and let M be a proper maximal ideal in $\mathfrak{A}_0$. Then $b + M = b(M)e + M$, for some complex number $b(M)$ (i.e. $\mathfrak{A}_0/M$ is the complex numbers), $(cb)(M) = c(M)b(M)$, $b(M)$ is in the spectrum of b relative to $\mathfrak{A}_0$, and if $b = b_1 + ib_2$ with b_1 and b_2 self-adjoint then $b^*(M) = b_1(M) - ib_2(M) = \overline{b(M)}$, since the spectra of b_1 and b_2 are real. Thus $1 = e(M) = (u^*u)(M) = |u(M)|^2$. Now $\| u \| \geq r(u) \geq 1$, and similarly, $\|u^*\| \geq 1$. But $1 = \| e \| = \| u^*u \| = \| u^* \| \cdot \| u \|$, so that $\| u \| = 1$. Hence

$$\| a \|^2 = \| u(a^*a)^{1/2} \|^2 \leq \| (a^*a)^{1/2} \|^2 = \| a^*a \| = \| a^* \| \cdot \| a \|;$$

and $\| a \| \leq \| a^* \|$, symmetrically, $\| a^* \| \leq \| a \|$, so that $\| a \| = \| a^* \|$ (i.e. the * operation is isometric on regular elements). If $\| b \| < 1$, then $e + b$ is regular, so that $\| b^* \| - 1 \leq \| e + b^* \| = \| e + b \| \leq 1 + \| b \| < 2$. Thus the * operation is continuous (bounded), and the self-adjoint elements in $\mathfrak{A}$ form a closed set.

If f is a state of $\mathfrak{A}$, the mapping $a, b \rightarrow f(b^*a)$ is a positive semi-definite inner product on $\mathfrak{A}$ (write (a, b) for the inner product of a and b). If a is a null vector then $(ba, ba) = (a, b^*ba) = 0$ (from the Schwarz inequality); so that ba is a null vector. Thus, the set $\mathscr{I}$ of null vectors is a left ideal in $\mathfrak{A}$ (the "left kernel" of f). The quotient vector space $\mathfrak{A}/\mathscr{I}$ has a positive definite inner product induced on it from that on $\mathfrak{A}$. Let $\mathscr{H}$ be the (Hilbert space) completion of $\mathfrak{A}/\mathscr{I}$ in this inner product. Define the operator $\phi(a)$ on $\mathfrak{A}/\mathscr{I}$ by $\phi(a)(b + \mathscr{I}) = ab + \mathscr{I}$, for each a in $\mathfrak{A}$. If $c \geq 0$, then $(b^*c^{1/2})(c^{1/2}b) \geq 0$. With $\| a^*a \|e - a^*a$ in place of c, we have $\| a^*a \| b^*b \geq b^*a^*ab$; so that $3 \| a \|^2 (b + \mathscr{I}, b + \mathscr{I}) \geq \| a^*a \| f(b^*b) \geq f(b^*a^*ab) = (\phi(a)(b + \mathscr{I}), \phi(a)(b + \mathscr{I}))$, and $\| \phi(a) \| \leq 3^{1/2} \| a \|$ ($\| \phi(a) \| \leq \| a \|$ if a is self-adjoint). Thus $\phi(a)$ has a unique extension to $\mathscr{H}$, with the same bound, which we denote again by $\phi(a)$. Since $(\phi(a)(b + \mathscr{I}), c + \mathscr{I}) = f(c^*ab) = (b + \mathscr{I}, \phi(a^*)(c + \mathscr{I}))$, $\phi(a)^* = \phi(a^*)$. It follows that ϕ is a *-representation of $\mathfrak{A}$ in the algebra of bounded operators on $\mathscr{H}$. If $\phi(a) = 0$ then $f(a) = (\phi(a)(e + \mathscr{I}), e + \mathscr{I}) = 0$.

If we perform this construction for each state of $\mathfrak{A}$, the direct sum ψ of the resulting *-representations is a *-isomorphism of $\mathfrak{A}$. In fact, if $\psi(a) = 0$, then $\psi(a^*a) = 0$, so that $f(a^*a) = 0$ for each state f of $\mathfrak{A}$. But there is a state f such that $f(a^*a) = \| a^*a \| = \| a^* \| \cdot \| a \|$. Thus $a = 0$. If b is self-adjoint $\| \psi(b) \| \leq \| b \|$, since each of the representations is norm decreasing on b. With f a state of $\mathfrak{A}$ such that $\| b \| = | f(b) | = | (\phi(b)(e + \mathscr{J}), e + \mathscr{J}) |$, however, we see that $\| \phi(b) \| \geq \| b \|$; so that $\| \psi(b) \| = \| b \|$. Since the self-adjoint elements in $\mathfrak{A}$ are closed (hence complete) they are complete (hence closed) in $\psi(\mathfrak{A})$; whence $\psi(\mathfrak{A})$ is closed (i.e. a C^*-algebra). If a is regular, $\| a \|^2 = \| a^*a \| = \| \psi(a^*a) \| = \| \psi(a) \|^2$.

Defining $||| \psi(b) |||$ to be $\| b \|$, $\psi(\mathfrak{A})$ has two norms ($||| \ |||$, and its operator norm $\| \ \|$) relative to which it is a B^*-algebra. These norms agree on self-adjoint and regular elements. If we show that they agree everywhere (i.e. that ψ is isometric) then the * operation is isometric on $\mathfrak{A}$ since it is preserved by ψ and is isometric on $\psi(\mathfrak{A})$. We write $\mathfrak{A}$ in place of $\psi(\mathfrak{A})$($\mathfrak{A}$ is a C^*-algebra with the two B^*-norms as described). As the first step, we establish the formula $\| A \| = \sup \{| f(UAV) |$: U and V unitary operators in $\mathfrak{A}$ and f a pure state of $\mathfrak{A}\}$. Since each state of $\mathfrak{A}$ has norm 1 (from the Schwarz inequality) relative to the operator norm, $| f(UAV) | \leq \| UAV \| \leq \| A \|$. On the other hand, if ϕ is the (irreducible) representation induced by f, $| f(U^*AV) | = |(\phi(A)\phi(V)x, \phi(U)x) |$, where x is a unit vector (in fact, the special one corresponding to $I + \mathscr{J}$). In view of Theorem 1 (or Theorem 2), $\sup \{| f(U^*AV) |$: U and V unitary operators in $\mathfrak{A}\} = \sup \{| (\phi(A)x, y) |: \| x \| = \| y \| = 1\} = \| \phi(A) \|$. Now the direct sum of the *-representations due to each pure state of $\mathfrak{A}$ is a *-isomorphism and hence an isometry of $\mathfrak{A}$; so that $\sup \{\| \phi(A) \|$: f a pure state of $\mathfrak{A}\} = \| A \|$, and our formula follows.

Each state of $\mathfrak{A}$ has norm 1 relative to the norm $||| \ |||$; for if $||| B ||| < 1$ and $f(B) = | f(B) | \alpha$ (where $| \alpha | = 1$), then $\bar{\alpha}B + I$ is regular. Hence, $| f(\bar{\alpha}\beta + I) | = | f(B) | + 1 \leq \| \bar{\alpha}B + I \| = ||| \bar{\alpha}\beta + I ||| < 2$. Thus $| f(UAV) | \leq ||| UAV ||| \leq ||| A |||$; and $\| A \| \leq ||| A |||$, for each A. But $\| A^* \| \cdot \| A \| = \| A^*A \| = ||| A^*A ||| = ||| A^* ||| \cdot ||| A |||$; so that $\| A \| = ||| A |||$ for each A. The proof is complete.

References

1. I. Gelfand and M. Neumark, *On the imbedding of normed rings into the ring of operators in Hilbert space*, Rec. Math. (Mat. Sbornik) N. S., **12** (1943), 197–213.

2. R. Kadison, *Irreducible operator algebras*, Proc. Nat. Acad. Sci. (USA), **43** (1957), 273–276.

3. I. Kaplansky, *A theorem on rings of operators*, Pacific J. Math., **1** (1951), 227–232.

4. I. Kaplansky, *Projections in Banach algebras*, Ann. Math., **53** (1951), 235-249.

5. G. Mackey, *Induced representations of locally compact groups* II. *The Frobenius reciprocity theorem*, Ann. Math., **58** (1953), 193-221.

6. F. Murray and J. von Neumann, *On rings of operators*, Ann. Math., **37** (1936), 116–229.

COLUMBIA UNIVERSITY

MEASURES DEFINED BY ABSTRACT L_p SPACES

HUGH GORDON

Let a linear space L of real-valued functions on a set E and a semi-norm on L be given. We shall consider when there exists a countably additive measure on E such that L is L_p with respect to this measure. We shall prove that certain conditions are sufficient for the measure to exist; it is obvious that these conditions are necessary. (We consider only the case where the constant function $1 \in L$.)

We need not assume that the elements of L are functions on a set. If we do not make this assumption, we use a theorem of Kakutani ([3], p. 998) to construct a representation for L as a space of continuous functions on a compact Hausdorff space. If, however, the elements of L are given as functions, we leave this preëstablished representation unchanged, even when it is not the one given by Kakutani's theorem.

The case where $p = 1$ and the elements of L are not given as functions was treated by Kakutani [2]. The case $p = 2$ will receive special attention at the end of the present paper. In this latter case, one may replace some of the hypotheses of the general case by the hypothesis that the semi-norm on L arises from a positive semi-definite bilinear form.

Let L be a Riesz space whose elements are functions on a set E. That is, let L be a set of real-valued functions on E which contains with f, g:

(a) $f + g$ defined by $(f + g)(x) = f(x) + g(x)$,

(b) αf defined by $(\alpha f)(x) = \alpha[f(x)]$, for each real number α,

(c) $f \wedge g$ defined by $(f \wedge g)(x) = \min(f(x), g(x))$,

and (d) $f \vee g$ defined by $(f \vee g)(x) = \max(f(x), g(x))$.

We denote $f \vee 0$ by f^+ and $(-f) \vee 0$ by f^-. (The case where L is an abstract Banach lattice will be considered shortly.)

Let p be a fixed real number ≥ 1. Throughout the paper, p will always stand for this fixed number. We suppose there is a semi-norm, which we denote by $\| \;\; \|$, defined on L. We further suppose:

(1) L is complete. That is, if $f_1, f_2, \cdots \in L$ are such that $\|f_n - f_m\|$ is small for large n, m; then there is a $g \in L$ such that $\|g - f_n\| \to 0$.

(2) For each $f \in L$, $\| |f| \| = \|f\|$.

(3) If f, g are positive, $\|f + g\|^p \geq \|f\|^p + \|g\|^p$.

(4) If f, g are positive and $f \wedge g = 0$, $\|f + g\|^p \leq \|f\|^p + \|g\|^p$.

Received April 27, 1959. The material of this paper was presented to the American Mathematical Society in August, 1957. It consists of part of the author's doctoral dissertation, which was prepared at Columbia Univeristy under the direction of Professor E. R. Lorch. Part of the work was done under National Science Foundation grant NSF G 1981.

(5) $1 \in L$ and $|| 1 || = 1$. (Here 1 denotes the constant function 1.)

We note that if $f, g \in L$ and $0 \leq f \leq g$, then $|| f || \leq || g ||$; since $|| f ||^p \leq || f ||^p + || g - f ||^p \leq || f + g - f ||^p = || g ||^p$ by (3) above. We also note that, for each $f \in L$, $|| f^+ || \leq || f ||$; since $|| f^+ || \leq || f^+ + f^- || = || \, | f | \, || = || f ||$ using (2) and the preceding remark.

We now briefly consider abstract L_p spaces. Let L be a Riesz space (i. e. a vector lattice), whose elements need not be functions. Suppose there is a norm on L. (If a semi-norm is given instead of a norm, we use, in place of L, the quotient space of L modulo the elements of norm 0. This quotient space will be a normed Riesz space provided the semi-norm satisfies (2) and (3) above.) Suppose, for some $p \geq 1$, that L has properties (1)-(4) above. Instead of (5), we suppose that L has a weak unit, i. e.:

(5′) There is a positive $e \in L$ such that $f \wedge e = 0$, $f \in L$ imply $f = 0$. (We suppose L is normalized so that $|| e || = 1$.)

Under these conditions we may call L an abstract L_p space. (In the case $p = 1$, an abstract L_1 space is thus an abstract (L)-space in the sense of Kakutani [2].)

We seek to represent abstract L_p spaces as function spaces. We recall from [1], p. 248, that a norm on a Banach lattice is called uniformly monotone when, given $\varepsilon > 0$, one can find $\delta > 0$ so small that if $f \geq 0$, $g \geq 0$, $|| f || = 1$ and $|| f + g || - 1 \leq \delta$, then $|| g || \leq \varepsilon$. It follows at once from (3) that the norm on L is uniformly monotone. Thus, since L is complete, it is completely reticulated ([1], p. 249); i. e. every non-empty subset of L bounded from above has a least upper bound. Hence, by a theorem of Kakutani ([3], p. 998) in the form given by Stone ([4], p. 85), L is isomorphic as a Riesz space to a space of continuous functions on a compact Hausdorff space, if we entirely ignore nowhere dense sets. If we do not ignore these sets, we obtain a space of functions with a semi-norm, defined by the norm on L, which satisfies the hypotheses given at the beginning of this section. Thus we may now return to these hypotheses without loss of generality.

We now define a collection N of functions, which we call null functions, by $f \in N$ if there are $f_1, f_2, \cdots \in L$ such that:

(a) $f_n \geq | f |$ for all n

and (b) $|| f_n || \rightarrow 0$.

Clearly if $f \in N \cap L$, $|| f || = 0$. It is also clear that N is a lattice ideal in the set of all functions on E; i. e. N is a linear subspace of this set with the property that $| f | \leq | g |$ and $g \in N$ imply $f \in N$.

We define $L' \supset L$ by $f \in L'$ if there are $g \in L$, $h \in N$ such that $f = g + h$. Clearly L' is a linear space. Suppose $f = g_1 + h_1 = g_2 + h_2$ with $g_i \in L$, $h_i \in N$ $(i = 1, 2)$. Then $h_1 - h_2 = g_2 - g_1 \in L \cap N$. Thus $|| g_2 - g_1 || = 0$. Hence $|| g_2 || = || g_1 + g_2 - g_1 || \leq || g_1 || + || g_2 - g_1 || = || g_1 ||$. Similarly $|| g_1 || \leq || g_2 ||$. Hence $|| g_1 || = || g_2 ||$. It follows that we may define a semi-norm on L' by defining $\|g + h\|$ to be $\|g\|$, where $g \in L$ and $h \in N$.

We next show that L' is a lattice; i. e. that $f_1 \wedge f_2 \in L'$ whenever $f_1, f_2 \in L'$. Let $f_1 = g_1 + h_1$, $f_2 = g_2 + h_2$ with $g_i \in L$, $h_i \in N$. Then $g_1 \wedge g_2 \in L$. We have $f_1 \wedge f_2 = (g_1 + h_1) \wedge (g_2 + h_2) \leq (g_1 + h_1^+) \wedge (g_2 + h_2^+) \leq g_1 \wedge g_2 + h_1^+ + h_2^+$. Thus $f_1 \wedge f_2 - g_1 \wedge g_2 \leq h_1^+ + h_2^+$. Similarly $f_1 \wedge f_2 - g_1 \wedge g_2 \geq -h_1^- - h_2^-$. Since N is a lattice ideal, $f_1 \wedge f_2 - g_1 \wedge g_2 \in N$. Hence $f_1 \wedge f_2 \in L'$.

It is easy to check that L' satisfies all the hypotheses imposed above on L. In addition, L' has the following property:

If $f_1, f_2, \cdots \in L'$ are positive, $f_n \uparrow f$ pointwise and $\|f_n\| < \alpha$ for all n, then $f \in L'$ and $\|f - f_n\| \to 0$. To see this we note that $\{\|f_n\|\}$ is an increasing sequence of real numbers bounded from above by α; hence it is a Cauchy sequence. Thus $\{\|f_n\|^p\}$ is also a Cauchy sequence. Whenever $n \geq m$ we have $\|f_n - f_m\|^p \leq \|f_n - f_m + f_m\|^p - \|f_m\|^p = \|f_n\|^p - \|f_m\|^p$ by (3) above. Thus there is an $f' \in L'$ such that $\|f' - f_n\| \to 0$ by (1) above. Since $f_n \leq f$ for all n, $f' - f_n \geq f' - f$ for all n. Since $f' - f_n \in L'$, $f' - f_n = g_n + h_n$ with $g_n \in L$, $h_n \in N$. By the definiton of N, we can find, for each n, a $g_n' \in L$ such that $g_n' \geq h_n$ and $\|g_n'\| \leq 1/n$. Let $f_n' = g_n + g_n'$. Then $f_n' \geq g_n + h_n = f' - f_n \geq f' - f$. Also $\|f_n'\| \leq \|g_n\| + \|g_n'\| \leq \|f' - f_n\| + 1/n \to 0$. By the definition of N, $f' - f \in N$. Thus $f \in L'$. Also $\|f - f_n\| \leq \|f - f'\| + \|f' - f_n\| \to 0$.

At this point, we replace L by L'; i. e. we write L for L'.

LEMMA. *Let $f \in L$ be positive. Let g be the characteristic function of the set on which f differs from 0. Then $g \in L$.*

Proof. Clearly $nf \wedge 1 \uparrow g$ pointwise. Since $\|nf \wedge 1\| \leq \|1\| = 1$ for all n, $g \in L$ by what has just been proved.

LEMMA. *Let $f \in L$ be positive. Then there are positive $f_1, f_2, \cdots \in L$ such that $f_n \uparrow f$ pointwise, $\|f - f_n\| \to 0$, and each f_n assumes only finitely many values.*

Proof. For each positive integer n, let f_n be defined by: $f_n(x) = 2^{-n}[2^n(f \wedge n)(x)]$ for all $x \in E$. (By $[\alpha]$ we mean the largest integer $\leq \alpha$.) For each $x \in E$, $f_n(x) = 2^{-n}[2^n f(x)]$ for large n; thus clearly $f_n(x) \to f(x)$. Hence $f_n \to f$ pointwise. We note

$$f_{n+1}(x) = \frac{1}{2^{n+1}}[2^{n+1}(f \wedge (n+1))(x)] \geq \frac{1}{2^{n+1}}[2^{n+1}(f \wedge n)(x)]$$
$$\geq \frac{1}{2^n}[2^n(f \wedge n)(x)] = f_n(x)$$

for each $x \in E$. Hence $f_n \uparrow f$ pointwise. If we show $f_n \in L$ for all n, we shall know $\|f - f_n\| \to 0$ and the lemma will be proved.

Let n be fixed. Let $g_1, g_2, \cdots$ be functions on E defined by:

$$g_i(x) = 1 \text{ if } x \text{ is such that } 2^n(f \wedge n)(x) \geq i$$

$$g_i(x) = 0 \text{ if } x \text{ is such that } 2^n(f \wedge n)(x) < i\ .$$

We note that $f_n(x) = 2^{-n} \sum_{i=1}^{\infty} g_i(x)$. Since $2^n(f \wedge n)(x) \leq 2^n n$, $g_i(x) = 0$ for all x when $i \geq 2^n n$. Thus $f_n(x) = 2^{-n} \sum_{i=1}^{2^n n} g_i(x)$. Clearly each $1 - g_i$ is the characteristic function of the set on which $(2^n(f \wedge n) - i)^-$ differs from 0. Since $(2^n(f \wedge n) - i)^- \in L$, $1 - g_i \in L$ by the previous lemma; hence $g_i \in L$. We note $f_n = 2^{-n} \sum_{i=1}^{2^n n} g_i$ which shows $f_n \in L$ and completes the proof.

We now define a measure μ on the set E. Let A be a subset of E. If f_A, the characteristic function of A, is in L, we call A measurable and put $\mu(A) = \| f_A \|^p$. The verification that μ is a countably additive measure is trivial, making use of conditions (3) and (4) of our hypothesis, except for the following: Let $A_1, A_2, \cdots \subset E$ be pairwise disjoint and measurable. Let f_n be the characteristic function of $A_1 \cup \cdots \cup A_n$ $(n = 1, 2, \cdots)$. Then $f_n \uparrow f$ pointwise, where f is the characteristic function of $\bigcup_{n=1}^{\infty} A_n$. By what has been shown above, $f \in L$ and $\| f - f_n \| \to 0$. Thus $\mu(A_1) + \cdots + \mu(A_n) = \| f_n \|^p \to \| f \|^p = \mu(\bigcup_{n=1}^{\infty} A_n)$.

Next we consider the space L_p defined by μ. The functions in L which assume only finitely many values are precisely the measurable functions which assume only finitely many values. Clearly the given semi-norm on L coincides with the L_p norm for such functions. It follows, by considering pointwise limits of increasing sequences of such functions, that the functions in L_p are precisely those in L and that the norms agree. Remembering that we modified the original L by introducing null functions, we have the following theorem:

THEOREM. *Let L be a Riesz space of functions on a set E. Suppose there is a semi-norm on L which satisfies conditions* (1)-(5) *above. Then there is a countably additive measure μ on E such that L is essentially L_p with respect to μ; i. e. such that*:

(a) *For every $f \in L$, $\| f \|^p = \int | f |^p \, d\mu$.*

and (b) *If $f \geq 0$ and $\int f^p \, d\mu < \infty$, then there is a $g \in L$ such that $f(x) = g(x)$ for almost all $x \in E$.*

In the case $p = 2$, we can modify the hypotheses above. We suppose that H is a Riesz space of functions. We also suppose that there is a positive semi-definite bilinear form defined on H and that H is complete in the semi-norm determined by this form. We also assume that $\| f \| \leq \| g \|$ whenever $f, g \in H$ and $0 \leq f \leq g$. Next suppose that $\| f^+ \| \leq \| f \|$ for all $f \in H$. Finally we suppose $1 \in H$ and $\| 1 \| = 1$. We prove the following lemmas to show that H satisfies, with $p = 2$, the hypotheses given at the beginning of the paper.

LEMMA. *If $f, g \in H$ are positive, then $(f, g) \geq 0$.*

Proof. We note $f + \alpha g \geq f \geq 0$ for all $\alpha > 0$. Thus $\|f + \alpha g\| \geq \|f\|$. Hence we have $0 \leq \|f + \alpha g\|^2 - \|f\|^2 = 2\alpha(f, g) + \alpha^2 \|g\|^2$. It follows that $2(f, g) \geq -\alpha \|g\|^2$ for all $\alpha > 0$. Hence $(f, g) \geq 0$.

LEMMA. *If $f, g \in H$ are positive and $f \wedge g = 0$, then $(f, g) = 0$.*

Proof. We note $f \wedge (\alpha g) = 0$ for all $\alpha > 0$. Hence $(f - \alpha g)^+ = f$. Thus we have $\|f\|^2 = \|(f - \alpha g)^+\|^2 \leq \|f - \alpha g\|^2 = \|f\|^2 - 2\alpha(f, g) + \alpha^2 \|g\|$. Hence $\alpha \|g\|^2 \geq 2(f, g)$ for all $\alpha > 0$. Thus $(f, g) \leq 0$. By the previous lemma, $(f, g) \geq 0$. Therefore $(f, g) = 0$.

LEMMA. *$\|f\| = \| \, |f| \, \|$ for all $f \in H$.*

Proof. We have $\|f\|^2 = \|f^+ - f^-\|^2 = \|f^+\|^2 - 2(f^+, f^-) + \|f^-\|^2 = \|f^+ + f^-\|^2 - 4(f^+, f^-) = \| \, |f| \, \|^2 - 4(f^+, f^-)$. But $(f^+, f^-) = 0$ by the previous lemma.

LEMMA. *$\|f + g\|^2 \geq \|f\|^2 + \|g\|^2$ whenever $f, g \in H$ are positive.*

Proof. $\|f + g\|^2 = \|f\|^2 + 2(f, g) + \|g\|^2 \geq \|f\|^2 + \|g\|^2$ since $(f, g) \geq 0$.

LEMMA. *If $f, g \in H$ are positive and $f \wedge g = 0$, then $\|f + g\|^2 = \|f\|^2 + \|g\|^2$.*

Proof. We have $\|f + g\|^2 = \|f\|^2 + 2(f, g) + \|g\|^2 = \|f\|^2 + \|g\|^2$.

Thus we have verified that H satisfies the hypotheses for L with $p = 2$. On this basis we prove:

THEOREM. *Let H be as described above. Then there is a countably additive measure μ on E such that H is essentially L_2 with respect to μ; i. e. such that:*

(a) *For every $f, g \in H$, $(f, g) = \int fg \, d\mu$.*

and (b) *If $f \geq 0$ and $\int f^2 \, d\mu < \infty$, then there is a $g \in H$ such that $f(x) = g(x)$ for almost all $x \in E$.*

Proof. In addition to what has been proved above, it is enough to note that the inner product may be expressed in terms of the norm in the usual way.

BIBLIOGRAPHY

[1] G. Birkhoff, *Lattice theory*, Revised edition, Amer. Math. Soc. Colloquium Publications, Vol. XXV, Amer. Math. Soc., New York 1948.

[2] S. Kakutani, *Concrete representation of abstract* (L)*-spaces and the mean ergodic theorem*, Annals of Math. **42**, (1941) 523–537.

[3] ———, *Concrete representation of abstract* (M)*-spaces*, Annals of Math. **42**, (1941) 994–1024.

[4] M. H. Stone, *A general theory of spectra* II, Proc. Nat. Acad. Sci. **27**, (1941) 83–87.

HARVARD UNIVERSITY

SEPARABLE CONJUGATE SPACES

ROBERT C. JAMES

A Banach space B is reflexive if the natural isometric mapping of B into the second conjugate space B^{**} covers all of B^{**}. All conjugate spaces of a reflexive separable space B are separable. The nonreflexive space $l^{(1)}$ is separable and its first conjugate space is (m), which is nonseparable. The space (c_0) is separable, its first conjugate space is $l^{(1)}$, and its second conjugate space is (m). An example is known of a nonreflexive Banach space whose conjugate spaces are all separable [4]. This space is pseudo-reflexive in the sense that its natural image in the second conjugate space has a finite-dimensional complement. The structure of such spaces has been studied carefully [2].

The main purpose of this paper is to show that the sequence started by $l^{(1)}$ and (c_0) can be extended to give a sequence $\{B_n\}$ of separable Banach spaces such that, for each n, the nth conjugate space of B_n is its first nonseparable conjugate space. The principal tool used is a theorem which states a sufficient condition on a space T for the existence of a space B with

$$B^{**} = \pi(B) \dotplus T ,$$

where $\pi(B)$ is the natural image of B in B^{**}. The following definition and notation will be used.

A *basis* for a Banach space B is a sequence $\{u^i\}$ such that, for each x of B, there is a unique sequence of numbers $\{a_i\}$ for which $\lim_{n\to\infty}\|x - \sum_1^n a_i u_i\| = 0$. A sequence $\{u_i\}$ is a basis for its closed linear span if and only if there is a number $\varepsilon > 0$ such that

$$\left\|\sum_1^{n+p} c_i x_i\right\| \geq \varepsilon \left\|\sum_1^{n} c_i x_i\right\|$$

for any numbers $\{c_i\}$ and positive integers n and p [1, page 111]. If ε can be $+1$, the basis is an *orthogonal basis*. It will be useful to classify bases as follows:

Type α. If $\{a_i\}$ is a sequence of numbers for which $\sup_n \|\sum_1^n a_i u_i\| < \infty$, then $\sum_1^\infty a_i u_i$ converges.

Type β. If f is a linear functional defined on B and $\|f\|_n$ is the norm of f on the closed linear span of $\{u_i \mid i \geq n\}$, then $\lim_{n\to\infty}\|f\|_n = 0$.

There are Banach spaces which have bases which are neither of type α nor of type β, while a basis is of both types if and only if the space

Received April 28, 1959.

is reflexive [3; Theorem 1].

The symbols C, (m), $l^{(1)}$, and (c_0) are used in the usual sense [1; pages 11, 12, 181]. The set of all $r + t$ with $r \in R$ and $t \in T$ is denoted by $R \dotplus T$. A space R is said to be *embedded* in a space S if R is mapped isomorphically and isometrically on a subspace of S; for $x \in R$, the image of x is indicated by $x^{(S)}$. In particular, $x^{(C)}$ is a continuous function defined on $[0, 1]$ and the value of $x^{(C)}$ at t is denoted by $x^{(C)}(t)$. If $w = (w_1, w_2, \cdots)$ is a sequence of numbers, then ${}^n w$ is the sequence obtained by replacing w_i by 0 if $i > n$. A *block* of w is a sequence ${}^n_m w$ obtained from w by replacing w_i by 0 if $i \leq m$ or $i > n$. Two blocks ${}^{n_1}_{m_1} w$ and ${}^{n_2}_{m_2} w$ are said to *overlap* if the intervals $(m_1, n_1]$ and $(m_2, n_2]$ overlap.

LEMMA 1. *Let T be a Banach space with an orthogonal basis $\{u_i\}$. Then T can be embedded in (m) in such a way that:*

(i) *if $x = \sum_1^\infty a_i u_i$, then the first $2N$ coordinates of $x^{(m)}$ are zero if and only if $a_i = 0$ for $i \leq N$;*

(ii) *if $\{a_i\}$ and $\{x_i^m\}$ are related by $x = \sum_1^\infty a_i u_i$ and $x^{(m)} = (x_1^m, x_2^m, \cdots)$, then $a_1, \cdots, a_N$ are each continuous functions of $x_1^m, \cdots, x_{2N}^m$ and $x_1^m, \cdots, x_{2N}^m$ are each continuous functions of $a_1, \cdots, a_N$;*

(iii) *if $x^{(m)} = (x_1^m, x_2^m, \cdots)$, then $\| x^{(m)} \| = \limsup | x_i^m |$.*

Proof. Let T be embedded in the space C. Let $\{t_i\}$ be a sequence of numbers in the interval $[0, 1]$ for which the sequence $\{t_{2i-1}\}$, $i = 1, 2, \cdots$, is dense in $[0, 1]$ and, for each i, $u_i^{(C)}(t_{2i}) \neq 0$. If $x = \sum_1^\infty a_i u_i$, let $x^{(m)}$ be the sequence $(x_1^m, x_2^m, \cdots)$ for which

$$x_{2k-1}^m = \sum_1^k a_i u_i^{(C)}(t_{2k-1}),\ x_{2k}^m = \sum_1^k a_i u_i^{(C)}(t_{2k}) .$$

Then for any $t \in [0, 1]$,

$$\left| \sum_1^k a_i u_i^{(C)}(t) \right| \leq \left\| \sum_1^k a_i u_i^{(C)} \right\| = \left\| \sum_1^k a_i u_i \right\| \leq \left\| x \right\| .$$

Hence $\| x^{(m)} \| \leq \| x \|$. But if $\varepsilon > 0$ and N is chosen so that $\| x - \sum_1^k a u_i \| < \varepsilon$ if $k > N$, then it follows from $\{t_{2k-1}\}$ being dense in $[0, 1]$ that

$$\| x^{(m)} \| \geq \sup_{k > N} \left| \sum_1^k a_i u_i^{(C)}(t_{2k-1}) \right| \geq \left\| x \right\| - \varepsilon .$$

Hence $\| x \| = \| x^{(m)} \|$ and T and its image in (m) are isometric. But if $x = \sum_{N+1}^\infty a_i u_i$, then $x_{2k-1}^m = x_{2k}^m = 0$ if $k \leq N$. If $x_i^m = 0$ for $i \leq 2N$, then the equations $x_{2k}^m = \sum_1^k a_i u_i{}^{(C)}(t_{2k}) = 0$, $k \leq N$, successively imply $0 = a_1 = a_2 = \cdots = a_N$, since $u_k^{(C)}(t_{2k}) \neq 0$. The conclusion (ii) follows from this system of equations and the continuity of $\sum_1^N a_i u_i$ in $a_1, \cdots, a_N$, while (iii) follows from $\{t_{2k-1}\}$ being dense in $[0, 1]$.

LEMMA 2. *Let T be a Banach space with an orthogonal basis $\{u_i\}$ and let T be embedded in (m) as described in Lemma 1. Then the following are equivalent:*

(i) *the basis $\{u_i\}$ is of type α;*

(ii) *if $w \in (m)$, then $w = v + t$, with v an element of (m) which has all coordinates zero after the Mth ($M \geq 0$) and t the image of an element of T, provided there is a sequence of elements $\{y_k\}$ of T for which* $\sup \| y_k \| < \infty$ *and*

$$\lim_{k\to\infty} y^m_{k,i} = w_i \text{ for } i > M ,$$

where $w = (w_1, w_2, \cdots)$ *and* $y_k^{(m)} = (y^m_{k,1}, y^m_{k,2}, \cdots)$.

Proof. Assume the basis $\{u_i\}$ is of type α and let $w = (w_i, w_2, \cdots)$ and $\{y_k\}$ satisfy the hypotheses of (ii). Since $\| y_k \|$ is bounded, there is a subsequence $\{z_k\}$ of $\{y_k\}$ such that

$$\lim_{k\to\infty} z^m_{k,i} = v_i$$

exists for $i \leq M$. Let $v = (w_1 - v_1, \cdots, w_M - v_M, 0, 0, \cdots)$. Also let $z_k = \sum_1^\infty a_i^k u_i$ for each k. It now follows from (ii) of Lemma 1 that $\lim_{k\to\infty} a_i^k = a_i$ exists for each i. Since the basis is orthogonal, $\| \sum_1^n a_i u_i \| \leq \sup \| z_k \|$. Since $\{u_i\}$ is a basis of type α, it then follows that $\sum_1^\infty a_i u_i$ is convergent. Also, $w - v = t$ is the (m)-image of $\sum_1^\infty a_i u_i$. This follows from the fact that the numbers a_i, $i \leq N$, continuously determine the first $2N$ coordinates of the (m)-image of $\sum_1^\infty a_i u_i$, while $z_k = \sum_1^\infty a_i^k u_i$, $\lim_{k\to\infty} a_i^k = a_i$, and $\lim_{k\to\infty} z^m_{k,i}$ exists and is the ith coordinate of $w - v$.

Now assume (ii) and let $\| \sum_1^n a_i u_i \|$ be a bounded function of n. Let $w = (w_1, w_2, \cdots)$ be the element of (m) whose first $2N$ coordinates are determined by $a_1, \cdots, a_N$. Take $M = 0$ and y_k to be the (m)-image of $\sum_1^k a_i u_i$. It then follows from (ii) that w is the (m)-image of some element of T, which can only be $\sum_1^\infty a_i u_i$.

THEOREM 1. *Let T be a Banach space which has an orthogonal basis of type α. Then there is a Banach space B which has a basis of type β and for which*

$$B^{**} = \pi(B) \dotplus T_1 ,$$

where $\pi(B)$ is the natural image of B in B^{**}, T and T_1 are isometric, and $\| r + t \| \geq \| t \|$ if $r \in \pi(B)$ and $t \in T_1$.

Proof. Let T_1 be the embedding of T in (m) as described in Lemma 1. The norm of (m) will be denoted by $\| \;\; \|$. For elements w of (m) which have only a finite number of nonzero coordinates, let

(1) $\theta(w) = \inf \| t \|$ for w a block of t, where t is either a member

of T_1 or has only one nonzero coordinate (note that $\theta(w)$ is defined only for elements w which are blocks of at least one $t \in T_1$ or which have only one nonzero coordinate);

(2) $h(w) = \{\inf \sum [\theta(b_i)]^2\}^{1/2}$, where $w = \sum b_i$, each b_i is a block of w, and no two blocks overlap.

(3) $|||x||| = \inf \sum h(w_j)$ for $x = \sum w_j$.

In the above, all sums have a finite number of terms. The triangular inequality for $|||\ |||$ is a direct consequence of (3). Also, $|||x||| \geq ||x||$, since $\theta(w) \geq ||w||$ and $h(w) \geq ||w||$. Let B be the completion of the space of sequences with a finite number of nonzero coordinates, using the norm $|||\ |||$. The sequence of elements $\{u_i\}$ for which u_i has all coordinates 0 except the ith, which is 1, is an orthogonal basis for B. This means that $|||\sum_1^{n+p} a_i u_i||| \geq |||\sum_1^n a_i u_i|||$, which follows by noting that, if $\sum_1^{n+p} a_i u_i = \sum w_j$, then $\sum_1^n a_i u_i = \sum {}^n w_j$ and $h({}^n w_j) \leq h(w_j)$ for each j, where ${}^n w_j$ is obtained from w_j by replacing each coordinate after the nth by 0.

The basis $\{u_i\}$ is of type β. For suppose there is a linear functional f for which $\lim_{n\to\infty} |||f|||_n = K \neq 0$ and choose N so that $|||f|||_N \leq 7/6K$. Then there are two elements $x = \sum_{n_1}^{n_2} a_i u_i$, $y = \sum_{n_3}^{n_4} a_i u_i$, for which $N < n_1 \leq n_2 < n_3 \leq n_4$, $|||x||| = |||y||| = 1$, $f(x) > 7/8K$ and $f(y) > 7/8K$. Then

$$\frac{7}{4}K < f(x) + f(y) \leq \left(\frac{7}{6}K\right)|||x + y||| \text{ and } |||x + y||| > \frac{3}{2}\,.$$

Since θ and h are both monotone decreasing as a block has coordinates at the ends replaced by zeros, there exists $\{x_j\}$ and $\{y_j\}$ such that $x = \sum x_j$, $y = \sum y_j$, $\sum h(x_j) < |||x||| + \varepsilon$, and $\sum h(y_j) < |||y||| + \varepsilon$, where each x_j has zero coordinates outside the index interval $[n_1, n_2]$ and each y_j has zero coordinates outside the index interval $[n_3, n_4]$. Now replace the sets $\{x_j\}$ and $\{y_j\}$ by $\{\bar{x}_j\}$ and $\{\bar{y}_j\}$ defined as follows: if $h(x_p)$ is the smallest of all the numbers $h(x_j)$ and $h(y_j)$, then let $\bar{x}_1 = x_p$ and $\bar{y}_1 = [h(x_p)/h(y_r)]y_r$ (for some r) and replace y_r by $[1 - h(x_p)/h(y_r)]y_r$. The analogous process is used if h takes on its minimum at one of the y_j's. This process creates two new elements and eliminates one old one at each step, until all of the x_j's or all of the y_j's are eliminated. If only x_j's remain, say x_{p_j}'s, then $\sum h(x_{p_j}) < \varepsilon$, and similarly $\sum h(y_{p_j}) < \varepsilon$ if only y_j's remain. Also

$$\sum h(\bar{x}_j) - \varepsilon = \sum h(\bar{y}_j) - \varepsilon < |||x||| = |||y||| = 1$$

and $h(\bar{x}_j) = h(\bar{y}_j)$ for each j. For each j, there are nonoverlapping blocks $\{\bar{x}_{ji}\}$ and $\{\bar{y}_{ji}\}$ such that

$$h(\bar{x}_j) = h(\bar{y}_j) = \{\textstyle\sum_i [\theta(\bar{x}_{ji})]^2\}^{1/2} = \{\textstyle\sum_i [\theta(\bar{y}_{ji})]^2\}^{1/2}\,.$$

Then

$$h(\bar{x}_j + \bar{y}_j) \leq \{\sum_i [\theta \bar{x}_{ji})]^2 + \sum_i [\theta(\bar{y}_{ji})]^2\}^{1/2} = \sqrt{2}\, h(\bar{x}_j) .$$

Hence

$$|||x + y||| \leq \sum h(\bar{x}_j + \bar{y}_j) + \varepsilon \leq \sqrt{2} \sum h(\bar{x}_j) + \varepsilon \leq \sqrt{2} + \varepsilon .$$

Since $|||x + y||| > 3/2$, this is contradictory if $\sqrt{2} + \varepsilon < 3/2$. It has therefore been shown that $\{u_i\}$ is a basis of type β.

Since $\{u_i\}$ is an orthogonal basis of type β for B, it follows that B^{**} consists of all sequences $F = (F_1, F_2, \cdots)$ for which

$$|||F||| = \lim_{n\to\infty} |||(F_1, \cdots, F_n, 0, 0, \cdots)|||$$

exists [4; page 174]. Note first that if $t = (t_1, \cdots) \in T_1$, then

$$|||(t_1, \cdots, t_n, 0, 0, \cdots)||| = ||(t_1, \cdots, t_n, 0, 0, \cdots)||$$

and $\lim_{n\to\infty} |||(t_1, \cdots, t_n, 0, 0, \cdots)||| = |||t||| = ||t||$. Thus $T_1 \subset B^{**}$. Also, the natural mapping of B into B^{**} is merely the mapping of a sequence in B onto the identical sequence in B^{**}. It then follows that $|||r + t||| \geq |||t|||$ if $r \in \pi(B)$ and $t \in T_1$, since r can be approximated by a sequence with a finite number of nonzero coordinates but (Lemma 1) $||t|| = \limsup |t_i|$.

Now suppose that $F = (F_1, F_2, \cdots)$ is a sequence for which $\lim_{n\to\infty} |||{}^nF|||$ exists; i.e., $F \in B^{**}$. It will be shown that there is an element v of $\pi(B) \dotplus T_1$ for which $|||F - v||| \leq 15/16\, |||F|||$. Successive application of this would then establish that $F \in \pi(B) \dotplus T_1$. For each n, there are ${}^nw_j$ and blocks $b^n_{j,i}$, which are either blocks of elements of T_1 or have only one nonzero coordinate, such that

$$|||{}^nF||| = \sum_j h({}^nw_j),\ {}^nF = \sum_j {}^nw_j, \text{ and } h({}^nw_j) = \{\sum_i [\theta(b^n_{j,i})]^2\}^{1/2} ,$$

where each ${}^nw_j$ and each $b^n_{j,i}$ have all coordinates zero after the nth. This follows by a limit argument, using the facts (1) that there are only a finite number K_n of ways of choosing division points for nonoverlapping blocks from the integers $1, 2, \cdots, n$ and (2) that it follows from Lemma 1 and the orthogonality of the basis for T that $\theta(b^{2N}_{j,i})$, for a block $b^{2N}_{j,i}$ which has zero coordinates beyond the $2N$th coordinate, can be evaluated by using only members of the span of the first N basis elements of T.

If $m < n$ and ${}^mw^n_j$ is obtained from ${}^nw_j$ by replacing coordinates after the mth by zeros, then

$$|||{}^mF||| \leq \sum_j h({}^mw^n_j) \leq |||{}^nF||| \leq |||F||| .$$

If ${}^mw^n_{j_1}$ and ${}^mw^n_{j_2}$ are of the "same type" in the sense that they are divided into blocks by using the same division points, then it follows by using these same division points for ${}^mw^n_{j_1} + {}^mw^n_{j_2}$ that

$$h({}^m w^n_{j_1} + {}^m w^n_{j_2}) \leq h({}^m w^n_{j_1}) + h({}^m w^n_{j_2}) .$$

For each $n > m$, let ${}^m\hat{w}^n_j$ be the sum of all ${}^m w^n_{j_i}$ of the "same type" as ${}^m\hat{w}^n_j$. A limit argument gives a sequence of integers $\{n_i\}$ such that $\lim {}^m\hat{w}^{n_i}_j = {}^m\overline{w}_j$ exists for each "type". If $m < n$, then there exist $\bar{b}^n_{j,i}$ such that

$$|||\,{}^mF\,||| \leq \textstyle\sum_j h({}^m\overline{w}_j) \leq \sum_k h({}^n\overline{w}_k) \leq |||\,F\,||| ,$$
$$h({}^m\overline{w}_j) = \{\textstyle\sum_i [\theta(\bar{b}^m_{j,i})]^2\}^{1/2},\ {}^mF = \sum {}^m\overline{w}_j ,$$

and ${}^m\overline{w}_j$ is equal to the sum of all ${}^m\overline{w}^n_j$ which are of the same type as ${}^m\overline{w}_j$ and are obtained from ${}^n\overline{w}_j$ by replacing all coordinates after the mth by zeros. The points used to divide ${}^m\overline{w}_j$ into the blocks $\bar{b}^m_{j,i}$ will be called the *division points* of ${}^m\overline{w}_j$.

Choose M so that $|||\,{}^mF\,||| > 15/16\,|||\,F\,|||$. Note that if ${}^m\overline{w}_j$ is of a particular type and $n > m$, then ${}^m\overline{w}_j$ is the sum of one or more elements obtained from the ${}^n\overline{w}_k$'s by replacing coordinates after the mth by zeros. For $n > m \geq M$, let ${}^nt$ be the sum of all ${}^n\overline{w}_k$'s which have no division points between M and n and let ${}^mt^n$ be obtained from ${}^nt$ by replacing coordinates after the mth by zeros. Let $\{n_i\}$ be chosen so that

$$\lim_{i\to\infty} {}^mt^{n_i} = {}^m\bar{t}$$

exists for each $m \geq M$. Let $\bar{t}$ be defined so as to have the same first m coordinates as ${}^m\bar{t}$. Then any finite block of $\bar{t}$ whose first M coordinates are zero is also approximately a block of an element of T_1 and these elements of T_1 are of bounded norm. It then follows from Lemma 2 that there is an element v_0, with a finite number of nonzero coordinates, such that $v_0 + \bar{t} \in T_1$. Thus

$$\bar{t} \in \pi(B) \dotplus T_1 .$$

First assume that $|||\,\bar{t}\,||| > 1/8\,|||\,F\,|||$ and choose N so that

$$|||\,{}^n\bar{t}\,||| > 1/8\,|||\,F\,||| \text{ if } n > N .$$

For $n > N$, choose $p > n$ so that

$$|||\,{}^n\bar{t} - {}^nt^p\,||| < \frac{1}{32}\,|||\,F\,||| .$$

Since $|||\,{}^nF\,||| \leq \sum_j h({}^n\overline{w}_j)$, discarding all ${}^n\overline{w}^p_j$ without division points between M and p gives

$$\begin{aligned}|||\,{}^nF - {}^nt^p\,||| &\leq \textstyle\sum h({}^n\overline{w}_j) - |||\,{}^nt^p\,||| \\ &\leq |||\,F\,||| - |||\,{}^nt^p\,||| .\end{aligned}$$

Hence $||| {}^nF - {}^n\bar{t} ||| < ||| F ||| - ||| {}^n\bar{t} ||| + 1/16 ||| F ||| < 15/16 ||| F |||$. Since n was an arbitrary integer with $n > N$, it follows that

$$||| F - \bar{t} ||| \leq \frac{15}{16} ||| F ||| .$$

Now assume that $||| \bar{t} ||| \leq 1/8 ||| F |||$. Then $||| {}^n\bar{t} ||| \leq 1/8 ||| F |||$ for all n. Choose q so that

$$||| {}^M\bar{t} - {}^Mt^q ||| < \frac{1}{16} ||| F ||| .$$

For each ${}^q\bar{w}_j$ which has a division point between M and q, let u_j^q be obtained from ${}^q\bar{w}_j$ by replacing all coordinates after the last such division point by zeros. Let

$$u = \sum_j u_j^q .$$

Choose $n > q$. Then ${}^nF = \sum {}^n\bar{w}_j$ and

$$\begin{aligned} ||| {}^MF ||| &\leq \sum h({}^M\bar{w}_j^n) \leq \sum h(u_j^q) + ||| {}^Mt^q ||| \\ &< \sum h(u_j^q) + \frac{3}{16} ||| F ||| . \end{aligned}$$

Since $||| {}^MF ||| > 15/16 ||| F |||$, we have $\sum h(u_j^q) > 3/4 ||| F |||$. Now consider $F - u$. Since $||| {}^nF ||| \leq \sum h({}^n\bar{w}_j)$, where $h({}^n\bar{w}_j) = \{\sum_i [\theta(b_{j,i}^n)]^2\}^{1/2}$, we have

$$\begin{gathered} {}^n(F - u) = \sum {}^n\bar{w}_j - \sum u_j^q = \sum {}^n\tilde{w}_j , \\ ||| {}^n(F - u) ||| \leq \sum h({}^n\tilde{w}_j) , \end{gathered}$$

where ${}^n\tilde{w}_j$ is obtained from ${}^n\bar{w}_j$ by replacing all coordinates before the last division point between M and q by zeros (if there is no such point, then ${}^n\tilde{w}_j = {}^n\bar{w}_j$). The following trivial facts will be used: If A and B are nonnegative and

$$\text{if } \sqrt{3}A < B, \text{ then } \sqrt{A^2 + B^2} > 2A ;$$

$$\text{if } \sqrt{3}A \geq B, \text{ then } B < \sqrt{A^2 + B^2} - \frac{1}{4}A .$$

Each ${}^n\bar{w}_j$ which has a division point between M and q makes a contribution to some u_j^q. For such an ${}^n\bar{w}_j$, let

$$h({}^n\bar{w}_j) = [\sum_r (A_r)^2 + \sum_s (B_s)^2]^{1/2} ,$$

where the A_r's and B_r's are, respectively, the values of $\theta(\bar{b}_{j,i}^n)$ for $\bar{b}_{j,i}^n$ a block of some u_j^q and $\bar{b}_{j,i}^n$ *not* a block of any u_j^q. Then

$$h(u_j^q) \leq \sum [\sum_r (A_r)^2]^{1/2} ,$$

where the sum is over all ${}^n\overline{w}_j$ which make a contribution to u_j^q. Let $\sum_r (A_r)^2$ be of class (1) or of class (2) according as

$$\sqrt{3}\, [\sum (A_r)^2]^{1/2} < [\sum (B_s)^2]^{1/2} \text{ or } \sqrt{3}\, [\sum (A_r)^2]^{1/2} \geq [\sum (B_s)^2]^{1/2} .$$

Since $\sum h(u_j^q) > 3/4 \,|||\, F \,|||$, the sum of all terms of class (1) is not larger than $1/2 \,|||\, F \,|||$ (otherwise we would have $\sum h({}^n\overline{w}_j) > |||\, F \,|||$) and the sum of all terms of class (2) is greater than $1/4 \,|||\, F \,|||$. But for a term of class (2),

$$[\sum (B_s)^2]^{1/2} < h({}^n\overline{w}_j) - \frac{1}{4} [\sum (A_r)^2]^{1/2} .$$

Adding these inequalities for each ${}^n\overline{w}_j$ and discarding each $\sum (A_r)^2$ which is of class (1) gives

$$\sum h({}^n\tilde{w}_j) < \sum h({}^n\overline{w}_j) - \frac{1}{16} \,|||\, F \,|||\, \text{ and } \,|||\, {}^n(F - u) \,|||\, < \frac{15}{16} \,|||\, F \,|||\, .$$

Since n was an arbitrary integer with $n > q$, it follows that

$$|||\, F - u \,|||\, \leq \frac{15}{16} \,|||\, F \,|||\, .$$

The importance of the assumption in Theorem 1 that T_1 have a basis of type α is made clear by the fact that the theorem breaks down if T_1 has a subspace isomorphic with (c_0). In fact, in this case there can not be a separable space B with

$$B^{**} = \pi(B) \dotplus T_1$$

and T_1 separable, whether or not B and T_1 have bases. This follows from the fact that if a conjugate space R^* contains a subspace isomorphic with (c_0), then R^* contains a subspace isomorphic with (m) and is not separable. To establish this fact, suppose that $\{F_n\}$ are continuous linear functionals defined on some Banach space B and that the closed linear span of $\{F_n\}$ is isomorphic with (c_0), the correspondence being

$$\sum_1^\infty a_i F_i \leftrightarrow (a_1, a_2, \cdots) .$$

For any bounded sequence $w = (w_1, w_2, \cdots)$, define F_w by

$$F_w(f) = \lim_{n\to\infty} \Big(\sum_1^n w_i F_i\Big)(f) ,$$

for each f of B. This limit exists, since if it did not there would exist

$\varepsilon > 0$ and $G_1 = \sum_1^{n_1} w_i F_i$, $G_2 = \sum_{n_2}^{n_3} w_i F_i, \cdots$, with $1 \le n_1 < n_2 \le n_3 < n_4 \le \cdots$, such that $G_i(f) > \varepsilon$. Then correct choice of signs would give

$$\sum_1^n \pm G_i(f) > n\varepsilon ,$$

which contradicts the boundedness of $\| \sum^n \pm G_i \|$. Clearly the correspondence with (c_0) is thus extended to a bicontinuous correspondence with (m).

THEOREM 2. *For any positive integer n, there is a Banach space B_n such that the nth conjugate space of B_n is the first nonseparable conjugate space of B_n.*

Proof. Let $B_1 = l^{(1)}$ and $B_2 = (c_0)$. Then B_1 has a basis of type α and B_2 has a basis of type β. In the following, the notation $R \dotplus S$ is used only if $\| r + s \| \ge \| s \|$ whenever $r \in R$ and $s \in S$. It follows from Theorem 1 that there is a separable Banach space B_3 with a basis of type β for which

$$B_3^{**} = B_3 \dotplus l^{(1)} = B_3 \dotplus B_2^*$$

Then B_3^{***} is nonseparable and B_3^* has a basis of type α [3, Theorem 3]. Now suppose that, for $k \le n$, B_k has been found for which

$$B_k^{**} = B_k \dotplus B_{k-1}^*$$

if $k \ge 3$, B_k has a basis of type β if $k \ge 2$, and the kth conjugate space of B_k is the first nonseparable conjugate space of B_k. Then B_n^* has a basis of type α and it follows from Theorem 1 that there exists a separable space B_{n+1} which has a basis of type β and for which

$$B_{n+1}^{**} = B_{n+1} \dotplus B_n^* .$$

Then $B_{n+1}^{***} = B_{n+1}^* \dotplus B_n \dotplus B_{n-1}^*$. The $(n-2)$nd conjugate space of B_{n-1}^* is the first nonseparable conjugate space of B_{n-1}^*, while the $(n-2)$nd conjugate space of B_n is separable. Hence the $(n+1)$st conjugate space of B_{n+1} is the first nonseparable conjugate space of B_{n+1}.

REFERENCES

1. S. Banach, *Théorie des opérations linéaires*, Warsaw, 1932.
2. P. Civin and B. Yood, *Quasi-reflexive spaces*, Proc. Amer. Math. Soc., **8** (1957), 906–911.
3. R. C. James, *Bases and reflexivity of Banach spaces*, Annals of Math., **52** (1950), 518–527.
4. ———, *A non-reflexive Banach space isometric with its second conjugate space*, Proc. Nat. Acad. Sci. U.S.A., **37** (1951), 174–177.

HARVEY MUDD COLLEGE

ON NON-ASSOCIATIVE ALGEBRAS ASSOCIATED WITH BILINEAR FORMS

W. E. JENNER

If $\mathfrak{V}_0$ is a vector space over a field k, then with any non-degenerate bilinear form f_0 on $\mathfrak{V}_0 \times \mathfrak{V}_0$ is associated the group $\mathfrak{G}$ of linear transformations of $\mathfrak{V}_0$ which keep f_0 invariant. In this paper a procedure is given for associating with such a bilinear form an algebra $\mathfrak{A}$, non-associative in general, whose automorphism group is isomorphic to $\mathfrak{G}$ and which is right and left simple provided $\mathfrak{V}_0$ has dimension at least 2. In case k is the field of real numbers, then $\mathfrak{G}$ is a Lie group and its Lie algebra is the Lie algebra of derivations of $\mathfrak{A}$. In case the form f_0 is degenerate, and either symmetric or alternating, then the analogue of the Wedderburn Principal Theorem holds for $\mathfrak{A}$. The results obtained apply, in particular, to the orthogonal and symplectic groups.

Let $\mathfrak{V}_0$ be a vector space of dimension n over a field k with basis $u_1, \cdots, u_n$. It is assumed that $\lambda v = v\lambda$ for all $v \in \mathfrak{V}_0$ and $\lambda \in k$. Suppose f_0 is a bilinear form on $\mathfrak{V}_0 \times \mathfrak{V}_0$. Define $\mathfrak{A}$ to be the algebra over k with basis $e_0, e_1, \cdots, e_n$ and multiplication table $e_0^2 = e_0$, $e_i \cdot e_0 = e_0 \cdot e_i = e_i$, $e_i \cdot e_j = f(e_i\, e_j)e_0$ for $i, j = 1, 2, \cdots, n$, where $f(e_i, e_j) = f_0(u_i, u_j)$. Let $\mathfrak{V}$ be the subspace of $\mathfrak{A}$ spanned be $e_1, \cdots, e_n$. Then f is a bilinear form on $\mathfrak{V} \times \mathfrak{V}$.

THEOREM 1. *Suppose that f is non-degenerate and that $n \geqq 2$. Then $\mathfrak{A}$ is right and left simple.*

Proof. Let $\mathfrak{U}$ be a non-zero left ideal of $\mathfrak{A}$ and let u be a non-zero element of $\mathfrak{U}$. Suppose first that $u \in \mathfrak{V}$. Then there exists an element $v \in V$ such that $f(v, u) \neq 0$. Then $v \cdot u = f(v, u)e_0$. Therefore $e_0 \in \mathfrak{U}$ and so $\mathfrak{U} = \mathfrak{A}$. Next suppose $u = \alpha e_0 + v$ where $\alpha \neq 0$ in k and $v \in V$. Then one can assume $\alpha = 1$. Since $n \geqq 2$ it follows that $e_1 \cdot u = e_1 + \lambda_1 e_0$ and $e_2 \cdot u = e_2 + \lambda_2 e_0$ where $\lambda_1, \lambda_2 \in k$. If $\lambda_1 = 0$ then $e_1 \in U$ and the first part of the proof applies; similarly if $\lambda_2 = 0$. Consequently one can suppose $\lambda_1\lambda_2 \neq 0$. Then $\lambda_2 e_1 u - \lambda_1 e_2 u = \lambda_2 e_1 - \lambda_1 e_2$ is a non-zero element in $\mathfrak{U} \cap \mathfrak{V}$. Thus the first part of the proof again applies and so $\mathfrak{U} = \mathfrak{A}$. Therefore $\mathfrak{A}$ is left simple; similarly $\mathfrak{A}$ is right simple.

If $\mathfrak{A}$ is any (non-associative) algebra over k then left (right) multiplication by an element $a \in \mathfrak{A}$ determines a linear transformation $L_a(R_a)$ of the underlying vector space of $\mathfrak{A}$ by $a \cdot u = L_a u (u \cdot a = R_a u)$, $u \in \mathfrak{A}$. The set of linear transformations L_a (R_a) for $a \in \mathfrak{A}$ generate an associative algebra $L(\mathfrak{A})$ $(R(\mathfrak{A}))$ over k. The algebras $L(\mathfrak{A})$ and $R(\mathfrak{A})$ together

Received May 8, 1959.

generate the transformation algebra $T(\mathfrak{A})$.

THEOREM 2. *If f is non-degenerate and $n \geqq 2$ then $L(\mathfrak{A}) = R(\mathfrak{A}) = T(\mathfrak{A}) = [k]_{n+1}$.*

Proof. The proof of Theorem 1 shows that for any $u \neq 0$ in $\mathfrak{A}$ there is an element of $L(\mathfrak{A})$ mapping u into any arbitrarily assigned element of $\mathfrak{A}$. Therefore $L(\mathfrak{A}) = [k]_{n+1}$; similarly for $R(\mathfrak{A})$, and so also for $T(\mathfrak{A})$.

Albert has introduced in [1] the concept of isotopy of non-associative algebras. Suppose $\mathfrak{A}$ is an algebra with left multiplications L_a defined by $a \cdot u = L_a u$. Then an isotope of $\mathfrak{A}$ is an algebra $\mathfrak{A}^0$ with the same underlying vector space and multiplication defined by $a \circ u = PL_{aQ}Su$ where P, Q, S are invertible linear transformations of the underlying vector space of $\mathfrak{A}$. An algebra $\mathfrak{A}$ is said to be isotopically left (right) simple if every isotope of $\mathfrak{A}$ is left (right) simple.

THEOREM 3. *Suppose f is non-degenerate and that $n \geqq 2$. Then $\mathfrak{A}$ is isotopically left and right simple.*

Proof. Suppose $\mathfrak{U}$ is a subspace of $\mathfrak{A}$ such that $PL_{xQ}S\mathfrak{U} \subseteqq \mathfrak{U}$ for all $x \in \mathfrak{A}$. Now choose $x \in \mathfrak{A}$ such that $L_{xQ} = L_{e_0} = I$, the identity transformation. Then $PS\mathfrak{U} \subseteqq \mathfrak{U}$. Therefore $PS\mathfrak{U} = \mathfrak{U}$ and $S\mathfrak{U} = P^{-1}\mathfrak{U}$ since P and S are invertible. Then for any $u \in \mathfrak{A}$, $L_{uQ}S\mathfrak{U} \subseteqq P^{-1}\mathfrak{U} = S\mathfrak{U}$ and so $S\mathfrak{U}$ is a left ideal of $\mathfrak{A}$. Therefore either $\mathfrak{U} = (0)$ or $\mathfrak{A}$. Consequently $\mathfrak{A}$ is isotopically left simple; similarly it is isotopically right simple.

REMARK. Bruck has shown in [2] that left and right isotopic simplicity follow from left and right simplicity if the algebra has a unit element. The proof has been given here for sake of completeness.

THEOREM 4. *Suppose that f is non-degenerate and that $n \geq 2$. Let $\mathfrak{G}$ be the group of linear transformations of $\mathfrak{B}$ which keep f invariant. Then the group of automorphisms of $\mathfrak{A}$ is isomorphic to $\mathfrak{G}$. In case k is the field of real numbers the Lie group $\mathfrak{G}$ has for its Lie algebra the Lie algebra of derivations of $\mathfrak{A}$.*

Proof. Let φ be an automorphism of $\mathfrak{A}$. It is understood that φ is a k-automorphism so that φ keeps scalar multiples of e_0 fixed. Suppose $\varphi e_i = \lambda_i e_0 + v_i$ where $\lambda_i \in k$, $v_i \in \mathfrak{B}$ and $i = 1, 2, \cdots, n$. Then each product $\varphi e_i \cdot \varphi e_j = \mu_{ij} e_0 + \lambda_i v_j, + \lambda_j v_i$, $\mu_{ij} \in k$, must be a scalar multiple of e_0. Therefore $\lambda_i v_j + \lambda_j v_i = 0$ and so $\varphi(\lambda_j e_i + \lambda_i e_j - 2\lambda_i\lambda_j e_0) = 0$, which implies that $\lambda_i = \lambda_j = 0$ if $i \neq j$. Therefore $\varphi\mathfrak{B} \subseteq \mathfrak{B}$. Then $\varphi e_i \cdot \varphi e_j = f(\varphi e_i, \varphi e_j)e_0 = \varphi(e_i \cdot e_j) = \varphi f(e_i, e_j)e_0 = f(e_i, e_j)e_0$ for $i, j = 1, \cdots, n$. Therefore $f(\varphi e_i, \varphi e_j) = f(e_i, e_j)$ *for* $i, j = 1, \cdots, n$. Therefore the restriction of φ to $\mathfrak{B}$ is an element of $\mathfrak{G}$. Conversely any element of $\mathfrak{G}$ can be extended uniquely to an automorphism of $\mathfrak{A}$. Thus $\mathfrak{G}$ is isomorphic to the group of automorphisms of $\mathfrak{A}$. Note that if these two groups are realized as groups of matrices with respect to the given basis, then the isomorphism is trivially birational and biregular in the sense of algebraic geometry, so that the groups are isomorphic as algebraic groups. The last statement of the theorem follows from a classical result in the theory of Lie groups (cf. [3] p. 137).

THEOREM 5 (Wedderburn Principal Theorem). *Suppose that f is degenerate and either symmetric or alternating. Then $\mathfrak{A}$ has a semisimple subalgebra $\mathfrak{A}_0$ and a nilpotent ideal $\mathfrak{N}$ such that $\mathfrak{A} = \mathfrak{A}_0 + \mathfrak{N}$ (vector space direct sum).*

Proof. If f is identically zero take $\mathfrak{N} = \mathfrak{B}$ and $\mathfrak{A}_0$ to be the subalgebra spanned by e_0. Otherwise let $\mathfrak{N}_0$ be the set of elements $u \in \mathfrak{B}$ such that $f(u, v) = 0$ for all $v \in \mathfrak{B}$. Choose a basis $e_1, \cdots, e_{r+1}, \cdots, e_n$ for $\mathfrak{B}$ such that $e_{r+1} \cdots, e_n$ span $\mathfrak{N}_0$. Suppose first that $r \geqq 2$. Then $e_0, e_1, \cdots, e_r$ span a subalgebra $\mathfrak{A}_0$ which is isotopically left and right simple by Theorem 3. Taking $\mathfrak{N} = \mathfrak{N}_0$ it follows that $\mathfrak{A} = \mathfrak{A}_0 + \mathfrak{N}$ with $\mathfrak{N}$ a nilpotent ideal of index two. Now suppose $r = 1$. Then $e_1^2 = \lambda e_0$ where $\lambda \neq 0$ in k. If the subalgebra spanned by e_0 and e_1 is semisimple, then $\mathfrak{A}_0$ and $\mathfrak{N}$ may be taken as before. Otherwise, suppose that $e_0 + \beta e_1, \beta \neq 0$ in k, spans the one-dimensional radical of this subalgebra. Then take $\mathfrak{N}$ to be the ideal of $\mathfrak{A}$ spannned by $e_0 + \beta e_1$, e_2, $\cdots, e_n$ and $\mathfrak{A}_0$ to be the subalgebra spanned by e_0.

REMARK. The use of the terms "semisimple" and "nilpotent ideal" does not seem yet to be standardized in the literature on non-associative algebras. Although in the present case all of the customary interpretations of these terms are equivalent, nevertheless it desirable to give explicit definitions. An algebra is said to be *semisimple* if it is a direct sum of simple algebras, none of which is the zero algebra of dimension 1. An ideal is said to be *nilpotent* if there is an integer $m > 0$ such that every product of m elements of the ideal, irrespective of the manner of bracketing, is zero.

REFERENCES

1. A. A. Albert, *Non-associative algebras I*, Ann. of Math. **43** (1942), 685–707.
2. R. H. Bruck, *Some results in the theory of linear non-associative algebras*, Trans. Amer. Math. Soc. **56** (1944), 141–199.
3. C. C. Chevalley. *Theory of Lie groups I*, Princeton 1946.

BUCKNELL UNIVERSITY

ON TERMINATING PROLONGATION PROCEDURES*

H. H. JOHNSON

In the classical treatments [3] of systems of differential equations there are two outstanding techniques—the Cauchy-Kowalewski theorem and completely integrable systems (the latter is really a special case of the former [1, p. 77]). In terms of systems of differential forms the Cauchy-Kowalewski theorem becomes the Cartan-Kahler theorem, and systems with independent variables which satisfy its conditions are called involutive.

Many systems are not involutive, and the central problem of prolongation theory is to construct a procedure by which one can reduce every system to an equivalent involutive system. For total prolongations Kuranishi's theorem [4, p. 44] gives a precise answer to the question of when total prolongations will lead to involutive systems. If S is the initial system in euclidean space E^n, $P^g(S)$ the g^{th} total prolongation in the space R_g, then for all points $x \in E^n$, except possibly on a proper subvariety, there is a number g_0 such that if $g \geqq g_0$ and $y \in R_g$ is a point over x, then $P^g(S)$ is involutive at y if and only if y is an ordinary integral point [4, p. 7] and the 1-forms of $P^g(S)$ do not imply any dependencies among the independent variables at integral points in a neighborhood of y. Then y is called a normal point.

The first part of this paper deals with an application of this theorem to certain types of differential systems. We show that under certain conditions the total prolongation process must result in normal points if there are to be *any* solutions. An application of this leads to a theorem often used in differential geometry [2, p. 14].

The second section is concerned with what can be done if normal points are not obtained for $P^g(S)$ as is the case with an example of Kuranishi. Here we must distinguish two cases. If $P^g(S)$ does not contain ordinary intergal points, so that its 0-forms are not a regular system of equations [4, p. 7] the Cartan-Kahler theory does not apply. Let us call such systems *singular*. We shall not consider this aspect of the problem in this paper.

If, however, the problem lies in a dependency among the independent variables implied by 1-forms of $P^g(S)$, at generic integral points, one would naturally think of restricting the system to those points where dependencies do not occur, since solutions must lie only in these points. Thus one obtains a sort of partial prolongation which could in turn be prolonged. Such a procedure was certainly what Cartan and Kuranishi had in mind. However, it is not clear that the process will ever result

Received March 2, 1959, and in revised form May 25, 1959. This work was supported in part by the U. S. Army Office of Ordnance Research, Contract DA 04-200-ORD-456.

in an involutive system. One might conceivably go on obtaining non-normal systems indefinitely.

Kuranishi has recently proved a generalization of his prolongation theorem which is used to show that the above procedure does in fact ultimately stop, barring the occurrence of singular systems somewhere along the line.

The first section of this paper is part of the author's doctoral thesis at the University of California at Berkeley, written under the direction of Professor Harley Flanders to whom the author would like to record here his appreciation.

All functions, forms, and manifolds are assumed to be real analytic.

1. Kuranishi's fundamental theorem [4, p. 44] concerns a certain general type of differential system (called *normal*) which is generated by 1-forms θ^α, $\alpha = 1, \cdots, \alpha_1$. If $\omega^1, \cdots, \omega^p$ is a basis of a system of independent variables and $\pi^1, \cdots, \pi^m$ any other 1-forms to fill out a basis, then the θ^α are normal if $d\theta^\alpha$ can be expressed as

$$d\theta^\rho \equiv \sum_{i=1}^{p} \sum_{\lambda=1}^{m} A_{\varphi,i\lambda}\omega^i \wedge \pi^\lambda + \sum_{i=1}^{p} \sum_{j=1}^{p} \tfrac{1}{2} B_{\varphi,ij}\omega^i \wedge \omega^j$$

modulo (θ^α). Suppose that these are defined on E^n where $n = \alpha_1 + p + m$ of variables $x^1, \cdots, x^n$. Then R_g is the euclidean space of variables

$$x^j,\ u^\lambda_{i_1},\ u^\lambda_{i_1 i_2},\ \cdots,\ u^\lambda_{i_1 \cdots i_g}\ ,$$

where $j = 1, \cdots, n$; $i_1, \cdots, i_g = 1, \cdots, p$; $\lambda = 1, \cdots, m$, and the $u^\lambda_{ij\cdots k}$ are symmetric in the lower indicies.

Then $P^g(S)$ can be taken to be the system on R_g generated by the 1-forms

$$\pi_g \begin{cases} \theta^\alpha, \\ d\pi^\lambda - \sum_{j=1}^{n} u^\lambda_j \omega^j, \\ du^\lambda_{j_1} - \sum_{j=1}^{p} u^\lambda_{j_1 j}\omega^j, \\ \quad\vdots \\ du^\lambda_{j_1 j_2 \cdots j_{g-1}} - \sum_{j=1}^{p} u^\lambda_{j_1 j_2 \cdots j_{g-1} j}\omega^j\ , \end{cases}$$

and certain functions

$$\Theta_{\varphi;ij;k_1\cdots k_t}, \qquad t \leqq g-1\ .$$

It turns out that for $t \leqq g-2$,

$$d\Theta_{\varphi;ij;k_1\cdots k_t} \equiv 0 \qquad \text{modulo } \pi_g\ ,$$

while

$$d\Theta_{\varphi;ij;k_1\cdots k_{g-1}} \equiv \sum_{\lambda=1}^{m}(A_{\varphi;i\lambda}du^{\lambda}_{jk_1\cdots k_{g-1}} - A_{\varphi;j\lambda}du^{\lambda}_{ik_1\cdots k_{g-1}}) + \sum_{k=1}^{p} B_{\varphi;ij;k_1\cdots k_{g-1}k}\omega^k ,$$

modulo π_g.

These B's are defined inductively by

$$B_{\varphi;ij;k1\cdots k_t} = \sum_{\lambda=1}^{m}[(D_{k_t}A_{\varphi;i\lambda})du^{\lambda}_{jk_1\cdots k_{t-1}} - (D_{k_t}A_{\varphi;j\lambda})du^{\lambda}_{ik_1\cdots k_{t-1}}] + D_{k_t}B_{\varphi;ij,k_1\cdots k_{t-1}} ,$$

where D_kF is defined as follows.

If F is any function on R_{t-1} it can be considered to be a function on R_s for all $s \geqq t-1$. If we form dF, then modulo π_s, when $s \geqq t$, dF involves only $\omega^1, \cdots, \omega^p$:

$$dF \equiv \sum_{k=1}^{p} F_k\omega^k \qquad \text{modulo } \pi_s ,$$

and the F_k are independent of s so long as $s \geqq t$. Then one defines D_kF to be F_k. D_kF is a function on R_t.

THEOREM 1. *Let S be a normal system where*

(1) *the $A_{\varphi;i\lambda}$ are constants,*

(2) *$dB_{\varphi;ij} \equiv 0$ modulo $(\omega^i, \theta^\alpha)$.*

Then if $P^g(S)$ is non-singular for all g, there is a g_0 such that $P^g(S)$ is involutive for all $g \geqq g_0$ at ordinary integral points, or else there exist no solutions.

Proof. If an ordinary integral point $y \in R_g$ is not normal, then there must be a dependency among $\omega^1, \cdots, \omega^p$ implied by the 1-forms of $P^g(S)$ at integral points y_1 arbitrarily near y. This can happen only if there is a relation of the type

$$\Sigma\Gamma^{\varphi;ij;k_1\cdots k_{g-1}}(y_1)(d\Theta_{\varphi;ij;k_1\cdots k_{g-1}})_{y_1} \equiv 0 \text{ modulo } (\omega^i) ,$$

where the left side does not vanish identically. This can only happen if

$$0 = \Sigma\Gamma^{\varphi;ij;k_1\cdots k_{g-1}}(y_1)[A_{\varphi;i\lambda}(y_1)(du^{\lambda}_{jk_1\cdots k_{g-1}})_{y_1} - A_{\varphi;j\lambda}(y_1)(du^{\lambda}_{ik_1\cdots k_{g-1}})_{y_1}] ,$$

while for some k,

$$\Sigma\Gamma^{\varphi;ij;k_1\cdots k_{g-1}}(y_1)B_{\varphi;ij;k_1\cdots k_{g-1}k}(y_1) \neq 0 .$$

Since the A depend only on $x^1, \cdots, x^n$, we can choose the Γ to be functions of $x^1, \cdots, x^n$.

Now, the functions in $P^{g+1}(S)$ have the form

$$\Theta_{\varphi;ij;k_1\cdots k_g} = \sum_{\lambda=1}^{m}(A_{\varphi;i\lambda}u^{\lambda}_{jk_1\cdots k_g} - A_{\varphi;j\lambda}u^{\lambda}_{ik_1\cdots k_g}) + B_{\varphi;ij;k_1\cdots k_g} .$$

Hence we have in $P^{g+1}(S)$ the function which is not in $P^g(S)$,

$$\Sigma\Gamma^{\varphi;ij;k_1\cdots k_{g-1}}\Theta_{\varphi;ij;k_1\cdots k_{g-1}k} = \Sigma\Gamma^{\varphi;ij;k_1\cdots k_{g-1}}B_{\varphi;ij;k_1\cdots k_{g-1}k} .$$

Consider now these B. Since the A are constants,

$$B_{\varphi;ij;k_1\cdots k_t} = D_{k_t}B_{\varphi;ij;k_1\cdots k_{t-1}} ;$$

where

$$dB_{\varphi;ij;k_1\cdots k_{s-1}} \equiv \sum_{l=1}^{p} D_k B_{\varphi;ij;k_1\cdots k_{t-1}}\omega^k$$

modulo π_s.

By assumption (2), $dB_{\varphi;ij}$ have the form

$$dB_{\varphi;ij} = \sum_{k=1}^{p} C_{\varphi;ij;k}\omega^k + \sum_{\beta=1}^{\alpha_1} E_{\varphi;ij;\beta}\theta^\beta ,$$

$$= \sum_{k=1}^{p} C_{\varphi;ij;k}\omega^k \qquad \text{modulo } \pi_1 ,$$

hence

$$B_{\varphi;ij;k} = D_k B_{\varphi;ij} = C_{\varphi;ij;k}$$

are functions of $x^1, \cdots, x^n$ alone. Obviously $dB_{\varphi;ij;k} \equiv 0$ modulo $(\theta^\alpha, \omega^i)$ also, so the argument can be repeated to show that the functions

$$B_{\varphi;ij;k_1\cdots k_t}$$

depend only on $x^1, \cdots, x^n$. But that means that

$$\text{(I)} \qquad \Sigma\Gamma^{\varphi;ij;k_1\cdots k_{g-1}}B_{\varphi;ij;k_1\cdots k_{g-1}k}$$

is a function in $P^{g+1}(S)$, not in $P^g(S)$, and dependent only on $x^1, \cdots, x^n$.

Now, to any integral manifold I of S there corresponds a unique integral manifold I^{g+1} of $P^{g+1}(S)$, such that if ρ^{g+1} is the natural fibre bundle mapping on R_{g+1} to E^n, then $\rho^{g+1}(I^{g+1}) = I$ [4, p. 15]. I^{g+1} must annihilate the function (1). Since it is a function of $x^1, \cdots, x^n$ alone, I must itself annihilate it.

We conclude; if there exist ordinary integral points in R_g where $P^g(S)$ is not normal, then the manifold of integral points of S where solutions can occur must satisfy an additional condition to any imposed by $P^t(S)$, $t < g$. Clearly, this can happen at most $n - p$ times if there are to be solutions

Since the $A_{\varphi;i\lambda}$ are constants, every point of E^n is regular of order 0 [4, p. 36], so by Kuranishi's fundamental theorem there exists an integer g_1 such that if y is an ordinary integral point in R_g for $g \geqq g_1$, then $P^g(S)$ is involutive at y if and only if y is normal. Taking $g_0 = g_1 + (n - p)$ one obtains the theorem.

Next an application of this theorem will be made to a certain type of system of differential equations.

Let E^n be the euclidean space of variables $x^1, \cdots, x^p, z^1, \cdots, z^m$. Consider the problem of finding m functions $f^\lambda(x^1, \cdots, x^p) = z^\lambda$ which will satisfy a given set of first order partial differential equations

$$\frac{\partial z^\alpha}{\partial x^i} = \psi_i^\alpha(x, z), \qquad \alpha = 1, \cdots, m;\ i = 1, \cdots, p\ .$$

In terms of differential forms this is the problem of finding integral manifolds of the system S generated by the 1-forms

$$\theta^\alpha = dz^\alpha - \sum_{i=1}^{p} \psi_i^\alpha(x, z) dx^i$$

with independent 1-forms $dx^1, \cdots, dx^p$. Here there are no π^λ. Then

$$d\theta^\alpha \equiv \frac{1}{2} \sum_{i=1}^{p} \sum_{j=1}^{p} \Big(\sum_{\beta=1}^{m} \frac{\partial \psi_i^\alpha}{\partial z^\beta} \psi_j^\beta - \sum_{\beta=1}^{m} \frac{\partial \psi_j^\alpha}{\partial z^\beta} \psi_i^\beta + \frac{\partial \psi_i^\alpha}{\partial x^j} - \frac{\partial \psi_j^\alpha}{\partial x^i} \Big) dx^j \wedge dx^i \qquad \text{modulo } (\theta^\alpha)\ .$$

If then

$$B_{\alpha;ij} = \sum_{\beta=1}^{m} \frac{\partial \psi_i^\alpha}{\partial z^\beta} \psi_j^\beta - \sum_{\beta=1}^{m} \frac{\partial \psi_j^\alpha}{\partial z^\beta} \psi_i^\beta + \frac{\partial \psi_i^\alpha}{\partial x^j} - \frac{\partial \psi_j^\alpha}{\partial x^i}$$

one can deduce the following theorem from the nature of the forms

$$d\Theta_{\phi;ij;k_1\cdots k_t} \equiv B_{\varphi;ij;k_1\cdots k_t k} \omega^k\ .$$

THEOREM 2. *In order that the system of differential equations*

$$\frac{\partial z^\alpha}{\partial x^i} = \psi_i^\alpha(x, z)$$

have a solution, given that the equations

$$B_{\varphi;ij;k_1\cdots k_t} = 0, \qquad t \leqq g$$

are non-singular for all g, it is necessary and sufficient that for all $\varphi, i, j, k_1, \cdots, k_g$,

$$B_{\varphi,ij,k_1\cdots k_g} \equiv 0 \text{ modulo } (B_{\theta;rs;h_1\cdots h_t} \mid t \leqq g - 1)$$

for some $g \leqq m - 1$. [See 2, p. 14].

2. In Theorem 1, the prolongation process had to yield an involutive system because whenever a non-normal prolongation occured, this implied additional restrictions on the original system. In general this need not happen. Kuranishi gives an example of a system in which $P^g(S)$ is not normal for any $g \geqq 1$ [4, p. 45].

Normality at integral points y of $P^g(S)$ involves two conditions; the set of 0-forms of $P^g(S)$ which define y must define a regular system of equations at y, and the 1-forms of $P^g(S)$ must imply no relations among the independent variables at integral points near y. This paper will ignore the first problem. It would seem to call for a more delicate approach to the Cartan-Kahler theorem. Let y be a non-normal integral point of $P^g(S)$ such that for all integral points y_1 near y there is a dependency of the type

$$\sum_{i=1}^{p} A_i(y_1)(\omega^i)_{y_1}$$

in $P^g(S)$. Then obviously solutions can occur only at points y_1 where $A_1(y_1) = A_2(y_1) = \cdots = A_p(y_1) = 0$. Hence a natural step to solving the system would be to add $A_1, \cdots, A_p$ as 0-forms to the system $P^g(S)$. One would obtain a system having the same solutions as $P^g(S)$.

Observe also that if $P^g(S)$ contains a 0-form which is a function on R_{g-1}, obviously any solution of $P^{g-1}(S)$ must annihilate that function; hence, adding it to $P^{g-1}(S)$ would generate a system having the same solutions as $P^{g-1}(S)$.

We introduce the following definition: let the system T in independent variables $x^1, \cdots, x^p$, and dependent variables $y^1, \cdots, y^r, z^1, \cdots, z^m$ be called *complete* if the 1-forms of T contain no forms of the type $\Sigma A_i\omega^i$, where $\omega^1, \cdots, \omega^p$ is a basis of independent variables, A_i not in T.

LEMMA. *Let S be any system with independent variables $x^1, \cdots, x^p$, and dependent variables $z^1, \cdots, z^m$. Then there exists a sequence $\{S^g\}$ of differential systems S^g, closed, on R_g such that*

(1) *S^g has the same solutions as $P^g(S)$,*

(2) *S^g is complete,*

(3) *$P(S^{g-1}) \subseteqq S^g$, and*

(4) *the 0-forms of S^g contain no functions on R_{g-1} except those in S^{g-1},*

(5) *S^g is generated by 0-forms, π_g, and their derivatives.*

Proof. Let X be the set of all sequences $\{T^g \mid g = 1, 2, \cdots\}$, where T^g is a closed differential system on R_g generated by 0-forms, π_g and their derivatives and having the same solutions as $P^g(S)$ and $P(T^{g-1}) \subseteqq T^g$. The elements of X can be partially ordered by inclusion: $\{U^g\} \geqq \{T^g\}$ if $U^g \supseteqq T^g$ for all $g = 1, 2, \cdots$ If $A = \{\{T^g_a\} \mid a \in A\}$ is a nest in X, then $\{T^g\}$, where T^g is the closed differential system generated by $U\{T^g_a | a \in A\}$, is in X and is $\geqq$ every element of A. Hence, X contains a maximal element, $\{S^g\}$. By definition, $\{S^g\}$ satisfies (1) and (3). If S^h were not complete, one could add to S^h the coefficients of forms of the type $\Sigma A_i \omega^i$ to obtain a still larger system $\bar{S}^h$, and $\{\bar{S}^g\}$, where $\bar{S}^g = S^g$ for $g < h$, and $\bar{S}^g = P^{g-h}(\bar{S}^h)$ for $g \geqq h$, would be properly greater than $\{S^g\}$. Similarly, if condition (4) did not hold for some S^h, we could enlarge S^{h-1}. Hence the lemma.

The construction of such a sequence, given S, could proceed as follows. Form $P(S)$ and complete it in the obvious way to form a system T^1. If the resulting system involves any functions on R_0 i.e., depending only on the coordinates of R_0, add these to the system S and begin again. Otherwise, form $P(T^1)$ and complete to form T^2. If T^2 contains functions on R_1, add these to T_1 and begin again at that step. Observe that the addition of new functions to any one system on, say, R_g, is limited by the dimension of R_g, since each such addition reduces the dimension of the variety of integral points, which must have at least dimension p if there are to be any solutions at all.

Granted that such a sequence $\{S^g\}$ as given in the lemma exists, it is still not clear whether any S^g is involutive. Of course, the 0-forms might not define a regular system of equations for the integral points. But barring this one can prove that for g sufficiently large, S^g is involutive. This follows from a recent extension of Kuranishi's prolongation theorem [5, Theorem III. 1], where the required conditions are precisely those of the lemma.

THEOREM 3. *Given a differential system S with independent variable $dx^1, \cdots, dx^p$, there exists a sequence $\{S^g\}$ of closed differential systems, where S^g is on R_g, $g = 1, 2, \cdots$, which have the same solutions as $P^g(S)$. Moreover, if for all $g \geqq g_0$, S^g is non-singular, then there exists a g_1 such that for $g \geqq g_1$, $P(S^{g-1}) = S^g$ and S^g is involutive.*

REFERENCES

1. E. Cartan, *Les systemes differentielles exterieurs et leurs applications geometriques*, Paris, 1945.

2. L. P. Eisenhart, *Non-Riemannian geometry*, New York, 1927.

3. E. Goursat, *Lecons sur l'integration des equations aux derivees partielles du premier order*, Paris, 1921.

4. M. Kuranishi, *On E. Cartan's prolongation theorem of exterior differential systems*, Amer. J. Math., vol. **79** (1957), pp. 1–47.

5. ———, *On the abstract approach to the local theory of continuous infinite pseudo groups*, Project Report, University of Chicago, 1957.

TWO REMARKS ON FIBER HOMOTOPY TYPE

JOHN MILNOR AND EDWIN SPANIER

Section 1 of this note considers the normal sphere bundle of a compact, connected, orientable manifold M^n (without boundary) differentiably imbedded in euclidean space R^{n+k}. (These hypotheses on M^n will be assumed throughout § 1.) It is shown that if k is sufficiently large then the normal sphere bundle has the fiber homotopy type of a product bundle if and only if there exists an S-map from S^n to M^n of degree one (i.e. for some p there exists a continuous map of degree one from S^{n+p} to the p-fold suspension of M^n). The proof is based on the fact that the Thom space of the normal bundle is dual in the sense of Spanier-Whitehead [8] to the disjoint union of M^n and a point.

Section 2 studies the tangent sphere bundle of a homotopy n-sphere. This has the fiber homotopy type of a product bundle if and only if n equals 1, 3 or 7. The proof is based on Adams' work [1].

If X is a space, S^kX will denote the k-fold suspension of X as in [8, 9]. If X has a base point x_0, then S_0^kX will denote the k-fold reduced suspension and is the identification space S^kX/S^kx_0 obtained from S^kX by collapsing S^kx_0 to a point (to be used as base point for S_0^kX). There is a canonical homeomorphism $S_0^kX \approx S^k \times X/S^k \vee X$.

Two fiber bundles with the same fiber and with projections $p_1\colon E_1 \to B$, $p_2\colon E_2 \to B$ have the same *fiber homotopy type* [3, 4, 10] if there exist fiber preserving maps $f_i\colon E_i \to E_{3-i}$ and fiber preserving[1] homotopies $h_i\colon E_i \times I \to E_i$ such that $h_i(x, 0) = f_{3-i}f_i(x)$, $h_i(x, 1) = x$.

Let ξ denote an oriented $(k-1)$-sphere bundle. The total space of ξ will be denoted by $\dot{E}$ and the total space of the associated k-disk bundle will be denoted by E. The *Thom space* $T(\xi)$ is the identification space $E/\dot{E}$ obtained from E by collapsing $\dot{E}$ to a single point (to be used as base point for $T(\xi)$). The following are easily verified:

(A) *If ξ_1, ξ_2 are $(k-1)$-sphere bundles of the same fiber homotopy type, then $T(\xi_1)$, $T(\xi_2)$ have the same homotopy type.*

(B) *If ξ is a product bundle, then $T(\xi)$ is homeomorphic to $S_0^k(B \cup p_0)$ (where $B \cup p_0$ is the disjoint union of B and a point, p_0, which is taken as the base point of $B \cup p_0$).*

1. The normal bundle.

If X and Y are spaces we let $[X, Y]$ denote the set of homotopy classes of maps of X into Y and we let

Received May 19, 1959. The authors were supported by the Sloan Foundation and by the United States Air Force (Contract No. AF 49(638)–393 monitored by the Air Force Office of Scientific Research), respectively, during the period when this paper was in preparation.

[1] The phrase "fiber-preserving" means that $p_{3-i}f_i(x) = p_i(x)$ and $p_ih_i(x, t) = p_i(x)$.

$\{X, Y\}$ denote the set of S-maps of X into Y as in [8]. Thus, $\{X, Y\}$ is defined to be the direct limit of the sequence

$$[X, Y] \xrightarrow{S} [SX, SY] \xrightarrow{S} \cdots \xrightarrow{S} [S^pX, S^pY] \xrightarrow{S} \cdots$$

There is a natural map

$$\phi\colon [X, Y] \longrightarrow \{X, Y\}$$

which assigns to every homotopy class $[f] \varepsilon [X, Y]$ the S-map $\{f\}$ represented by any map of $[f]$. The following gives a sufficient condition for ϕ to be onto $\{X, Y\}$.

LEMMA 1. *Let Y be a k-connected CW-complex ($k \geq 1$) and let X be a finite CW-complex with*[2] $H^q(X) = 0$ *for* $q > 2k+1$. *Then* $\phi([X, Y]) = \{X, Y\}$.

Proof. It suffices to prove that under the hypotheses of the lemma the map $S\colon [X, Y] \to [SX, SY]$ is onto $[SX, SY]$ because then, for each $p \geq 0$, the map $S\colon [S^pX, S^pY] \to [S^{p+1}X, S^{p+1}Y]$ is onto $[S^{p+1}X, S^{p+1}Y]$ (because S^pY is $(p + k)$-connected and $H^q(S^pX) = 0$ for $q > 2k + p + 1$ and $2(k + p) + 1 \geq 2k + p + 1$).

Choose base points $x_0 \varepsilon X$, $y_0 \varepsilon Y$ and let $[X, Y]'$ denote the set of homotopy classes of maps $(X, x_0) \to (Y, y_0)$. Since Y is simply-connected the natural map $[X, Y]' \to [X, Y]$ is a 1–1 correspondence. Since X, Y are CW-complexes the collapsing maps $SX \to S_0X$ and $SY \to S_0Y$ are homotopy equivalences (Theorem 12 of [11]) so there are 1–1 correspondences

$$[S_0X, S_0Y] \approx [S_0X, SY] \approx [SX, SY] .$$

Since S_0Y is simply connected, we also have a 1–1 correspondence $[S_0X, S_0Y]' \approx [S_0X, S_0Y]$. Hence, it suffices to show that $S_0([X, Y]') = [S_0X, S_0Y]'$.

Let ΩS_0Y denote the space of closed paths in S_0Y based at y_0. There is a canonical 1–1 correspondence $[S_0X, S_0Y]' \approx [X, \Omega S_0Y]'$ and a natural imbedding $Y \subset \Omega S_0Y$ such that the map $S_0\colon [X, Y]' \to [S_0X, S_0Y]'$ corresponds to the injection (see § 9 of [7])

$$[X, Y]' \longrightarrow [X, \Omega S_0Y]' .$$

Hence, it suffices to show this injection is onto or, equivalently, that the natural injection (without base point condition) $[X, Y] \to [X, \Omega S_0Y]$ is onto.

[2] When no coefficient group appears explicitly in the notation for a homology or cohomology group it is to be understood that the coefficient group is the group of integers. In dimension 0 the groups will be taken reduced.

Since Y is k-connected it follows from the suspension theorem (see § 7 of [9]) that

$$S_0\colon \pi_i(Y) \longrightarrow \pi_{i+1}(S_0 Y)$$

is 1–1 if $i \leq 2k$ and is onto if $i \leq 2k+1$. Since S_0 corresponds to the injection map $\pi_i(Y) \to \pi_i(\Omega S_0 Y)$, this is equivalent to the statement that

$$\pi_i(\Omega S_0 Y,\ Y) = 0 \text{ for } i \leq 2k+1 .$$

Since Y is simply-connected the groups $\pi_i(\Omega S_0 Y,\ Y)$ form a simple system for every i. Now the groups $H^i(X;\ \pi_i(\Omega S_0 Y,\ Y))$ vanish for every i because for $i \leq 2k+1$ the coefficient group vanishes while for $i > 2k+1$ the groups vanish because of the assumption on the cohomology of X. By Theorem 4.4.2 of [2] it follows that any map $X \to \Omega S_0 Y$ is homotopic to a map $X \to Y$, completing the proof.

REMARK. If in Lemma 1 we assume that $H^q(X) = 0$ for $q > 2k$, then a similar argument shows that ϕ is 1–1, however we shall not need this result.

Let $M^n \subset R^{n+k}$ be as in the introduction (i.e. M^n is a differentiably imbedded manifold which is compact, connected, orientable, and without boundary). The following result relates the normal bundle of M^n to M^n itself by means of duality.

LEMMA 2. *Let ξ be the normal $(k-1)$-sphere bundle of M^n in R^{n+k}. Then the Thom space $T(\xi)$ is weakly $(n+k+1)$-dual to the disjoint union $M^n \cup p_0$.*

Proof. Regard S^{n+k} as the one point compactification of R^{n+k}. Let E be a closed tubular neighborhood of M^n and assume E is contained in a large disk D^{n+k}. Then (D^{n+k}-interior E) is a deformation retract of $R^{n+k} - M^n = S^{n+k} - (M^n \cup (\text{point at infinity}))$. Using standard homotopy extension properties and the contractibility of D^{n+k} it follows that if $\dot{E}$ denotes the boundary of E then

$$T(\xi) = E/\dot{E} = D^{n+k}/(D^{n+k} - \text{interior } E)$$

has the homotopy type of the suspension $S(D^{n+k} - \text{interior } E)$. Since $(D^{n+k} - \text{interior } E)$ is an $(n+k)$-dual of $M^n \cup (\text{point at infinity})$, and the suspension of an $(n+k)$-dual is an $(n+k+1)$-dual, this completes the proof.

REMARK. Lemma 2 shows that the S-type of $T(\xi)$ depends only on that of M^n. If k is sufficiently large this implies that the homotopy type of $T(\xi)$ depends only on that of M^n. This suggests the conjecture that the fiber homotopy type of the normal bundle of any manifold $M^n \subset R^{n+k}$, k large, is completely determined by the homotopy type of M^n. A similar conjecture can be made for the tangent bundle.

THEOREM 1. *Let $M^n \subset R^{n+k}$ be as before and assume that $H_q(M^n) = 0$ for $q < r$ and that $k \geq \min(n - r + 2,\ 3)$. The following statements are equivalent:*

(1) *There is an S-map $\alpha \varepsilon \{S^n,\ M^n\}$ such that*

$$\alpha_*\colon H_n(S^n) \approx H_n(M_n)\ .$$

(2) *The normal sphere bundle of $M^n \subset R^{n+k}$ has the fiber homotopy type of a product bundle.*

(3) *The disjoint union $M^n \cup p_0$ is weakly $(n + k + 1)$-dual to $S_0^k(M^n \cup p_0)$.*

Proof. (1) $\Rightarrow$ (2). Let N denote the complement in S^{n+k} of an open tubular neighborhood of M^n. Then N is $(n + k)$-dual to M^n. The S-map α is $(n+k)$-dual to an S-map $\beta \varepsilon \{N, S^{k-1}\}$ such that $\beta^*\colon H^{k-1}(S^{k-1}) \approx H^{k-1}(N)$. Since $H^p(N) \approx H_{n+k-p-1}(M^n)$, we see that $H^p(N) = 0$ if $p > n + k - r - 1$. Since S^{k-1} is $(k - 2)$-connected, $k - 2 \geq 1$, and $k \geq n - r + 2$ (so $2(k-2)+1 \geq n+k-r-1$), it follows from Lemma 1 that there is a map $f\colon N \to S^{k-1}$ representing β. Then $f^*\colon H^{k-1}(S^{k-1}) \approx H^{k-1}(N)$. Let $\dot{E}$ be the boundary of N (so $\dot{E}$ is the normal $(k - 1)$-sphere bundle of M^n), and let F be a fiber of $\dot{E}$. Then the inclusion map $F \subset N$ induces an isomorphism $H^{k-1}(N) \approx H^{k-1}(F)$ (because by Corollaries III. 15 and I.5 of [10] or by Theorems 14 and 21 of [5] we have $H^{k-1}(\dot{E}) \approx H^{k-1}(M^n) + Z$ and the injection $H^{k-1}(N) \to H^{k-1}(\dot{E})$ maps isomorphically onto the second summand while the injection $H^{k-1}(\dot{E}) \to H^{k-1}(F)$ maps the second summand isomorphically.) Therefore, the map $f \mid \dot{E}\colon \dot{E} \to S^{k-1}$ has the property that its restriction to a fiber F induces an isomorphism of the cohomology of S^{k-1} onto that of F so is a homotopy equivalence of F with S^{k-1}. This implies (by Corollary 2 on p. 121 of [3]) that $\dot{E}$ has the same fiber homotopy type as a product bundle.

(2) $\Rightarrow$ (3). By Lemma 2, $T(\xi)$ is weakly $(n + k + 1)$-dual to $M^n \cup p_0$. If ξ is of the same fiber homotopy type as a product bundle, it follows from (A), (B) that $T(\xi)$ is of the same homotopy type as $S_0^k(M^n \cup p_0)$. Combining these two statements gives the result.

(3) $\Rightarrow$ (1) assume $M^n \cup p_0$ is weakly $(n + k + 1)$-dual to $S_0^k(M^n \cup p_0)$. The map $M^n \cup p_0 \to S^0$ collapsing each component of $M^n \cup p_0$ to a single point represents an S-map $\beta\colon S_0^k(M^n \cup p_0) \to S_0^k(S^0) = S^k$ such that $\beta^*\colon H^k(S^k) \approx H^k(S_0^k(M^n \cup p_0))$. By duality there is an S-map $\alpha \varepsilon \{S^n, M^n \cup p_0\}$ such that $\alpha_*\colon H_n(S^n) \approx H_n(M^n \cup p_0) \approx H_n(M^n)$. Since.

$$\{S^n,\ M^n \cup p_0\} \approx \{S^n,\ M^n\} + \{S^n,\ S^0\}\ ,$$

the result is proved.

As a corollary we obtain the following result proved by Massey [4].

COROLLARY. *Let M^n be a homology sphere. Then the normal bundle of M^n in R^{n+k} has the same fiber homotopy type as a product bundle.*

Proof. Since $r = n$, the case $k \geq 3$ follows from the theorem. For the cases $k = 1,\ 2$ it is well known that the normal bundle is, in fact, trivial.

REMARK. Puppe [6] calls a manifold "sphere-like" if the unstable group $\pi_{n+1}(SM^n)$ contains an element of degree one. (The group $\pi_n(M^n)$ can contain an element of degree one if and only if M^n is a homotopy sphere.) Theorem 1 shows that the normal sphere bundle of a sphere-like manifold $M^n \subset R^{n+k}$ has the fiber homotopy type of a product bundle provided k is sufficiently large. An example of a manifold with trivial normal bundle which is not sphere-like is provided by the real projective 3-space.

2. The tangent bundle. Let M^n be as above (i.e. compact, connected, orientable, and without boundary), but let E denote a closed tubular neighborhood of the diagonal in $M^n \times M^n$. If the tangent bundle has the fiber homotopy type of a product bundle, then there exists a map $\dot{E} \rightarrow S^{n-1}$ (where $\dot{E}$ is the boundary of E) having degree one on each fiber. This gives rise to a map $(E,\ \dot{E}) \rightarrow (D^n,\ S^{n-1}) \rightarrow (S^n,$ point) of degree one and, hence, to a map

$$M^n \times M^n \longrightarrow M^n \times M^n/(M^n \times M^n\text{-interior } E) = E/\dot{E} \longrightarrow S^n$$

which has degree (1, 1) (the degree is (1, 1) because a generator of $H^n(S^n)$ maps, under the homomorphism induced by the above composite, into a cohomology class of $M^n \times M^n$ dual under Poincaré duality to the diagonal class of $H_n(M^n \times M^n)$).

THEOREM 2. *Suppose that M^n has the homotopy type of an n-sphere. Then the tangent bundle has the fiber homotopy type of a product bundle if and only if n equals* 1, 3 *or* 7 (*and in this case the tangent bundle is a product bundle*).

Proof. If a map $S^n \times S^n \rightarrow S^n$ of degree (1, 1) exists, then according to Adams n must be equal to 1, 3 or 7 (see Theorem 1a of [1]).

Conversely, if n equals 1, 3 or 7 then $\pi_{n-1}(SO(n)) = 0$. Using obstruction theory it follows that any homotopy n-sphere is parallelizable. This completes the proof.

REFERENCES

1. F. Adams, *On the nonexistence of elements of Hopf invariant one*, Bull. Amer, Math. Soc., **64** (1958), 279–282.
2. A. L. Blakers and W. S. Massey, *The homotopy groups of a triad* I, Ann. of Math., **53** (**1951**), 161–205.
3. A. Dold, *Über fasernweise Homotopieäquivalenz von Faserräumen*, Math. Zeits., **62** (1955), 111–136.
4. W. S. Massey, *On the normal bundle of a sphere imbedded in Euclidean space* (to appear).
5. J. Milnor, *Lectures on characteristic classes* (mimeographed), Princeton University, 1958.
6. D. Puppe, *Homotopiemengen und ihre induzierten Abbildungen* II, Math. Zeits., **69** (1958), 395–417.
7. E. Spanier, *Infinite symmetric products, function spaces, and duality*, Ann. of Math., **69** (1959), 142–198.
8. E. Spanier and J. H. C. Whitehead, *Duality in homotopy theory*, Mathematika, **2** (1955), 56–80.
9. E. Spanier and J. H. C. Whitehead, *The theory of carriers and S-theory*, in Algebraic Geometry and Topology (a symposium in honor of S. Lefschetz), Princeton University Press, 1957, pp. 330–360.
10. R. Thom, *Espaces fibrés en sphères et carrés de Steenrod*, Ann. Ecole Norm. Sup., **69** (1952), 109–182.
11. J. H. C. Whitehead, *Combinatorial Homotopy* I, Bull, Amer. Math. Soc., **55** (1949), 213–245.

PRINCETON UNIVERSITY
UNIVERSITY OF CHICAGO AND
THE INSTITUTE FOR ADVANCED STUDY

A NOTE ON ASSOCIATIVITY

DONALD A. NORTON

1. Introduction. In a groupoid with binary operation $(\cdot)$ the constraints that the groupoid be a quasigroup[1] and that it be associative are not independent. This note defines three forms of associativity in order of descending strength and shows that in a groupoid they are essentially independent while in a quasigroup (with minor limitations on the number of elements) the stronger implies the weaker. Let us define:

A groupoid is *tri-associative* if for every triple x, y, z of distinct elements

$$x \cdot (y \cdot z) = (x \cdot y) \cdot z \; ; \tag{1}$$

A groupoid is *di-associative*[2] if in (1) above, exactly two of the elements are distinct;

A groupoid is *mono-associative* if (1) is true when all three x, y and z are equal.

The next section shows that a tri-associative quasigroup Q which contains sufficient elements (seventeen are adequate) for which $Q^2 = \{q^2, \text{ all } q \in Q\}$ also contains sufficient elements (seventeen are again adequate) is di-associative. Further, any di-associative quasigroup is mono-associative. The restrictions on the minimum number of elements in Q and Q^2 are necessitated by the method of proof for which there does not seem any essential improvement but Theorem II is probably true for all quasigroups. An examination of all possibilities indicates its validity if Q contains no more than 5 elements.

The final section illustrates, by examples, the falseness of these theorems if the assumption that Q is a quasigroup is removed.

2. Associativity conditions. We shall first prove a theorem of interest in its own right but which contributes little to the main theorems-Theorems II and III.

THEOREM I. *A tri-associative quasigroup Q has a unity element.*

Before proving the theorem it is convenient to have

LEMMA. *There exists no idempotent tri-associative quasigroup Q containing at least 2 elements.*

Proof of Lemma. We shall use product as our operation in Q with

Received May 21, 1959.

[1] For definitions of groupoid and quasigroup see, for instance, [**1**, pp. 1, 8, 15].

[2] This definition differs from the one used by this author [**2**, p. 59] in which di-associativity included power-associativity, and thereby mono-associativity. Theorem III shows that, in a quasigroup, this distinction is vacuous.

the usual conventions of juxtaposition of u and x to mean the binary product of u and x and the notation $a \cdot ux$ to mean $a(ux)$.

Suppose that $q^2 = q$, all $q \in Q$. For fixed $q \in Q$ let $u \in Q$, $u \neq q$. Then if x is the solution of $q = ux$, it is true that $x \neq q, u$; for if:

(a) $x = q$, $q = uq = q^2$ implies $u = q$;

(b) $x = u$, $q = u^2$. But $u^2 = u$ implies $u = q$.

Either is a contradiction.

Now consider $q^2 = q$. Since $q = ux$, substitution yields $q \cdot ux = ux$. Since $u \neq q \neq x \neq u$, tri-associativity implies $qu \cdot x = ux$, from which $qu = u = u^2$. So $q = u$; a contradiction. We are now ready for:

Proof of Theorem I. If Q contains 1, 2, or 3 elements an examination of possibilities yields the theorem. So suppose that Q contains at least 4 elements.

Q is not idempotent by preceding lemma so there is an $a \in Q$ so that $a^2 \neq a$. Let $ae = a$ whence $e \neq a$. Now choose some $b \neq a, e$. Tri-associativity yields $a \cdot eb = ae \cdot b = ab$; and since Q is a quasigroup

(1) $$eb = b \text{ for all } b \neq a, e .$$

Finally choose $c \in Q$, $c \neq b, e$. As before $cb = c \cdot eb = ce \cdot b$ and

(2) $$ce = e \text{ for all } c \neq b, e .$$

Therefore, combining (1) and (2), we see that e is a unity except perhaps for the products ea, ee, and be. Listing the possible values of the products from (1):

I (a) $ea = a$; $ee = e$; I (b) $ea = e$; $ee = a$;

and from (2):

II (a) $be = b$; $ee = e$; II (b) $be = e$; $ee = b$.

Now I(b) and II(b) are inconsistent since $a \neq b$. Similarly II(a) and I(b) or I(a) and II(b) are inconsistent since $e \neq a$, and $e \neq b$ respectively.

This leaves I(a) and II(a), or $ea = a$

$$ee = e$$
$$be = b$$

and e is a unity element.

We can now prove

THEOREM II. *Let Q be a tri-associative quasigroup for which both Q and $Q^2 = \{q^2;\ all\ q \in Q\}$ contain a "sufficient number" of elements, then Q is di-asssociative.*

Proof. There are 3 equalities to show, where $a \neq b$:

(1) $a \cdot ab = a^2 \cdot b$;

(2) $a \cdot ba = ab \cdot a$;

(3) $b \cdot a^2 = ba \cdot a$.

Because of the symmetry of the postulates, it is necessary to prove only one of (1) and (3). We shall prove (1) and (2).

As the proof will be given, each step of it has restrictions on the elements which will be listed and considered at the end.

(1) *Proof*	*Restrictions on elements*
$a \cdot ab$	
$= xy \cdot ab$	(a) $xy = a$
$= x(y \cdot ab)$	(b) $x \neq ab \neq y \neq x$
$= x(ya \cdot b)$	(c) $y \neq a \neq b \neq y$
$= (x \cdot ya)b$	(d) $x \neq ya \neq b \neq x$
$= (xy \cdot a)b$	(e) $x \neq y \neq a \neq x$
$= a^2b.$	

Let us now consider the restrictions:

(a) Since Q is a quasigroup, given either x, or y the other can always be found.

(b) If Q contains sufficient elements it is always possible to find x and y; $x \neq ab$, $y \neq ab$.

We next note that if Q^2 contains n elements, there will be at least n or $n-1$ pairs, x, y, $x \neq y$ for which $xy = a$, (the number depending on whether or not $a \in Q^2$).

(c) Conditions $y \neq a, b$ can always be satisfied if Q contains sufficient pairs to satisfy (a) and Q^2 enough to also satisfy (b) as well.

(d) The same as (c) may be said about the conditions $ya \neq b$ and $x \neq b$. Consider now the condition $x \neq ya$. Then $x^2 \neq x \cdot ya$.

Before proceeding we can also satisfy (e) which is a condition similar to (c).

Now since $x \neq y$; $x, y \neq a$.

$$x^2 \neq x \cdot ya = xy \cdot a = a^2$$

Conversely, if

$$x^2 \neq a^2 = xy \cdot a = x \cdot ya$$

then $x \neq ya$.

So the remaining condition of (d) can be satisfied if Q^2 contains an adequate number of elements.

The proof of (2) is parallel.

(2)
$$\begin{aligned} & a\cdot ba \\ &= xy\cdot ba \\ &= x(y\cdot ba) \\ &= x(yb\cdot a) \\ &= (x\cdot yb)a \\ &= (xy\cdot b)a \\ &= ab\cdot a \end{aligned}$$

(a) $xy = a$

(b) $ba \neq x \neq y \neq ba$.

(c) $y \neq a \neq b \neq y$.

(d) $x \neq yb \neq a \neq x$,

(e) $x \neq y \neq b \neq x$.

Condition (a), (c), (e) and $a \neq x$ of (d) have already been met previously. Condition (b) is a condition similar to (b) of the previous part and can be similarly met if Q contains adequate elements. The condition

$$\begin{aligned} & x \neq yb \text{ of part (d) yields} \\ & x^2 \neq x\cdot yb \\ & x^2 \neq xy\cdot b \\ & x^2 \neq ab \ . \end{aligned}$$

Again if Q^2 contains a sufficient number of elements, this may be met.

To complete this section we shall prove

THEOREM III. *If a quasigroup Q satisfies the constraint $x\cdot xy = x^2y$ when $x \neq y$, then Q is mono-associative.*

Proof. We must show that $q\cdot q^2 = q^2\cdot q$, all $q \in Q$. Since Q is a quasigroup, $\exists\, x$ so that

$$a\cdot a^2 = a^2x \ .$$

If $x \neq a$, from the condition of the theorem

$$a\cdot ax = a^2x \ .$$

Then $a^2 = ax$ since Q is a quasigroup and $a = x$, a contradiction. So it must be that $a = x$.

COROLLARY. *A di-associative quasigroup is mono-assoiciative.*

3. Associativity conditions for groupoids.

EXAMPLE I. The groupoid whose multiplication table is displayed is trivially tri-associative since any triple of distinct elements must contain c and so the product must be c. However, it is not di-associative since

$\cdot$	a	b	c
a	b	b	c
b	a	b	c
c	c	c	c

$$ab\cdot a = ba = a \text{ while } a\cdot ba = a^2 = b \ ;$$

nor is it mono-associative since

$$a^2\cdot a = ba = a \text{ while } a\cdot a^2 = ab = b \ .$$

EXAMPLE II. The groupoid whose multiplication table is displayed is di-associative as an examination of all possible triple products containing two distinct elements will reveal but it is not mono-associative since

$$aa^2 = ab = y \text{ while } a^2a = ba = x .$$

$\cdot$	a	b	x	y
a	b	y	a	a
b	x	a	b	b
x	a	b	x	x
y	a	b	x	x

These examples illustrate that for the groupoid the "stronger" associativity assumption does not imply the weaker, while examples of power-associative and Moufang loops illustrate that, even for quasigroups the "weaker" do not imply the "stronger".

REFERENCES

1. R. H. Bruck, *A survey of binary systems*, Springer-Verlag. Berlin, 1958.
2. D. A. Norton, *Hamiltonian loops*, Proc. Amer. Math. Soc., **3** (1952), 56–65.

UNIVERSISY OF CALIFORNIA
DAVIS, CALIFORNIA

ON THE EXTENSIONS OF A TORSION MODULE

R. J. NUNKE

This paper concerns the structure of $\mathrm{Ext}(A, T) = \mathrm{Ext}_R^1(A, T)$ where A is a torsion-free and T is a torsion module over a Dedekind ring R. In the first section it is shown that for a given torsion-free module A the structure of $\mathrm{Ext}(A, T)$ is completely determined by the basic subgroup of T. If in addition T is primary the structure of $\mathrm{Ext}(A, T)$ depends on a single known invariant of T, called by Szele [4] the critical number. The rest of the paper is devoted to showing the nature of this dependence in the special case in which A is the quotient field of R and T is primary. The problem reduces to that of computing the rank of certain complete modules over a discrete valuation ring. This computation is carried out in section two and the description of $\mathrm{Ext}(A, T)$ is given in section three.

Throughout the paper R is assumed to be a Dedekind ring other than a field. A consequence of this assumption, used in section two, is that R is infinite. An exact sequence $0 \to A' \to A \to A'' \to 0$ and a module C give rise to two exact sequences. We follow S. MacLane in calling the one beginning $0 \to \mathrm{Hom}(A'', C)$ the *first exact sequence* and the one beginning $0 \to \mathrm{Hom}(C, A')$ the *second exact sequence*.

1. In this section it is shown that whenever A is torsion-free and C is a torsion module, then the structure of $\mathrm{Ext}(A, C)$ depends only on the basic submodule of C.

LEMMA 1.1. *If A, B, C are modules with A torsion-free and if there is a homomorphism of B into C with divisible cokernel, then* $\mathrm{Ext}(A, C)$ *is a direct summand of* $\mathrm{Ext}(A, B)$.

Proof. Suppose that $f: B \to C$ is a homomorphism with $\mathrm{Coker}\, f = C/Im f$ divisible. Let f be factored into an epimorphism g followed by a monomorphism $h: f = hg$. We get two exact sequences

$$\begin{array}{l} 0 \longrightarrow Im f \xrightarrow{h} C \longrightarrow \mathrm{Coker}\, f \longrightarrow 0 \\ 0 \longrightarrow \mathrm{Ker} f \longrightarrow B \xrightarrow{g} Im\ f \longrightarrow 0\,, \end{array}$$

and the relevant parts of the associated second exact sequences are

Received May 6, 1959.

$$\begin{array}{l} \operatorname{Hom}(A, \operatorname{Coker} f) \longrightarrow \\ \operatorname{Ext}(A, Im\, f) \xrightarrow{h^*} \operatorname{Ext}(A, C) \longrightarrow \operatorname{Ext}(A, \operatorname{Coker} f) \longrightarrow 0 \\ \operatorname{Ext}(A, \operatorname{Koker} f) \longrightarrow \operatorname{Ext}(A, B) \xrightarrow{g^*} \operatorname{Ext}(A, Im\, f) \longrightarrow 0 . \end{array}$$

Since A is torsion-free all the terms with Ext in them are divisible. But the divisibility of Coker f implies that Hom $(A, \operatorname{Coker} f)$ is also divisible. For suppose that $\varphi : A \to \operatorname{Coker} f$ is a given homomorphism and r is any nonzero element of R. Since A is torsion-free, division by r in A is unique; hence there is a homomorphism $\psi: rA \to \operatorname{Coker} f$ defined by $\psi(ra) = \varphi(a)$ for a in A. Since Coker f is divisible ψ can be extended to all of A. Then $r\psi(a) = \psi(ra) = \varphi(a)$ so that $r\psi = \varphi$ and φ is divisible by r.

Hence all the modules in the last two exact sequences are divisible and the images of the various homomorphisms are direct summands. In addition Ext $(A, \operatorname{Coker} f) = 0$ because Coker f is divisible. It follows that Ext (A, C) is a direct summand of Ext $(A, Im\, f)$ which is in turn a direct summand of Ext (A, B).

COROLLARY 1.2. *If A is torsion-free and each of B and C has a homomorphism into the other with divisible cokernel, then*

$$\operatorname{Ext}(A, B) \approx \operatorname{Ext}(A, C) .$$

Proof. A divisible R-module is a direct sum of submodules each of which is isomorphic to Q or to a primary component of Q/R, the number of summands of each type being independent of the decomposition.

THEOREM 1.3. *If A is torsion-free, C is a torsion module, and B is a basic submodule of C, then*

$$\operatorname{Ext}(A, C) \approx \operatorname{Ext}(A, B) .$$

Proof. A basic submodule of a torsion module is a pure submodule for which the factor module is divisible and which is a direct sum of cyclic modules. Hence there is a homomorphism of B into C with divisible cokernel. On the other hand Szele has shown in [4] that B is a homomorphic image of C (Szele's proof is for primary groups but the generalization to this case is trivial). Hence the hypotheses of Corollary 1.2 are satisfied and the conclusion follows.

Suppose now that P is a prime ideal of R and that T is a P-primary module. The order ideal of an element x of T has the form $P^{e(x)}$ with $e(x)$ a nonnegative integer which we will call the *exponential order* of x. The submodule of T consisting of those elements with exponential order ≤ 1 is a vector space over the field R/P; its dimension will be called the P-rank of T and will be denoted by $r_P(T)$. If B is a basic submodule of T, the minimum of the numbers $r_P(P^nB)$ with n ranging over the non-negative integers is independent of the choice of B because the basic submodules of T are all isomorphic. This number is thus an invariant of T. We shall follow Szele in calling it the *critical number* of T.

If the basic submodule B of T is decomposed into the direct sum of cyclic modules, then $r_P(P^nB)$ is the number of summands whose generators have exponential order $> n$. Hence $r_P(P^nB)$ finite implies that the orders of the elements of B are bounded and the critical number of T is then 0. Thus the critical number of T is either 0 or infinite, and if 0, B is a direct summand of T which is therefore a direct sum of a divisible module and a module all of whose elements have bounded order.

THEOREM 1.4. *Let T be a P-primary module with critical number $\aleph$ and let A be torsion-free.*

(i) *If* $\aleph = 0$, *then* $\mathrm{Ext}(A, T) = 0$.

(ii) *If $\aleph$ is infinite and M is the direct sum of $\aleph$ copies of $\sum_n R/P^n$, then* $\mathrm{Ext}(A, T)$ *and* $\mathrm{Ext}(A, M)$ *are isomorphic. Thus the module structure of* $\mathrm{Ext}(A, T)$ *depends only on the critical number of T.*

Proof. Since the maximal divisible submodule of T is a direct summand of T and contributes neither to $\mathrm{Ext}(A, T)$ nor to the critical number of T, we may as well assume T reduced. In the paragraph preceding the theorem it was shown that if $\aleph = 0$, the orders of the elements of T are bounded. Any extension of T having a torsion-free factor module contains T as a pure submodule. Hence it splits and $\mathrm{Ext}(A, T)=0$ in this case.

Suppose now that $\aleph$ is infinite and M is the direct sum of $\aleph$ copies of $\sum_n R/P^n$. By Theorem 1.3 $\mathrm{Ext}(A, T) \approx \mathrm{Ext}(A, B)$ where B is a basic submodule of T. We write $B = \sum_n B_n$ where each B_n is a direct sum of copies of R/P^n. There is a natural number m such that $\aleph = r_P(P^mB)$ and $B = B' + B''$ where B' is the sum of the B_n with $n \leq m$ and B'' is the sum of the remaining B_n. Since $P^mB' = 0$ and A is torsion-free, $\mathrm{Ext}(A, B') = 0$. Then the additivity of Ext implies that $\mathrm{Ext}(A, B) \approx \mathrm{Ext}(A, B'')$. The module B'' is the direct sum of cyclic modules and $r_P(B'') = r_P(P^mB'') = \aleph$ so that B'' is generated by $\aleph$ elements. Hence it is a homomorphic image of M. On the other hand B'' can be expressed as a direct sum $B'' = C + \sum_\gamma C_\gamma$ where the summands C_γ are $\aleph$ in number and each C_γ is the direct sum of a sequence of cyclic modules whose orders are strictly increasing. It follows that M is also a homomorphic image of B'', hence $\mathrm{Ext}(A, B'') \approx \mathrm{Ext}(A, M)$ by Corollary 1.2. This proves (ii).

2. In this section we assume that R is a discrete valuation ring with prime p. If M is an R-module for which the submodules p^nM have intersection 0 (i. e. if M has no elements of infinite height), then these submodules are a base at 0 for a topology called the *p-adic* topology. The completion of M in this topology will be denoted by M^*. The p-adic topology on M induces a topology on each submodule N which may or may not coincide with the p-adic topology on N. The two topologies will certainly coincide if N is pure in M for then $p^nN = N \cap p^nM$ for all n.

The problem to be solved in this section is that of determining the rank of M^* where M is a direct sum of copies of $\Sigma_n R/p^nR$.

A subset X of an R-module A is called independent if $r_1x_1 + \cdots + r_nx_n = 0$ implies $r_1 = \cdots = r_n = 0$ whenever $x_1, \cdots, x_n$ are distinct elements of X and $r_1, \cdots, r_n$ are elements of R. The cardinal $|X|$ of a maximal independent subset of A is an invariant of A called its *rank* (denoted by $r(A)$); the rank of A is in fact the dimension of $A \otimes_R Q$ as a vector space over Q. The rank formula

$$r(A) = r(B) + r(A/B)$$

holds for any R-modules A and B with B a submodule of A. If A is torsion-free its cardinal $|A|$ and its rank are connected by the relation

$$|A| = r(A)\,|R|\,.$$

In particular $r(A) = |A|$ wherenever A is torsion-free and $r(A) \leq |R|$. (The properties mentioned is this paragraph hold for any Dedekind ring.)

LEMMA 2.1. *If $M = \Sigma_\gamma M_\gamma$ is the direct sum of the modules M_γ, each of which is without elements of infinite height then M^* is isomorphic to the submodule of the direct product $\Pi_\gamma M^*_\gamma$ consisting of those sequences $u = (u_\gamma)$ such that (*) for each natural number n, $u_\gamma \in p^nM^*_\gamma$ for all but a finite set of indices.*

The condition (*) implies that $u_\gamma = 0$ for all but a countable set of indices.

Proof. For each index γ M_γ is pure in M which is pure in M^*. Hence M_γ is pure in M^*. By Lemma 20 of [2] the closure $M_{\bar\gamma}$ of M_γ in the p-adic topology is also pure in M^*. Therefore M^* induces the p-adic topology on $M_{\bar\gamma}$ and, since a closed subspace of a complete space is complete, $M_{\bar\gamma} = M^*_\gamma$.

We next show that the sum $\Sigma_\gamma M^*_\gamma \subseteq M^*$ is direct. Suppose $\Sigma_\gamma x_\gamma = 0$ where $x_\gamma \in M^*$ and γ belongs to a finite set σ of indices. For each natural number n and each $\gamma \in \sigma$ there is an $x_{\gamma n} \in M_\gamma$ such that $x_{\gamma n} - x_\gamma \in p^nM^*_\gamma$, hence $\Sigma_\gamma x_{\gamma n} = \Sigma_\gamma(x_{\gamma n} - x_\gamma) \in p^nM^*$. Since $\Sigma_\gamma M_\gamma$ is pure in M^* it is pure in $\Sigma_\gamma M^*_\lambda$ so that $\Sigma_\gamma x_{\gamma n} \in (\Sigma_\gamma M_\gamma) \cap p^n \Sigma_\gamma M^*_\gamma = p^n \Sigma_\gamma M_\gamma$. Then $x_{\gamma n} \in p^n M$ for each $\gamma \in \sigma$ because the sum $\Sigma_\gamma M_\gamma$ is direct. Thus for each $\gamma \in \sigma$, $x_{\gamma n} \to 0$ and $x_\gamma = 0$.

Let S be the submodule of $\Pi_\gamma M^*_\gamma$ defined by (*). We shall define an isomorphism φ of M^* onto S. Let x be any element of M^*. Since $\Sigma_\gamma M^*_\gamma$ is dense in M^* there is, for each natural number n, an element $x_n \in \Sigma_\gamma M^*_\gamma$ such that $x_n - x \in p^n M^*$. We express each x_n as a sum $x_n = \Sigma_\gamma x_{\gamma n}$ with $x_{\gamma n} \in M^*_\gamma$ where $x_{\gamma n} = 0$ for all γ not in some finite set τ_n. Since x_n converges to x, the arguments of the preceding paragraph show that, for each γ, $x_{\gamma n}$ converges to some $u_\gamma \in M^*$. It is easily shown that the elements u_γ depend only on x. We set $\varphi(x) = (u_\gamma)$.

It is necessary to show that u lies in S. Consider a fixed natural number i and assume that γ is not in τ_i so that $x_{\gamma i} = 0$. Then, for $j > i$, $x_{\gamma j} = x_{\gamma j} - x_{\gamma i} \in p^i M^* \cap M^*_\gamma = p^i M^*_\gamma$. Passing to the limit we have $u_\gamma \in p^i M^*_\gamma$ because $p^i M^*_\gamma$ is closed in M^*. Since each τ_i is finite, u_γ satisfies (*) and is in S as required.

To prove φ epimorphic suppose $u \in S$. For each n let τ_n be a finite set of indices such that $u_\gamma \in p^n M^*_\gamma$ for all γ not in τ_n and let x_n be the sum (in M^*) of the u_γ for $\gamma \in \tau_n$. The existence of τ_n is insured by (*). Since $\tau_n \subseteq \tau_m$ for $m \leq n$, $x_m - x_n \in p^n M^*$. Hence the x_n converge to an element x in M^*. Moreover $x_n - x \in p^n M^*$. An examination of the definition of φ shows that $x_{\gamma n} = u_\gamma$ if $\gamma \in \tau_n$ and $x_{\gamma n} = 0$ otherwise. Hence $\varphi(x) = u$ and φ is epimorphic.

Finally suppose that $\varphi(x) = 0$. Refering to the definition of φ we have, for fixed n and all $i > n$, $(\Sigma_{\gamma i} - x_{\gamma n}) = x_i - x_n \in p^n M^*$. Since $\Sigma_\gamma M^*_\gamma$ is pure in M^* and the sum is direct, this implies that $x_{\gamma i} - x_{\gamma n} \in p^n M^*_\gamma$ for each index γ and each $i > n$. We are assuming all $u_\gamma = 0$ so that $x_{\gamma i} \in p^n M^*_\gamma$ for large i, hence $x_{\gamma n} \in p^n M^*_\gamma$. But then $x_n = \Sigma_\gamma x_{\gamma n} \in p^n M^*$ and $x_n \to 0$, $x = 0$. This shows that φ is a monomorphism and completes the proof.

LEMMA 2.3. *If $M = \Pi_\gamma M_\gamma$ where γ ranges over a set of cardinal $\aleph$ and the M_γ are all torsion-free with the same rank, then*

$$r(M) = |\, M_\gamma \,|^{\aleph} .$$

Proof. Note first that for each γ $|\, M_\gamma \,| = r(M_\gamma)\, |\, R\, |$ so that all the M_γ have the same power. If we can show that $r(M) \geq |\, R\, |$, then $r(M) = |\, M\, | = |\, M_\gamma \,|^{\aleph}$ as required.

Suppose the indices are the natural numbers and that each $M_\gamma = R$. Consideration of a suitable Vandermonde determinant shows that the elements $(1, r, r^2, \cdots) \in M$ with r ranging over R are independent so that $r(M) \geq |\, R\, |$ in this case. In the general case $\aleph$ is infinite and each M_γ contains a copy of R so that M contains a countable product of copies of R, hence $r(M) \geq |\, R\, |$ in all cases.

LEMMA 2.3. *Suppose that N is a submodule of M and that, for each natural number n, M_n and N_n are copies of M and N respectively. If $\varphi: \Pi_n M_n \to M$ is a homomorphism such that $\varphi^{-1}(N) \subseteq \Pi_n N_n$, then*

$$r(M/N) = r(M/N)^{\aleph_0} .$$

Proof. Since φ maps $\varphi^{-1}(N)$ into N, it induces a monomorphism

$$(1) \qquad 0 \to \Pi_n M_n/\varphi^{-1}(N) \to M/N .$$

Since $\varphi^{-1}(N) \subseteq \Pi_n N_n$, there is an epimorphism

$$(2) \qquad \Pi_n M_n/\varphi^{-1}(N) \to \Pi_n(M_n/N_n) \to 0 .$$

Rank does not increase on passing to submodules or to homomorphic images, hence (1) and (2) imply

$$(3) \qquad r(M/N) \geq r(\Pi_n M_n/\varphi^{-1}(N)) \geq r(\Pi_n(M_n/N_n)) .$$

By the definition of rank M/N contains a free module F such that $r(F) = r(M/N)$. For each n let F_n be a copy of F in M_n/N_n. Then $\Pi_n F_n \subseteq \Pi_n(M_n/N_n)$ and Lemma 2.2 implies

$$(4) \qquad r(\Pi_n(M_n/N_n)) \geq r(\Pi_n F_n) = |F|^{\aleph_0} \geq r(F)^{\aleph_0} = r(M/N)^{\aleph_0} .$$

Thus (3) and (4) imply the conclusion of the lemma.

THEOREM 2.4. *If M is the direct sum of $\aleph$ copies of $\Sigma_n R/p^n R$, then $r(M^*) = (\aleph\,|R|)^{\aleph_0}$.*

Proof. We first consider the case $\aleph = 1$. It will be convenient to replace $R/p^n R$ by the isomorphic module $R(p^n)$ which consists of all elements of Q/R annihilated by p^n, for then $R(p^n) \subseteq R(p^m)$ for all $m \geq n$. Each element $a \neq 0$ in $R(p^n)$ has a height $h_n(a)$ in $R(p^n)$ where $h_n(a)=i$ if $a \in p^i R(p^n)$ but a is not in $p^{i+1}R(p^n)$. The height and exponential order of a are related by $h_n(a) + e(a) = n$. We let $C = \Sigma_n R(p^n)$ and $D= \Pi_n R(p^n)$. Then C^* consists of those elements $x = (x_n) \in D$ such that $h_n(x_n)$ goes to ∞ with n.

We show first that $r(C^*) = r(D)$. The inequality $r(C^*) \leq r(D)$ holds because $C^* \subseteq D$. To prove the opposite inequality we define $\rho: D \to C^*$ by

$$\rho(x)_n = \begin{cases} 0 & \text{if} \quad n = 2k+1 , \\ x_k & \text{if} \quad n = 2k . \end{cases}$$

Since $R(p^k) \subseteq R(p^{2k})$, ρ is a homomorphism into D. Since $e(x_k) \leq k$ and $h_{2k}(x_k) + e(x_k) = 2k$, $h_{2k}(x_k) \geq k$ so that $\rho(x)$ lies in C^*. The map ρ is clearly a monomorphism so $r(D) \leq r(C^*)$ as required.

The next step is to show that

$$r(D) = r(D)^{\aleph_0} .$$

Let $\sigma_1, \sigma_2, \cdots$ be an infinite partition of the set of natural numbers into infinite subsets. For each n let D_n be a copy of D. An element $u \in \Pi_n D_n$ is a sequence $(u_1, u_2, \cdots)$ with $u_n = (u_{ni}) \in D$. We define $\xi : \Pi_n D_n \to D$ by $\xi(u)_k = u_{ni}$ if k is the ith element of σ_n; $u_{ni} \in R(p^k)$ because $k \geq i$. The hypotheses of Lemma 2.3 are satisfied with $M = D$ and $N = 0$ which shows that $r(D) = r(D)^{\aleph_0}$.

The module D can be represented as the module of all infinite sequences $(x_1, x_2, \cdots)$ of elements of R modulo the sequences of the form $(b_1p, b_2p^2, b_3p^3, \cdots)$. Thus Lemma 2.2 and the fact that rank does not increase on passing to homomorphic images imply that $r(D) \leq |R|^{\aleph_0}$. We shall show that $r(D) \geq |R|$. Then $r(D) = r(D)^{\aleph_0} \geq |R|^{\aleph_0}$ and we get

$$r(D) = |R|^{\aleph_0} .$$

To show that $r(D) \geq |R|$ let $\alpha(r) = (1, r, r^2, \cdots)$ for each $r \in R$ and let $\bar{\alpha}(r)$ be the image of $\alpha(r)$ in D. We show that the elements $\bar{\alpha}(r)$ for $r \in R - (p)$ are independent. Suppose $r_1, \cdots, r_n$ are distinct elements of R not in (p), and suppose $a_1, \cdots, a_n \in R$ such that

$$a_1\bar{\alpha}(r_1) + \cdots + a_n\bar{\alpha}(r_n) = 0 .$$

Then elements $b_1, b_2, \cdots$ exist in R such that

$$a_1\alpha(r_1) + \cdots + a_n\alpha(r_n) = (b_1p, b_2p^2, \cdots) .$$

Hence, for each k, the a_i satisfy a system of n equations

$$\begin{aligned} &a_1r_1^k + \cdots + a_nr_n^k = b_kp^k \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ &a_1r_1^{k+n-1} + \cdots + a_nr_n^{k+n-1} = b_{k+n-1}p^{k+n-1} . \end{aligned}$$

The determinant Δ of this system is $r_1^k \cdots r_n^k d$ where d is the Vandermonde determinant of $r_1, \cdots, r_n$; $d \neq 0$ because the r's are distinct. We set $d = p^m s$ with s prime to p and $t = r_1^k \cdots r_n^k s$. Then $\Delta = p^m t$ where t is prime to p because $r_1, \cdots, r_n, s \in R - (p)$. Then by Cramer's rule each a_i satisfies an equation of the form $p^m t a_i = p^k c_i$. Hence, for $k > m$, p^{k-m} divides ta_i and therefore divides a_i because it is prime to t. Since this is true for all $k > m$, $a_i = 0$ for each i. Therefore the $\bar{\alpha}(r)$ with r ranging over $R - (p)$ is an independent subset of D so $r(D) \geq |R - (p)|$. But $R - (p)$ is the disjoint union of cosets of (p) so that $|R - (p) \geq |(p)| = |R|$; hence $|R - (p)| = |R|$.

We now have $r(C^*) = r(D) = |R|^{\aleph_0}$ which completes the proof in the case $\aleph = 1$.

Now suppose $\aleph$ arbitrary, let Γ be a set with cardinal $\aleph$ and let $M = \Sigma_\gamma M_\gamma$ where, for each $\gamma \in \Gamma$, $M_\gamma = C = \Sigma_n R(p^n)$. In view of Lemma 2.1 and the remark following it M^* is contained in the submodule A of all sequences $x \in \Pi_\gamma M_\gamma^*$ with $x_\gamma = 0$ for all but a countable number of indices. Each such sequence is determined by the set σ of indices γ such that $x_\gamma \neq 0$ and a function $f: \sigma \to C^* - \{0\}$. From this it follows easily that $|A| \leq (\aleph |C^*|)^{\aleph_0}$. Since $C^* \subseteq D$ and D is a homomorphic image of the direct product of $\aleph_0$ copies of R, $|C^*| \leq |R|^{\aleph_0}$. Since $|R|^{\aleph_0} = r(C^*) \leq |C^*|$ we have $|C^*| = |R|^{\aleph_0}$. Hence

$$r(M^*) \leq r(A) \leq |A| \leq (\aleph |R|)^{\aleph_0} .$$

Using Lemma 2.1 again we have $\Sigma_\gamma M_\gamma^* \subseteq M^*$ so that

$$r(M^*) \geq r(\Sigma_\gamma M_\gamma^*) = |\Gamma| r(C^*) = \aleph |R|^{\aleph_0} .$$

These last two sets of inequalities combine to give

$$\aleph |R|^{\aleph_0} \leq r(M^*) \leq (\aleph |R|)^{\aleph_0}.$$

If $\aleph$ is finite this completes the proof. If $\aleph$ is infinite, the proof will be complete once we show that $r(M^*)^{\aleph_0} = r(M^*)$. To show this assume $\aleph$ infinite and partition the index set Γ into a countable sequence Γ_1, $\Gamma_2, \cdots$ of disjoint subsets such that $|\Gamma_n| = |\Gamma| = \aleph$ and set $M_n = \Sigma\{M_\gamma \mid \gamma \in \Gamma_n\}$. Then $M_n \approx M$ and $M_n^* \approx M^*$ for each n. Our purpose will be achieved if we can define a monomorphism $\varphi : \Pi_n M_n^* \to M^*$, for then $\varphi^{-1}(tM^*) = t(\Pi_n M_n^*) \subseteq \Pi_n tM_n^*$, where tM^* is the torsion submodule of M^*. Now Lemma 2.3 applies to give $r(M^*/tM^*) = r(M^*/tM^*)^{\aleph_0}$. But $r(M^*) = r(M^*/tM^*)$ so $r(M^*) = r(M^*)^{\aleph_0}$.

Earlier in the proof of this theorem we defined a monomorphism $\rho : D \to C^*$. For each k we now define a monomorphism $\psi_k : D \to D$ by

$$\psi_k(x)_i = \begin{cases} 0, & i \leq k , \\ x_{i-k}, & i > k . \end{cases}$$

For $i > k$ we have $e(x_{i-n}) \leq i - k$ so that $h_i(x_{i-k}) = i - e(x_{i-k}) \geq k$. Hence $\psi_k(D) \subseteq p^k D$ so that $\rho\psi_k$ maps D into $p^k C^*$. We define φ_k: $C^* \to p^k C^*$ to be the restriction of $\rho\psi_k$ to C^* and note that it is a monomorpoism.

We now use Lemma 2.1 to identify M^* with the submodule of $\Pi_\gamma M_\gamma^*$ described by the condition (*). An element x of $\Pi_n M_n^*$ is a sequence $(x_1, x_2, \cdots)$ where $x_n \in M_n^* \subseteq \Pi\{M_\gamma^* \mid \gamma \in \Gamma_n\}$. We define φ by $\varphi(x)_\gamma = \varphi_n(x_{n\gamma})$ for $\gamma \in \Gamma_n$. Then $\varphi : \Pi_n M_n^* \to \Pi_\gamma M_\gamma^*$ and is a monomorphism because each φ_n is one. There remains the task of showing that $\varphi(x)$ lies in M^*. Let n be a natural number. For each $k < n$ there is by Lemma 2.1 a finite subset τ_k of Γ_k such that $x_{k\gamma} \in p^n M_\gamma^*$ for $\gamma \in \Gamma_k$

but not in τ_k. By the definition of φ_k, $\varphi_k(x_{k\gamma}) \in p^n M_\gamma^*$ for all $\gamma \in \Gamma_k$ with $k \geq n$. Hence $\varphi(x)_\gamma \in p^n M_\gamma^*$ for all not in $\tau_1 \cup \cdots \cup \tau_{n-1}$ which is a finite set. Thus $\varphi(x)$ satisfies (*) of Lemma 2.1 and is in M^* as required.

3. Let R once more be an arbitrary Dedekind ring and let P be a prime ideal of R. For any R-module T, $\operatorname{Ext}(Q, T)$ is a vector space over Q and is therefore completely described by its dimension over Q or equivalently its rank over R. According to Theorem 1.4 this dimension is a function of the critical number of T if T is primary.

THEOREM 3.1. *If T is a P-primary R-module with infinite critical number $\aleph$, then the rank of* $\operatorname{Ext}(Q, T)$ is $(\aleph \mid R \mid)^{\aleph_0}$.

Proof. In order to make the results of section two available we change rings. The module T, being P-primary, can be considered as a module over the ring S consisting of all elements of the form a/b in Q with a and b in R and b prime to P. The theory of P-primary modules is left unchanged by the shift from R to S. In particular the critical number of T is $\aleph$ in both cases.

Since S is torsion-free as an R-module Proposition 4.1.3. of [1] applies to give a natural isomorphism

$$\operatorname{Ext}_R(Q, T) \approx \operatorname{Ext}_S(S \otimes_R Q, T) .$$

Since R and S have the same quotient field Q, $Q = S \otimes_R Q$ and

$$\operatorname{Ext}_R(Q, T) \approx \operatorname{Ext}_S(Q, T) .$$

These are both vector spaces over Q and the isomorphism is a Q-isomorphism; hence the two modules have the same dimension over Q. Let M be the direct sum of $\aleph$ copies of $\Sigma_n S/p^n S$ where p is the prime of S. According to Theorem 1.4

$$\operatorname{Ext}_S(Q, T) \approx \operatorname{Ext}_S(Q, M) .$$

Since M is a basic submodule of tM^*, Theorem 1.3 gives

$$\operatorname{Ext}_S(Q, M) \approx \operatorname{Ext}_S(Q, tM^*) .$$

By Theorem 7.4 of [3], $\operatorname{Ext}_S(Q, M^*) = 0$ because M^* is complete, while $\operatorname{Hom}_S(Q, M^*) = 0$ because M^* is reduced. Hence the second exact sequence associated with Q and $0 \to tM^* \to M^* \to M^*/tM^* \to 0$ reduces to

$$0 \to \operatorname{Hom}_S(Q, M^*/tM^*) \to \operatorname{Ext}_S(Q, tM^*) \to 0 .$$

Since M^*/tM^* is torsion-free divisible

$$\operatorname{Hom}_S(Q, M^*/tM^*) \approx M^*/tM^* .$$

It follows that $\mathrm{Ext}_R(Q, T)$ and M^*/tM^* have the same dimension over Q. This dimension is $(\aleph \mid S \mid)^{\aleph_0}$ by Theorem 2.5. Moreover $\mid R \mid = \mid S \mid$. Hence the theorem is proved.

Since the integers are the most important example of a Dedekind ring it is appropriate to interpret the last theorem for this special case. Since rank and cardinality coincide for torsion-free abelian groups of infinite rank, we can say that *if T is a p-primary abelian group with infinite critical number $\aleph$, there are $\aleph^{\aleph_0}$ inequivalent extensions of T by the rational numbers.*

REFERENCES

1. H. Cartan and S. Eilenberg, *Homological Algebra*, Princeton 1956.

2. I Kaplansky, *Infinite Abelian Groups*, Ann Arbor, 1954.

3. R. J. Nunke, *Modules of extensions over Dedekind rings*. Ill. Journ. of Math., **3**, (1959), 222–241.

4. T. Szele, *On the basic subgroups of abelian p-groups*, Acta. Math. Acad. Sci. Hung., **5** (1954), 129-141.

UNIVERSITY OF WASHINGTON

MIXED MODULES OVER VALUATION RINGS

JOSEPH ROTMAN

1. Introduction. A p-primary abelian group is a module over the p-adic integers; thus Ulm's theorem can be viewed as a classification of reduced countably generated torsion modules over the p-adic integers, or, more generally, over a complete discrete valuation ring. It is with this point of view that Kaplansky and Mackey [4] generalized Ulm's theorem to cover mixed modules of rank 1. In this paper their result is generalized in various ways, sometimes to modules of finite rank, sometimes to modules over possibly incomplete rings. The structure theorems obtained are applied to solve square-root, cancellation, and direct summand problems.

The main idea is to squeeze as much information as possible from the proof of Ulm's theorem in [4]. In order to understand our procedure, we sketch that proof. Order, once for all, generating sets of the modules T and T': $t_1, t_2, \cdots; t'_1, t'_2, \cdots$ The plan is to build an isomorphism stepwise up these lists. The crucial point is then, given a height-preserving isomorphism $f: S \to S'$, S finitely generated, to extend f to a height-preserving isomorphism of $\{t_i, S\}$ and a suitable submodule of T' containing S'. In order to construct this extension it is necessary to normalize t_i in two ways:

(i) assume $pt_i \in S$;

(ii) assume that t_i has maximal height in the coset $t_i + S$. If T is torsion, both of these normalizations are always possible. Now the possibility of extra generality arises precisely at these two points. If T is mixed and (ii) is satisfied, then the proof will go through if T/S is torsion; this is what Kaplansky and Mackey did in their paper. In this paper, we define a class of modules in which (ii) can always be satisfied, and it is this class of modules which we shall consider.

2. Definitions. A *discrete valuation ring* (DVR) is a principal ideal domain R with a unique prime ideal (p). $\bigcap_{n=1}^{\infty}(p^n) = (0)$. Hence if $r \in R$ is non-zero, there is a maximal n, depending on r, such that $r \in (p^n)$. Define $|r| = e^{-n}$; define $|0| = 0$. $|\ |$ is a norm which satisfies the strong triangle inequality: $|r + r'| \leq \max |r|, |r'|$. This norm induces a metric on R. R is a *complete* DVR if it is complete in this metric. If R is incomplete, we may form its completion R^*, and R^* is a complete DVR. The p-adic integers is a complete DVR; it is also compact as a metric space.

Let Q be the quotient field of R. We define the *rank* of a module M (often called the 'torsion-free rank') to be the dimension of the Q

Received June. 22, 1959.

vector space $Q \otimes M$. Thus if M is torsion, rank $M = 0$. The rank can also be defined as the cardinality of a maximal independent subset of M. Note that every element in an independent subset has infinite order.

The word module will mean unitary module over a DVR. All abelian group-theoretic notions can be found in [2; 3; 5].

3. *KM* modules. In this section we shall define a certain family of modules and determine some members of this family.

DEFINITION. A *semi-KM module* is a reduced countably generated module of finite rank.

DEFINITION. A module M *has the coset property* if the coset $x+S$ has an element of maximal height whenever S is a finitely generated submodule of M.

DEFINITION. A *KM module* is a semi-*KM* module with the coset property.

The coset property is the crucial part of the definition of a *KM* module; for later use, we now give a characterization of this property.

DEFINITION. Let S be a submodule of M; if $x \in M$, let x^* denote the image of x in M/S under the natural homomorphism. S is *copure* if any $x^* \in M/S$ has a pre-image x such that $h(x^*) = h(x)$. ($h(x)$ denotes the height of the element x).

LEMMA 3.1. *S is copure in M if and only if every coset of S has an element of maximal height.*

Proof. Induction on $h(x)$ that $h(x) = h(x^*)$ if x has maximal height in $x + S$.

COROLLARY 3.2. *M has the coset property if and only if every finitely generated submodule is copure.*

LEMMA 3.3. *If R is complete, a reduced module M with no elements of infinite height has the coset property.*

Proof. Let $S = \{y_1, \cdots, y_s\}$. It must be shown that $x + S$ contains an element of maximal height. We may assume that $x \notin S$, otherwise 0 has maximal height in $x + S$. Under this assumption we show by induction on s that $y + S$ contains only finitely many distinct heights.

Let $s = 1$. If $h(x + a_n y) = \alpha_n$ is strictly increasing, then $h(b_n y) = \alpha_n$, where $b_n = a_{n+1} - a_n$. Hence $h(b_{n+1}y) > h(b_n y)$. Let $(p^{m(n)})$ be the smallest ideal containing b_n. Then $m(n+1) > m(n)$, i.e., $m(n) \to \infty$, and so $b_n \to 0$. Hence $\{a_n\}$ is a Cauchy sequence and $a_n \to a$, since R is complete. Now $x + ay = x + a_n y + (a - a_n)y$. If $h((a - a_n)y) \geq \alpha_n$ for all n, then $x + ay$ has infinite height and is thus 0, contradicting $x \notin S$. Therefore we may assume that $h((a - a_n)y) = h((a - a_m)y)$ for all $m \geq n$. But then $a - a_n$ and $a - a_m$ are associates, contradicting $a - a_m \to 0$. Hence $\{\alpha_n\}$ cannot be strictly increasing, i.e., there can only be a finite number of heights in the coset.

For the general case, suppose $h(x + a_1^n y_1 + \cdots + a_s^n y_s) = \alpha_n$ is strictly increasing. Suppose further that each coordinate sequence $\{a_i^n\}$ is Cauchy, and so $a_i^n \to a_i$ for each i. Then

$$x + a_1 y_1 + \cdots + a_s y_s = (x + a_1^n y_1 + \cdots + a_s^n y_s) + (a_1 - a_1^n)y_1 + \cdots + (a_s - a_s^n)y_s .$$

The height of the first term on the right is α_n while the height of the remaining terms gets arbitrarily large. Hence $x + a_1 y_1 + \cdots + a_s y_s$ has infinite height and so must be 0, contradicting $x \notin S$. Hence $\{\alpha_n\}$ cannot be strictly increasing, i.e., there are only a finite number of heights.

Therefore we may assume $\{a_1^n\}$ contains no Cauchy subsequence, and so we may assume further that it consists of incongruent units. Now

$$h(a_1^{n+1}(x + \Sigma a_j^n y_j) - a_1^n(x + \Sigma a_j^{n+1} y_j)) = \alpha_n = h((a_1^{n+1} - a_1^n)x + \Sigma b_k^n y_k) ,$$

where

$$b_k^n = a_1^{n+1} a_k^n - a_1^n a_k^{n+1}$$

and $k \geq 2$. Since $a_1^{n+1} - a_1^n$ is a unit, and since multiplication by a unit does not alter heights, we may assume it is 1. But there are only $s - 1$ y's occurring, and so the inductive hypothesis applies. Hence there can only be a finite number of heights, and so $\{\alpha_n\}$ cannot be strictly increasing. Thus $x + S$ contains only finitely many distinct heights.

LEMMA 3.4. *If R is compact and M is a reduced module of rank 2, then M has the coset property.*

Proof. Let S be a finitely generated submodule with $x \notin S$. By the method of [4], it suffices to consider the case when S is generated by two elements of infinite order, y and z. Moreover, we may assume $h(x + a_n y + b_n z) = \alpha_n$, where $\{\alpha_n\}$ is strictly increasing. Since R is compact, we may assume that $a_n \to a$ and $b_n \to b$. $x + ay + bz = (x + a_n y + b_n z) + ((a - a_n)y + (b - b_n)z)$. Now the height of the first term on the right is α_n. If the other term has height $\geq \alpha_n$ for all n, then $h(x + ay + bz) \geq \alpha_n$ for all n and $x + ay + bz$ is the desired element. Hence we may suppose that $h((a - a_n)y + (b - b_n)z) = \beta < \alpha_n$. This equation must hold for all $m \geq n$. If a sequence $\{c_n\}$ converges to c, there is a subsequence $\{c_{n_i}\}$ such that $c - c_{n_i}$ and $c_{n_{i+1}} - c_{n_i}$ are associates, In our case, there are units u_n and v_n such that $(a - a_n)y = u_n(a_{n+1} - a_n)y$ and $(b - b_n)z = v_n(b_{n+1} - b_n)z$. (We have assumed, for notation, that $\{a_n\}$ and $\{b_n\}$ are the subsequences). Hence

$$(a - a_n)y + (b - b_n)z = u_n(x + a_{n+1}y + b_{n+1}z)$$
$$- u_n(x + a_n y + b_n z) + (v_n - u_n)(b_{n+1} - b_n)z \, .$$

Hence $h((v_n - u_n)(b_{n+1} - b_n)z) = \beta$ for large n. Therefore, $(v_n - u_n)(b_{n+1} - b_n)$ are associates, and non-zero since $\beta < \alpha_n < \infty$. Hence there must be a maximal power of p dividing any of them, contradicting the fact that $b_{n+1} - b_n \rightarrow 0$.

LEMMA 3.5. *If R is complete and M is reduced of rank* 1, *then M has the coset property.*

Proof. Kaplansky and Mackey [4].

To this point, all modules with the coset property have been modules over a complete DVR. We shall now exhibit modules over a possibly incomplete ring which have the coset property. For this purpose we consider tensor products. All tensor products will be taken over the ring R.

LEMMA 3.6. *Let R be a* DVR *with completion R^*. Any R-module M can be imbedded as a pure R-submodule in $R^* \otimes M$; moreover, the torsion submodule T of M coincides with $R^* \otimes T$, which is the torsion submodule of $R^* \otimes M$.*

Proof. R^* is a torsion-free R-module, and R is a pure submodule [3]. Further, if $\delta + R \in R^*/R$, there is an $r \in R$ such that $\delta - r = p\delta'$, $\delta' \in R^*$. Therefore $\delta + R = p\delta' + R$ and so $p(R^*/R) = R^*/R$. Hence R^*/R is torsion-free and divisible.

Exactness of the sequence $0 \rightarrow R \rightarrow R^* \rightarrow R^*/R \rightarrow 0$ induces exactness of $\operatorname{Tor}(R^*/R, M) \rightarrow R \otimes M \rightarrow R^* \otimes M \rightarrow (R^*/R) \otimes M \rightarrow 0$. $R \otimes M = M$ and, since R^*/R is torsion-free, $\operatorname{Tor}(R^*/R, M) = 0$. Thus $x \rightarrow 1 \otimes x$ is an imbedding of M into $R^* \otimes M$. But the sequence also implies that $(R^* \otimes M)/M \approx (R^*/R) \otimes M$. Since R^*/R is torsion-free and divisible, we have $(R^*/R) \otimes M$ torsion-free. Hence M is pure in $R^* \otimes M$ and contains the torsion submodule of $R^* \otimes M$. We already know that $x \to 1 \otimes x$ is a monomorphism; this last remark shows it is an epimorphism when restricted to T. Thus $T \approx R^* \otimes T$, which is the torsion submodule of $R^* \otimes M$.

LEMMA 3.7. *If R is a* DVR *with completion R^*, and if M is an R-module of rank* 1 *with no elements of infinite height, then $R^* \otimes M$ has no elements of infinite height.*

Proof. Suppose $z = \Sigma\delta_i \otimes m_i \in R^* \otimes M$ has infinite height. By the preceding lemma, z has infinite order. Let $x \in M$ have infinite order. Since rank $M = 1$, there is an n such that for all i, $p^n m_i = r_i x$, $r_i \in R$. $p^n z = \Sigma\delta_i r_i \otimes x$. As any element in R^*, $\Sigma\delta_i r_i$ can be expressed as γp^k, where γ is a unit. But then $h(z) < h(p^n z) = h(\gamma \otimes p^k x) = h(1 \otimes p^k x) = h(p^k x)$ which is finite. This contradiction completes the proof.

LEMMA 3.8. *If M is an R-module of rank* 1 *with no elements of infinite height, then M has the coset property.*

Proof. Let S be a finitely generated submodule of M, and let $x \notin S$. Then $R^* \otimes S$ is a finitely generated R^*-submodule of $R^* \otimes M$. We now show $1 \otimes x \notin R^* \otimes S$. Consider the following commutative diagram with exact rows:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & S & \longrightarrow & M & \xrightarrow{\alpha} & M/S & \longrightarrow & 0 \\ & & \downarrow & & \downarrow{\scriptstyle i} & & \downarrow{\scriptstyle j} & & \\ 0 & \longrightarrow & R^* \otimes S & \longrightarrow & R^* \otimes M & \xrightarrow{\beta} & R^* \otimes (M/S) & \longrightarrow & 0 \end{array}$$

where the downward maps are $y \to 1 \otimes y$. Then

$$\beta(1 \otimes x) = \beta i(x) = j\alpha(x) = j(x + S) .$$

But

$$\gamma\colon (R^* \otimes M)/(R^* \otimes S) \longrightarrow R^* \otimes (M/S)$$

defined by

$$\gamma(r^* \otimes m + R^* \otimes S) = \beta(r^* \otimes m)$$

is an isomorphism. In particular,

$$\gamma(1 \otimes x + R^* \otimes S) = \beta(1 \otimes x) = j(x + S) .$$

Since $x \notin S, x + S \neq 0$. Since j is a monomorphism, by Lemma 3.6, $j(x + S) \neq 0$. Therefore $1 \otimes x + R^* \otimes S \neq 0$, i.e., $1 \otimes x \notin R^* \otimes S$. Hence $1 \otimes x + R^* \otimes S$ contains only finitely many distinct heights, by Lemma 3.3. Therefore the pure subset $x + S$ of $1 \otimes x + R^* \otimes S$ car contain only finitely many distinct heights, and so it has an element of maximal height.

We now sum up the results of this section in the following theorem.

THEOREM 3.9. *A semi-KM module is a KM module if any of the following conditions hold*:

(i) *R is complete and M has no elements of infinite height*;

(ii) *R is compact and rank $M = 2$*;

(iii) *R is complete and rank $M = 1$*;

(iv) *rank $M = 1$ and M has no elements of infinite height.*

It is an open question whether these are all the semi-*KM* modules with the coset property. Later we shall give an example of a module of rank 2 with no elements of infinite height over an incomplete ring which does not have the coset property.

4. The Structure theorem. The main result of this section is the classification of all *KM* modules.

DEFINITION. A *strand* is a function from the cartesian product of s copies of R into the ordinals and the symbol ∞, where R is a DVR and s is finite.

DEFINITION Two strands f and $g: R \times \cdots \times R \to$ ordinals and ∞ are *equivalent*, denoted $f \equiv g$, in case there is an s by s non-singular matrix A over R and non-negative integers m and n such that $g(p^{m+n}(r_1, \cdots, r_s)) = f(p^n(r_1, \cdots, r_s)A)$ for all $r_i \in R$. The argument of f is obtained by regarding $(r_1, \cdots, r_s)$ as a 1 by s matrix.

It is easy to verify that $f \equiv g$ is an equivalence relation. If M is a reduced module of finite rank s, then any ordered independent set of elements $x_1, \cdots, x_s$ determines a strand f by $f(r_1, \cdots, r_s) = h(\Sigma r_i x_i)$. f is *the strand determined by the* x's. It is straightforward to see that two strands determined by different ordered maximal independent subsets of M are equivalent in the above sense. Thus M determines an equivalence class of strands, which we denote $S(M)$. Clearly $S(M)$ is an invariant of M.

LEMMA 4.1. *Let M and M' be KM modules. Let S and S' be finitely generated submodules of M and M' respectively, let f be a height-preserving isomorphism of S onto S', and let $x \in M$ with $px \in S$. Then f can be extended to a height-preserving isomorphism between $\{x, S\}$ and a suitable submodule of M' which contains S'.*

Proof. Exactly as in [4].

LEMMA 4.2. *Let M and M' be KM modules with $S(M) = S(M')$. Then there are maximal independent subsets in M and in M' which determine the same strand.*

Proof. Let $y_1, \cdots, y_s$ be independent in M with strand f; let $y'_1, \cdots, y'_s$ be independent in M' with strand g. Since $S(M) = S(M')$, $f \equiv g$. Hence there are non-negative integers m and n and a non-singular matrix (a_{ij}) over R such that

$$g(p^{m+n}(r_1, \cdots, r_s)) = f(p^n(r_1, \cdots, r_s)(a_{ij})) ,$$

i.e.,

$$h(p^{m+n}\Sigma r_i y'_i) = h(p^n \Sigma\Sigma r_i a_{ij} y_j) .$$

Set $x_i = p^n \Sigma a_{ij} y_j$ and set $x'_i = p^{m+n} y'_i$.

THEOREM 4.3. *Let M and M' be KM modules. M and M' are isomorphic if and only if they have the same Ulm invariants and $S(M) = S(M')$.*

Proof. By Lemma 4.2, there are maximal independent subsets $x_1, \cdots, x_s$ in M, $x'_1, \cdots, x'_s$ in M' such that $h(\Sigma r_i x_i) = h(\Sigma r_i x'_i)$ for all $r_i \in R$. Let S be the submodule of M generated by the x's and let S' be the submodule of M' generated by the x''s. Define $f: S \to S'$ by $f(x_i) = x'_i$. Since S and S' are free on generators x_i, respectively x'_i, f is a well-defined isomorphism. Moreover, our choice of generators makes f height-preserving. This isomorphism is now extended stepwise to an isomorphism of M and M' by Lemma 4.1. To ensure catching all of M and M', we take fixed countable sets of generators for each and alternate between adjoining an element of M and an element of M'. Since the elements of M and M' have finite order modulo S and S' respectively, we can suppose that at each step we are adjoining an element x such that px lies in the preceding submodule. This is precisely the situation of Lemma 4.1.

COROLLARY 4.4. *Let M and M' be isomorphic KM modules. Then any height-preserving isomorphism between finitely generated submodules S and S' of M and M' respectively* (*rank S = rank M*) *can be extended to an isomorphism of M with M'.*

As first applications of the structure theorem, we now solve a square-root problem and a cancellation problem.

THEOREM 4.5. *Let M and M' be KM modules of rank 1 with $M \oplus M \approx M' \oplus M'$. Then $M \approx M'$.*

Proof. It is a corollary of Ulm's theorem that the above is true when M and M' are torsion. Hence the torsion submodules of M and M' are isomorphic. Let $x \in M$ have infinite order. Then there is an element $(a, b) \in M' \oplus M'$ such that $h(rx) = h(ra, rb)$ for all $r \in R$. Since rank $M' = 1$, there are non-negative integers m and n such that $p^m a = p^n ub = y$, where u is a unit in R. We assume $m \geq n$. Thus, for large k, we have $h(p^k x) = h((p^{k-m}y, p^{k-n}u^{-1}y)) = h(p^{k-m}y)$. Hence $S(M) = S(M')$. Therefore, $M \approx M'$ by the structure theorem.

I have been unable to prove the analogous result in the case of higher rank, and I conjecture it is false.

THEOREM 4.6. *Let M and M' be KM modules, and let T be a reduced countably generated torsion module such that $U_\alpha(T)$ is finite for all α, where $U_\alpha(T)$ is the αth Ulm invariant of T. Then $T \oplus M \approx T \oplus M'$ implies $M \approx M'$.*

Proof.

$$S(M) = S(T \oplus M) = S(T \oplus M') = S(M') .$$

By Ulm's Theorem, we may cancel T to obtain that the torsion submodules of M and M' are isomorphic. By the structure theorem, $M \approx M'$.

$S(M)$ is a rather cumbersome invariant. We make the following definition in order to rephrase Theorem 4.4.

DEFINITION. Two modules M and M' are *almost isomorphic* if there exist torsion modules T and T' such that $T \oplus M \approx T' \oplus M'$.

THEOREM 4.7. *Two KM modules M and M' are isomorphic if and only if they are almost isomorphic and they have the same Ulm invariants.*

Proof. The necessity is obvious. For sufficiency, note that if M and M' are almost isomorphic, then $S(M) = S(M')$. Since M and M' have the same Ulm invariants, $M \approx M'$ by 4.3.

5. Modules over incomplete rings. At present we have a structure theorem for KM modules, and the only KM modules over incomplete rings that we know are those of rank 1 with no elements of infinite height. In Lemmas 3.6, 3.7, and 3.8, however, we saw that we could obtain information about a module M by examining $R^* \otimes M$, which we henceforth denote M^*. We now investigate this situation more closely.

LEMMA 5.1. *The rank of M as an R-module = the rank of M^* as an R^*-module.*

Proof. Rank $M \geq \text{rank } M^*$, for if $x_1, \cdots, x_s$ is a maximal independent subset of M, then $1 \otimes x_1, \cdots, 1 \otimes x_s$ is a maximal independent subset of M^*. For the other inequality, let S be a free submodule of M with rank $S = \text{rank } M$. Since R^* is torsion-free, exactness of $0 \to S \to M$ implies exactness of $0 \to S^* \to M^*$. Since tensor product commutes with direct sums, rank $M \leq \text{rank } M^*$

LEMMA 5.2. *Let W be an R^*-module of finite rank s, with torsion submodule T. Let M and M' be R-modules of rank s contained in W satisfying*:

(i) $T \subset M \cap M'$;

(ii) *there is an independent subset $x_1, \cdots, x_s$ in $M \cap M'$;*

(iii) *if f is the strand determined by the x's in M, and if g is the strand they determine in M', then $f = g$. Under these conditions, $M = M'$.*

Proof. Let $x \in M$. Since rank $W = s$, $p^k x = \Sigma c_i x_i$, $k \geq 0$, and $c_i \in R^*$. But each $c_i \in R$, lest $\Sigma c_i x_i, x_1, \cdots, x_s$ are $s+1$ independent (over R) element in M, contradicting rank $M = s$. Hence $p^k x \in M \cap M'$. In M, $h(\Sigma c_i x_i) \geq k$. By (iii), $h(\Sigma c_i x_i) \geq k$ in M' Thus there is a $y \in M'$ such that $p^k y = \Sigma c_i x_i$. Hence $p^k(x - y) = 0$, and so $x - y \in T$. Thus $x = y + (x - y) \in M'$. The other inclusion is proved similarly.

LEMMA 5.3. *Let M and M' be reduced R-modules; let $x_1, \cdots, x_s$ be a maximal independent subset in M, $x'_1, \cdots, x'_s$ a maximal independent subset of M' such that $h(\Sigma r_i x_i) = h(\Sigma r_i x'_i)$ for all $r_i \in R$. If $c_i \in R^*$, then $h(\Sigma c_i \otimes x_i) = h(\Sigma c_i \otimes x'_i)$ if either is finite; also, if one of these heights is infinite, so is the other.*

Proof. We shall be done if we can prove $h(\Sigma c_i \otimes x_i) \geq k$ implies $h(\Sigma c_i \otimes x'_i) \geq k$, for any finite k. Choose $r_i \in R$ such that $c_i - r_i \in p^k R^*$. Then $\Sigma c_i \otimes x_i = \Sigma(c_i - r_i) \otimes x_i + \Sigma r_i \otimes x_i$ Hence $h(\Sigma r_i \otimes x_i) \geq k$. By Lemma 3.6, $h(\Sigma r_i \otimes x_i) = h(\Sigma r_i x_i) = h(\Sigma r_i x'_i) = h(\Sigma r_i \otimes x'_i)$. Hence $h(\Sigma r_i \otimes x'_i) \geq k$. But $\Sigma c_i \otimes x'_i = \Sigma(c_i - r_i) \otimes x'_i + \Sigma r_i \otimes x'_i$. Thus $h(\Sigma c_i \otimes x'_i) \geq k$.

DEFINITION. Let M be a module with no elements of infinite height. M is *taut* if length $M = \text{length } M^*$; otherwise M is *slack*.

Note that length ([3, page 26]) may be defined for not necessarily reduced modules. Thus M is taut if and only if the reduced part of M^* has no elements of infinite height. It is an open question whether slack modules exist; it is easy, however, to give an example in which M has no elements of infinite height while M^* has a proper divisible submodule. Let M be an indecomposable torsion-free R-module of rank 2 of the type exhibited in [3, page 46]. M is reduced (and so has no elements of infinite height, being torsion-free), but $M^* \approx R^* \oplus Q^*$, Q^* being the quotient field of R^*.

LEMMA 5.4. *Any direct sum of taut modules is taut.*

LEMMA 5.5. *Any module of rank* 1 *with no elements of infinite height is taut.*

Proof. Lemma 3.7.

DEFINITION. A module is *completely decomposable* if it is the direct sum of modules of rank 1.

COROLLARY 5.6. *Any completely decomposable module with no elements of infinite height is taut.*

LEMMA 5.7. *Any reduced torsion-free module is taut.*

Proof. There is a unique solution to the equation $py = x$.

THEOREM 5.8. *Let M and M' be taut semi-KM modules. Then M and M' are isomorphic if and only if they have the same Ulm invariants and $S(M) = S(M')$.*

Proof. Since M and M' have isomorphic torsion submodules, so do M^* and M'^*, by 3.6. By 4.2 there are maximal independent subsets $x_1, \cdots, x_s$ in M, $x'_1, \cdots, x'_s$ in M' such that $h(\Sigma r_i x_i) = h(\Sigma r_i x'_i)$ for all $r_i \in R$. Since both M and M' are taut, the reduced parts of M^* and M'^* have no elements of infinite height. By 5.3, $\{1 \otimes x_i\}$ and $\{1 \otimes x'_i\}$ determine the same strand. In particular, the divisible parts of M^* and M'^* have the same rank and hence are isomorphic, since they are torsion-free. By 4.3, $M^* \approx M'^*$. By Corollary 4.4, there is an isomorphism $f\colon M^* \to M'^*$ such that $f(1 \otimes x_i) = 1 \otimes x'_i$ for all i. But now M' and $f(M)$ satisfy the conditions of 5.2. Hence $M' = f(M)$. Thus M and M' are isomorphic.

Theorem 5.8 suggests that taut modules have the coset property. We now exhibit a counter-example.

EXAMPLE 5.9. There exist taut modules which do not have the coset property.

Proof. Let M be an indecomposable torsion-free R-module of rank

2, where R is (necessarily) incomplete. M is taut, by 5.7. Let S be a pure submodule of rank 1. Since M is reduced, S must be cyclic. Further, $M/S \approx Q$. Thus S cannot be copure. Hence M does not have the coset property, by 3.2.

6. Completely decomposable modules. We begin this section with the study of the simplest completely decomposable modules: those of rank 1. We have already seen that if we assume no elements of infinite height, modules of rank 1 are taut. Using results of the last section, we can now prove a cancellation law.

THEOREM 6.1. *Let M and M' be semi-KM modules of rank 1 with no elements of infinite height. Then $M \approx M'$ if and only if $M^* \approx M'^*$.*

Proof. By 3.6. $M^* \approx M'^*$ implies that the torsion submodules of M and M' are isomorphic. If x has infinite order in M, x' has infinite order in M', then the strands determined by $1 \otimes x$ and $1 \otimes x'$ are equivalent. But equivalence for modules of rank 1 is via two non-negative integers and a one-by-one matrix over R^*, i.e., an element of R^*. But any element of R^* has the form up^k where u is a unit. Since multiplication by a unit does not alter heights, we may assume that the one-by-one matrix lies in R. But then we are calculating equivalence over R. The purity of the imbedding of M into M^* yields $S(M) = S(M')$. Hence $M \approx M'$.

If rank $M = 1$, then $S(M)$ has a representative f: $R \to$ ordinals and ∞, where $f(r) = h(rx)$ for some element x of infinite order. But we know that if r and r' are associates in R, then $f(r) = f(r')$. Hence f is completely determined by its values at $p^k, k = 0, 1, 2, \cdots$ Thus $S(M)$ can be looked upon as an equivalence class of sequences of ordinals. Indeed, these ordinal sequences are the extra invariant Kaplansky and Mackey discovered in their paper.

DEFINITION. A sequence of ordinals $\{\alpha_n\}$ *has a gap at* α_n if $\alpha_{n+1} > 1 + \alpha_n$.

LEMMA (Kaplansky). *If $\{\alpha_n\}$ is the Ulm sequence of x* ([3, page 57]), *and if $\{\alpha_n\}$ has a gap at α_n, then the α_nth Ulm invariant of $M \neq 0$.*

Proof. Since $h(p^n x) = \alpha_n$ and $h(p^{n+1}x) = \alpha_{n+1} > 1 + \alpha_n$, there is a $y \in M$ such that $h(p^n y) \geq \alpha_n$ and $p^{n+2}y = p^{n+1}x$. Set $t = p^{n+1}y - p^n x$. Then t has order p and height α_n. Thus the α_nth Ulm invariant of M is non-zero.

Suppose we are given a monotone increasing sequence of non-negative integers and a torsion module T. Is there a module of rank 1 possessing these as invariants? Kaplansky's lemma provides a link between these two objects, and the following theorem shows it is the only restriction.

THEOREM 6.2. *Let T be a countaly generated torsion module with no elements of infinite height; let $\{\alpha_n\}$ be a strictly increasing sequence of non-negative integers such that $\{\alpha_n\}$ has a gap at α_n implies $U_{\alpha_n}(T)$ is non-zero. Then there exists a KM module M of rank 1 whose torsion submodule is isomorphic to T and such that $S(M)$ is the equivalence class of $\{\alpha_n\}$.*

Proof. In this proof we often denote p^k by $\exp k$. If $\{\alpha_n\}$ has only a finite number of gaps, equivalence allows us to assume that $\alpha_n = n$ for all n. Then $M = T \oplus R$ is the desired module. Therefore we may assume $\{\alpha_n\}$ has an infinite number of gaps. Let $\{\alpha_{n_i}\}$ be the subsequence of gaps. The conditions on T imply T is the direct sum of cyclic modules. Further the compatibility condition tells us that T has a cyclic summand C_i of order $(\exp(\alpha_{n_i} + 1))$; let a_i be a generator of C_i. There is a B such that $T \approx B \oplus \Sigma C_i$. We first construct a certain submodule M' of ΠC_i.

Define $x = \{u_i a_i\}$ where $u_i = \exp(\alpha_{n_i} - n_i)$. x has infinite order; for $p^m x = 0 \Longleftrightarrow p^m u_i a_i = 0$ for all $i \Longleftrightarrow \exp(m + \alpha_{n_i} - n_i) a_i = 0$ for all $i \Longleftrightarrow m + \alpha_{n_i} - n_i \geq \alpha_{n_i} + 1$ for all $i \Longleftrightarrow m \geq n_i + 1$ for all i. This is impossible since $n_i \to \infty$. We claim that if $p^k u_i a_i \neq 0$, then $p^k u_i \in (\exp \alpha_k)$. In other words, if $k + \alpha_{n_i} - n_i < \alpha_{n_i} + 1$, then $k + \alpha_{n_i} - n_i \geq \alpha_k$. Equivalently, if $n_i \geq k$, then $\alpha_{n_i} - \alpha_k \geq n_i - k$. But $\alpha_{n_i} - \alpha_k = (\alpha_{n_i} - \alpha_{n_i - 1}) + \cdots + (\alpha_{k+1} - \alpha_k) \geq n_i - k$. Thus for each k we may define an element x_k with the property that $(\exp \alpha_k) x_k = p^k x$: set $x_k = \{u_i^k a_i\}$, where $u_i^k = 0$ if $k \geq n_i + 1$, while $u_i^k = \exp(-\alpha_k + k + \alpha_{n_i} - n_i)$ otherwise.

Set $M' =$ the submodule of ΠC_i generated by the x_k's. Note that $h(p^k x) = \alpha_k$ in M'. It can be no greater, since the height of an element of ΠC_i is the smallest power of p which occurs in one of its coordinates. Hence $h(p^k x) = \alpha_k$ in ΠC_i, and so can be no larger in the submodule M'.

We still must determine the torsion submodule T' of M'. Given any two x_k's, multiplication of each by a suitable power of p makes their coordinates equal from some point on. Hence any element of finite order in M' cannot have an infinite number of non-zero coordinates. But it may be verified that for all i,

$$a_i = \exp(\alpha_{n_{i+1}} - \alpha_{n_i} - n_{i+1} + n_i) x_{n_{i+1}} - x_{n_i} .$$

Hence $T' = \Sigma C_i$. Thus $M = M' \oplus B$ is the module we seek, where B

is the module we originally found satisfying $T' \oplus B \approx T$.

We now prove the existence of minimal modules possessing an element of a given Ulm sequence.

COROLLARY. *Let* $\{\alpha_n\}$ *be a strictly increasing sequence of non-negative integers, and let* $\{\alpha_{n_i}\}$ *be its subsequence of gaps. Let* T *be the direct sum of cyclic modules* C_i *of order* $(\exp(\alpha_{n_i}+1))$. *Then there exists a KM module* 1 *with torsion submodule* T *and which contains an element* x *such that* $h(p^n x) = \alpha_n$. *Further,* M *is a direct summand of any KM module* M' *of rank* 1 *which contains an elements whose Ulm sequence is* $\{\alpha_n\}$.

Proof. We need only prove the last statement, since the existence of M with the prescribed invariants follows immediately from Theorem 6.2. Let T' be the torsion submodule of M'. By Kaplansky's lemma, the α_{n_i}th Ulm invariant of T' is non-zero. Hence there are cardinals U_n such that $U_n(T') = U_n + U_n(T)$. Let V be the torsion module with Ulm invariants given by U_n. By Ulm's theorem, $T' \approx V \oplus T$. The *KM* module $V \oplus M$ has torsion submodule $V \oplus T$ and $S(V \oplus M) = S(M')$. Hence $V \oplus M$ and M' are isomorphic, by the structure theorem.

Thus there is an uncountable number of non-isomorphic *KM* modules of rank 1 with no elements of infinite height. In particular we have exhibited modules of rank 1 which do not split.

DEFINITION. $x_1, \cdots, x_s$ is a *decomposition set for* M if it is a maximal independent subset of M and $h(\Sigma r_i x_i) = \min h(r_i x_i)$ for all $r_i \in R$. A *subdecomposition set* is a not necessarily maximal independent subset satisfying the above condition on heights.

DEFINITION. A decomposition set has *k gaps at level n* if k of its elements have Ulm sequences which have a gap at n.

LEMMA 6.3. *Let* $X = x_1, \cdots, x_s$ *be decompostion set with* k *gaps at level* n. *Then the the nth Ulm invariant of* $M \geq k$.

Proof. If $x_1, \cdots, x_s$ is a decompostion set for M, so is $r_1x_1, \cdots, r_sx_s$ where $r_i \neq 0$ for all i. Hence we may assume that $h(x_i) = n$ and $h(px_i) > n+1$ for $i \leqq k$. Thus there are elements y_i, $i \leq k$, such that $h(py_i) \geq n+1$ and $p^2 y_i = px_i$. Set $t_i = py_i - x_i$. We now have k elements of order p and of height n. It remains to prove that they are independent over $R/(p)$. Suppose $\sum_{i=1}^{k} r_i t_i = 0$, where r_i is either 0 or a unit in R. By the definition of the t_i, $\Sigma r_i(py_i - x_i) = 0$ which implies that $p\Sigma r_i y_i = \Sigma r_i x_i$. Since X is a decomposition set, $h(\Sigma r_i x_i) = \min h(r_i x_i) = n$ or ∞. But $h(p\Sigma r_i y_i) \geq n + 1$. Hence $\Sigma r_i x_i = 0$. The independence of the x's implies that each $r_i = 0$; hence the t_i are independent over $R/(p)$.

THEOREM 6.4. *Let M be a taut semi-KM module. M is completely decomposable if and only if M contains a decomposition set.*

Proof. If M is completely decomposable, the assertion is trivial. Suppose M contains a decomposition set $x_1, \cdots, x_s$. Define functions U_i: non-negative integers $\rightarrow$ cardinals $\leq \aleph_0$, $i = 1, 2, \cdots, s$ as follows: $\sum_{i=1}^{s} U_i(n) = n$th Ulm invariant of M; if the Ulm sequence of x_i has a gap at n, then $U_i(n) \neq 0$. By Lemma 6.3, the Ulm invariants of M are sufficiently large to allow this construction. Let T_i be the torsion module with Ulm invariants given by U_i. By Theorem 6.2, there exists a KM module of rank 1, M_i, having torsion submodule T_i and with $S(M_i)$ the equivalence class of the Ulm sequence of x_i. Consider ΣM_i. Since Ulm invariants are additive, the first condition in the definition of the U_i coupled with Ulm's theorem yields the fact that the torsion submodules of M and of ΣM_i are isomorphic. Further, $S(M) = S(\Sigma M_i)$. By Corollary 5.6, ΣM_i is a taut semi-KM module. By Theorem 5.8, $M \approx \Sigma M_i$.

LEMMA 6.5. *Let M and M' be taut semi-KM modules of rank* 1 *such that $S(M) = S(M')$. Then M and M' are almost isomorphic.*

Proof. Let T and T' be the torsion submodules of M and M' respectively. Then $M \oplus T'$ and $M' \oplus T$ are isomorphic, by 5.8.

We now prove a technical lemma which will allow us to obtain our first direct summand theorem.

LEMMA 6.6. *Let M be a reduced module of finite rank s. Let $x_1, \cdots, x_s$ be a decomposition set such that each x_i has the same Ulm sequence. Suppose also that $x_i = w_{i1}a_1 + \cdots + w_{is}a_s$, and, for all i, $|w_{i1}| \leq |w_{11}|$. Under these conditions, $y_i = w_{i1}x_1 - w_{11}x_i$, $i \geq 2$, is a subdecomposition set and each y_i is in A, the submodule generated by $a_2, \cdots, a_s$.*

Proof. Rank $M = s$ while rank $A \leq s - 1$. Hence not all the w_{i1} are 0 lest we have s independent elements $x_1, \cdots, x_s$ lying in A. Thus w_{11} is non-zero.

First we show the y_i's are independent. Suppose $\Sigma r_i y_i = 0$. Then $0 = (\Sigma r_i w_{i1})x_1 - \Sigma r_i w_{11} x_i$ which implies $r_i w_{11} = 0$ for all $i \geq 2$, since the x's are independent. Since $w_{11} \neq 0$, we must have $r_i = 0$ for all $i \geq 2$;

hence the y's are independent.

Next we show that the y's satisfy the required condition on their heights.

$$h(\Sigma r_i y_i) = h(\Sigma r_i w_{i1} x_1 - \Sigma r_i w_{11} x_i) = \min h(\Sigma r_i w_{i1} x_1), h(r_i w_{11} x_i) .$$

But

$$|\Sigma r_i w_{i1}| \leq \max |r_i w_{i1}| = \max |r_i||w_{i1}| \leq \max |r_i||w_{11}| = \max |r_i w_{11}| .$$

Hence there is an i such that $|\Sigma r_i w_{i1}| \leq |r_i w_{11}|$. Therefore, $h(\Sigma r_i w_{i1} x_1) \geq h(r_i w_{11} x_i)$ for that i. Hence $h(\Sigma r_i y_i) = \min h(r_i w_{11} x_i)$. On the other hand, $\min h(r_i y_i) = \min h(r_i w_{i1} x_1 - r_i w_{11} x_i) = \min h(r_i w_{i1} x_1), h(r_i w_{11} x_i)$. But for all i, $|r_i w_{i1}| \leq |r_i w_{11}|$. Therefore, $h(r_i w_{i1} x_1) \geq h(r_i w_{11} x_i)$. Hence $\min h(r_i y_i) = \min h(r_i w_{11} x_i)$. Hence $h(\Sigma r_i y_i) = \min h(r_i y_i)$.

THEOREM 6.7. *Let M be a completely decomposable semi-KM module with no elements of infinite height. Let $M = \Sigma M_i$, all the M_i isomorphic and of rank* 1. *If $M = A \oplus B$, then B is completely decomposable. In fact, B is almost isomorphic to a direct sum of copies of M_i.*

Proof. We first prove that any two elements in M of infinite order have equivalent Ulm sequences. Let $x_i \in M_i$ have infinite order. Clearly these x's form a decomposition set. Further, since all the M_i are isomorphic, we may assume that all the x_i's have identical Ulm sequences. Let $z \in M$ have infinite order. There is an $m \geq 0$ such that $p^m z = \Sigma r_i x_i$. Suppose $|r_i| \leq |r_1|$. Then $h(p^{m+k} z) = h(p^k \Sigma r_i x_i) = h(p^k r_1 x_1)$ for any non-negative k.

Choose $a_1, \cdots, a_{s-k}$ independent in A, $a_{s-k+1}, \cdots, a_s$ independent in B. We are now in the situation of the lemma. Applying the lemma k times (after each application, we must normalize the y's obtained so that they have identical Ulm sequences), we obtain $s - k$ independent elements in $\{a_{s-k+1}, \cdots, a_s\} \subset B$ which is a subdecompostion set of M. By the purity of B, and since rank $B = s - k$, these elements constitute a decomposition set for B. By Theorem 6.4, B is completely decomposable. Hence $B = \Sigma B_j$, and our initial remarks imply that $S(B_j) = S(M_i)$ for all i and j. By Lemma 6.5, B_j and M_i are almost isomorphic. Hence B is almost isomorphic to a direct sum of copies of M_i.

I have been unable to discover the truth of Theorem 6.7 in the event all the M_i are not isomorphic to each other.

COROLLARY 6.8. *Let $M = \sum M_\alpha$ (α in some index set), each M_α a semi-KM module of rank* 1 *with no elements of infinite height. If all the M_α are isomorphic, any direct summand B of M of finite rank is almost isomorphic to a direct sum of copies of M_α's.*

Proof. Let $x_\alpha \in M_\alpha$ have infinite order. Let $y_1, \cdots, y_s$ be a maximal independent subset of B. There is a finite subset $x_{\alpha_1}, \cdots, x_{\alpha_k}$ of the x_α's such that $p^m y_i$ lies in the submodule they generate, for all i. Let B' be the submodule of M generated by B and $M_{\alpha_1}, \cdots, M_{\alpha_k}$. Since B' is countably generated and of finite rank, $B' = \sum M'_{\alpha_j}$, where $S(M'_{\alpha_j}) = [Ux_{\alpha_j}]$. (If $x \in M$, Ux is its Ulm sequence and $[Ux]$ is the equivalence class of Ux). Hence all the $S(M'_{\alpha_j})$'s are the same. Since B is a direct summand of M, it is a direct summand of B'. By 6.7, B is completely decomposable. Since all elements of infinite order have equivalent Ulm sequences, B is almost isomorphic to a direct sum of copies of M_α.

We are now in a position to consider uniqueness of a decomposition of a module into the direct sum of modules of rank 1. The unpredictability of the torsion submodules does not allow one to find pairs of isomorphic summands from two different decompositions. For example, if C is cyclic of order (p) and $M = R \oplus C \oplus C \oplus R$, different associations yield different decompositions of M as a direct sum of modules of rank 1 whose terms are not pairwise isomorphic. However, the two decompositions do have isomorphic refinements.

THEOREM 6.9. *Let $M = \sum_{i=1}^n M_i$, M_i a KM module of rank 1, all the M_i isomorphic. Any two decompositions of M into summands of rank 1 have isomorphic refinements.*

Proof. We saw in the proof of Theorem 6.7 that any two elements in M of infinite order have equivalent Ulm sequences. Hence if $M = \sum_{i=1}^n M'_i$, all the M'_i of rank 1, then $S(M_i) = S(M'_i)$ for all i. By the existence theorem, there are modules N_i of rank 1 such that:

(i) $N_i \oplus T_i \approx M_i$, $N_i \oplus T'_i \approx M'_i$ for some torsion T_i, T'_i;

(ii) if W_i is the torsion submodule of N_i, then the Ulm invariants of W_i are 0's and 1's. Now $\sum W_i \oplus \sum T_i \approx \sum W_i \oplus \sum T'_i$. By Ulm's Theorem and condition (ii), we may cancel and obtain $\sum T_i \approx \sum T'_i$. Since any two decompositions of a module which is the direct sum of cyclic modules have isomorphic refinements, $\sum T_i$ and $\sum T'_i$ have isomorphic refinements. This completes the proof.

As a corollary, we have another proof of the square root problem, Theorem 4.5.

This paper is part of a dissertation written at the University of Chicago. I wish to thank Professor I. Kaplansky for his guidance.

REFERENCES

1. H. Cartan and S. Eilenberg, *Homological Algebra*, Princeton, 1956.
2. L. Fuchs, *Abelian Groups*, Publishing House of the Hungarian Academy of Sciences, 1958.
3. I. Kaplansky, *Infinite Abelian Groups*, University of Michigan Press, 1956.
4. I. Kaplansky and G. Mackey, *A Generalization of Ulm's theorem*, Summa Brasil. Math. **2** (1951), 195–202.
5. A. Kurosh, Theory of Groups, Vol. I. Chelsea, 1955.

UNIVERSITY OF ILLINOIS

THEORIE DES SYSTEMES DEMOSIENS DE GROUPOIDES

ALBERT SADE

1. Introduction. L'idée première de ces recherches est dans deux papiers de M. Schauffler consacrés à l'étude des codes avec un vocabulaire ayant un nombre uniforme de figures, [40], [41]. Dans le second, il construit un tel code au moyen d'une population Ω de quasigroupes Q_i, définis sur le même ensemble fini, $E = (1, 2, 3, \cdots, n)$ et satisfaisant à une associativité qu'il appelle "im Ganzen", $\forall x, y, z \in E$, $\forall Q_1, Q_2 \in \Omega$, $\exists Q_3, Q_4 \in \Omega$, $xQ_1(yQ_2z) = (xQ_3y)Q_4z$. Il montre que l'ensemble de tous les quasigroupes construits sur E ne peut être associatif "im Ganzen" si n surpasse 3. Dans le présent travail on se propose, sans préoccupations cryptographiques immédiates, d'étudier systématiquement les ensembles (finis ou non) de groupoïdes construits sur un support commun et satisfaisant à quelque relation *demosienne* analogue à l'associativité "im Ganzen". De telles considérations ont déjà été abordées dans un précédent travail de l'auteur ([35], p. 156, N°2, iv, p. 161, N°8, IV). Elles ne sont pas seulement susceptibles de conduire à des applications dans le domaine du "chiffre", elles présentent encore un intérêt en soi dans le champ de la spéculation pure. De tels *ensembles multistructurés*, c'est-à-dire munis de plusieurs lois de composition, se rencontrent à chaque pas en algèbre. On sait que les *anneaux*, *corps*, *clusters*, *narings* et *neofields* ([32], p. 296, III) possédent deux lois de composition. Skolem ([43], p. 53) donne un système de quatre semigroupes idempotents, abéliens, et qui sont deux à deux mutuellement distributifs. Les ensembles de groupoïdes engendrés par deux *groupoïdes orthogonaux* ([31], p. 231, N°6), les *ringoïdes* ([7], p. 203, N°2) en possèdent un nombre quelconque. Dans [21 b], Hasse, p 27, définite un ensemble muni de quatre opérations.

Le fait que le même ensemble soit muni de toute une population de lois de composition suggère le nom de *demosiennes*, déjà introduit dans [35] pour qualifier les identités entre éléments de tels ensembles. On trouve dans la littérature maints exemples de relations contenant à la fois plusieurs lois de composition. A peu près toutes les égalités de l'algèbre classique font intervenir six opérations usuelles. L'équation d'associativité mutuelle ([36], déf. 17, [43], p. 47), $(x \times y) \ast z = x \times (y \ast z)$, l'équation de Kolmogoroff ([2], équ. I'), $st \times tu = su$, l'équation de Cacciopoli [9], Ghermănescu [15], Gyires [19], Aczél [1], [3], $f(x \times y) = f(x) \ast f(y)$, où f est une application de E sur lui-même, l'équation de Ghermănescu [16], $(x \ast a) \times (y \ast b) = x \times y$, celle de Kurepa [29], $[(a \times b) \ast c] \times (a \ast b) = (b \ast c) \times [a \ast (b \times c)]$, l'équation de distibuti-

Received April 27, 1959.

vité [23], [17], ([43], p. 52, équ. 4), $(x \ast y) \times z = (x \times z) \ast (y \times z)$, l'équation d'Aczél-Hosszú [4], $(x \times y) \times z = x \times (y \circ z)$, contiennent toutes deux opérations différentes. Hosszú ([25], p. 206), considère une identité avec quatre lois de composition, $F[x, G(y, z)] = H[K(x, y), z]$.

Mais la plupart des auteurs qui ont traité de ces relations ont regardé les lois qu'elles contiennent comme des fonctions de deux variables définies sur un corps (celui des réels, en général). Si l'on cesse de considérer ces égalités comme des équations fonctionnelles, pour les interpréter comme des conditions entre éléments d'une même support, muni de plusieurs lois, c'est-à-dire comme des équations entre groupoïdes, alors ce changement de point de vue peut amener, avec une généralité plus grande, des simplifications notables et inattendues. Pareil fait n'est pas nouveau. Scherk [42], développe sur dix pages de pesantes considérations d'analyse pour établir une proposition dont la démonstration directe tient en quelques lignes. Ici, soit par exemple l'équation de Cauchy-Cacciopoli, $f(x \times y) = f(x) \ast f(y)$, qui a fait l'objet de nombreux mémoires couvrant plus d'une centaine de pages. Si E est un ensemble quelconque, muni de deux opérations $(\times)$ et $(\ast)$, cette équation exprime que le groupoïde $G' = E(\ast)$ est homomorphe de $G = E(\times)$. Soit $T = [x \rightarrow f(x)]$ l'homomorphisme qui fait passer de G à G' et A le groupe d'automorphisme de G; alors, toutes les solutions en f sont données par les éléments du coset A_T. Le problème n'est possible que si G' est homomorphe de G. L'affirmation d'Aczél [1], que G et G' sont en même temps bissymétriques ou non devient évidente puisque $G \sim G'$. Le théorème d'Aczél ([3], p. 329), que (8) est un groupe continu, devient immédiat puisque G' est alors homomorphe au groupe additif des réels. De même l'équation ci-dessus de Hosszú, mise sous la forme $x \times (y \ast z) = (x \ominus y) \circ z$, a été complètement résolue par Belousov [6], dans le cas des quasigroupes, par Hosszú [26], et par Rado (Cluj). Les quatre quasigroupes sont isotopes d'un même groupe. (Voir, N°7.2 une solution différente de ce problème, et [37] une extension aux multigroupoïdes).[1] On aura un autre exemple de telles simplifications à propos de l'équation de distributivité (Ci-après, N°8) et de celle de transitivité [38].

Il est certain qu'un pareil sujet déborde le cadre d'une simple note; nous nous bornerons à l'esquisser ici. Les questions abordées sont, en se limitant à quelques identités classiques, les systèmes satisfaisant à une équation fonctionnelle ou à une loi demosienne particulière donnée, les conséquences de l'existence de deux lois demosiennes, les systèmes demosiens dont les éléments sont dérivés d'un même groupe par des isotopies ayant pour composantes des translations de ce groupe, (systèmes (G, K, T)), un essai sur les systèmes demosiens organisés au moyen d'une loi de composition entre les groupoïdes qui les forment.

La nomenclature, les définitions et notations sont celles des trois papiers [34], [35], [27], auxquels le lecteur est prié de se reporter. En outre les symboles suivants seront utilisés:

$a \prec b$, Relation d'ordre,
φ_i, symbole opératoire d'une loi de composition,
Φ, ensemble des φ_i, sur un même ensemble E,
$\langle a, b, c, \cdots \rangle$, ensemble ordonné,
(ξ, η, ζ), isotopie de composantes ξ, η, ζ.

Un index, avant la bibliographie, renvoie aux N°.

Chapitre I

SYSTÈMES ADMETTANT UNE LOI DEMOSIENNE DÉTERMINÉE

2. **Définitions.** Un ensemble quelconque, fini ou non, $E = (x, y, \cdots)$ muni de plusieurs lois de composition φ, formant une population finie ou non, $\Phi = (\varphi_1, \varphi_2, \cdots)$, est dit *multistructuré* ou *demosien*, et noté (E, Φ). Les lois φ sont supposées partout définies et homogènes ([35], p. 156, N°2, équ. 55) et, sauf stipulation contraire, il n'est fait aucune supposition particulière (commutativité, associativité, axiome d'absorption, [7], p. 18) sur la nature de ces lois. Une *expression sur* (E, Φ) est un assemblage d'éléments $\in E$, séparés par les signes opératoires $\varphi_i \in \Phi$ et par des parenthèses, crochets et accolades. Elle définit une suite d'opérations à effectuer, dans un ordre déterminé, sur ces éléments, aboutissant comme résultat final à un élément bien défini $\in E$. Deux expressions sont *égales* si elles définissent le même élément. Si deux expressions sont égales quels que soient les éléments qu'elles contiennent, elles forment une *identité*. Si l'égalité n'a lieu que pour certains choix des lettres, on obtient une *équation*. Mais les choses dépendent aussi des signes opératoires, et les significations des deux vocables empiètent; il convient de préciser dans chaque cas quels sont les éléments ou symboles qui peuvent, on non, être choisis arbitrairement dans une égalité. De telles relations sont dites *demosiennes* parce qu'elles ont lieu sur toute une population de groupoïdes ([35], p. 156, N°2 iv; p. 161, N°8, IV, p. 172, N°30).

Remarquons qu'un système de groupoïdes homogènes satisfaisant à une identité demosienne, quels que soient les éléments et les symboles, se réduirait en général à un seul groupoïde. Soit par exemple

$$\forall x, y, z \in E\,, \qquad \forall \varphi_i \in \Phi\,, \qquad (x\varphi_1 y)\varphi_2 z = x\varphi_3(y\varphi_4 z)\,.$$

Si les $E(\varphi)$ sont homogènes on peut $\forall a, b \in E$, choisir $x = a$ et y, z tels que $y\varphi_4 z = b$. Laissant tous les φ, sauf φ_3, invariables et posant $(x\varphi_1 y)\varphi_2 z = c$, on aurait donc $\forall \varphi_3 \in \Phi$, $c = a\varphi_3 b$, $a, b, c =$ Constantes. Le produit $a\varphi b$ serait le même dans tous les groupoïdes. Cela pouvant étre répété $\forall a, b$, tous les $E(\varphi)$ coïncideraient avec un même groupoïde, qui serait évidemment un semigroupe. Plus généralement, soit A une expression dépendant des éléments $x, y, z, \cdots \in E$, associés au moyen des lois $\varphi_1, \varphi_2, \varphi_3, \cdots \in \Phi$ et B une expression analogue. Supposons l'identité $A = B$ vérifiée $\forall x, y, \cdots \in E$, $\forall \varphi_1, \varphi_2, \cdots \in \Phi$, Alors, laissant tous les x constants et tous les φ, sauf un, φ_i, invariables, on aura, en désignant par a, b, c trois constantes, $c = a\varphi_i b$. Ainsi, le produit $a\varphi b$ sera le même dans tous les groupoïdes du système, Si toutes les lois de Φ sont homogènes, on pourra toujours choisir les $x, y, \cdots$ de manière à attribuer à a et b deux valeurs choisies d'avance (l'homogénéité n'est même pas toujours nécessaire, comme par exemple dans le cas de la commutativité demosienne). Alors

$$\forall a, b \in E , \quad \forall \varphi_i \in \Phi , \quad a\varphi_i b = \text{constante.}$$

Le produit $a\varphi b$ étant le même dans tous les groupoïdes, Φ se réduit à une seule loi.

3. Commutativite demosienne. DÉFINITION 3.1. *Un système (E, Φ) admet la commutativité demosienne si*

$$\forall x, y \in E , \quad \forall \varphi \in \Phi , \quad \exists \varphi' , \quad x\varphi y = y\varphi' x .$$

Quand φ' ne dépend ni de x, ni de y, la commutativité est *forte*. Quand φ' dépend à la fois de φ, de x et de y, la commutativité est *faible*. Il est clair qu'*un système satisfaisant à la commutativité forte contient, avec tout groupoïde, $G = E(\varphi)$, le groupoïde conjoint*, ([35], p. 155, N°2, ii), $x\varphi y = z \rightleftarrows y\varphi' x = z$. Dans le cas fini, les deux tables de Cayley de $G = E(\varphi)$ et de $G' = E(\varphi')$ sont deux matrices dont chacune est transposée de l'autre.

EXEMPLE 3.2. Le système de quasigroupes, défini sur le corps R des réels par $x\varphi_\lambda y \equiv (a\lambda + b)x + (c\lambda + d)y + f\lambda + g$, $a, b, c \cdots, g$, constantes $\in R$, admet la commutativité faible.

DÉFINITION 3.3. *Un sous-système d'un système (E, Φ) satisfaisant à une ou plusieurs relations demosiennes est un système (E', Φ'), avec $E' \subseteq E$, $\Phi' \subseteq \Phi$, et satisfaisant aux mêmes relations demosiennes.*

EXEMPLE 3.4. Dans le système (E, Φ) ci-dessus (3.1), les mêmes équations de définition, appliquées au corps Q des fractions relationnelles, fournissent le sous-système demosien (Q, Φ'), $\lambda, a, b, \cdots, g \in Q$.

Question 3.5. *Un système commutatif faible contient-il toujours deux groupoïdes conjoints* (distincts ou non)?

DÉFINITION 3.6. *Etant donné un quasigroupe $Q(\times)$, l'ensemble $(\cdots, \Delta_i^{-1}\Delta_j, \cdots)$, $i, j \in Q$, où $\Delta_i = (x \rightarrow x \times i)$, s'appelle le complexe* *relatif aux translations à droite, l'ensemble* $(\cdots, \Gamma_i^{-1}\Gamma_j, \cdots)$, *le complexe relatif aux translations à gauche.*

Ces deux complexes engendrent deux groupes qui sont des diviseurs des groupes engendrés par les translations elles-mêmes et introduits par Albert [5], p. 509). L'intérêt de ces deux groupes est qu'ils restent invariants si l'on soumet le quasigroupe Q à une isotopie quelconque de forme $(\xi, \eta, 1)$ Si l'on fait subir à $Q(\times)$ l'isotopie principle $x \times y = x\xi * y\eta$, on aura

$$\Delta_i^{-1}\Delta_j = (x \times i \rightarrow x \times j)$$

et

$$\Delta_{i\eta}^{-1}\Delta_{j\eta} = (u * i\eta \rightarrow u * j\eta) = (x\xi * i\eta \rightarrow x\xi * j\eta)$$
$$= (x \times i \rightarrow x \times j) = \Delta_i^{-1}\Delta_j \,.$$

Le calcul est le même à gauche.

THÉORÈME 3.7. *Le système demosien dérivé d'un groupoïde abélien en le soumettant à toutes les isotopies possibles est commutatif demosien fort.*

Preuve. Si G est un groupoïde abélien, les isotopies (ξ, η, ζ) et (η, ξ, ζ) le transforment en deux groupoïdes conjoints. A toute isotopie appliquée à G en correspond une autre (pouvant coïncider avec la première) et pour laquelle les deux isotopes obtenus sont conjoints. Ainsi, le système possède la commutativité forte. La réciproque n'est pas vraie comme le montre l'exemple suivant. Soit Q le quasigroupe du 5° ordre défini par ses translations à droite $\Delta_0 = 1$, $\Delta_1 = (01)(234)$, $\Delta_2 = (04132)$, $\Delta_3 = (03124)$, $\Delta_4 = (02143)$. Par l'isotopie $(1, 1, (24))$, Q devient son propre conjoint Q'. Si l'on soumet Q et Q' à une commune isotopie on obtiendra deux quasigroupes conjoints et le système (Q, Φ) aura la commutativité forte. Pourtant aucun isotope de Q ne sera abélien. Il suffit pour s'en assurer d'examiner les isotopies $(1, \eta, 1)$; or aucune ne rend Q abélien. On a toutefois la condition.

THÉORÈME 3.8. *Pour que le système obtenu en soumettant un quasigroupe Q à toutes les isotopies admette la commutativité demosienne il faut et il suffit que, dans Q, les complexes relatifs aux translations à droite,* $(\cdots\Delta_i^{-1}\Delta_j\cdots)$ *et à gauche* $(\cdots\Gamma_i^{-1}\Gamma_j\cdots)$, *soient isomorphes.*[2]

Preuve. La condition est nécessaire. Dans toute isotopie, chacun des complexes reste évidemment isomorphe à lui-même. Si (E, Φ) est commutatif demosien, il contient, avec tout quasigroupe K, son conjoint K'. Soient G et D les complexes à gauche et à droite de K, G' et D' ceux de K'. On aura, puisque K et K' sont isotopes, $G \cong G'$, $D \cong D'$. Mais, d'autre part, K et K' étant conjoints, $G = D'$, $D = G'$, donc $D \cong G$.

Elle est suffisante. Soit S la permutation de l'ensemble Q qui projette le premier complexe sur le second. On peut d'abord par une isotopie $(\xi, 1, 1)$ s'arranger de manière que $\forall i, j,\ \Delta_i^{-1}\Delta_j \cong \Gamma_i^{-1}\Gamma_j$; en faisant alors l'isotopie $(1, 1, S)$, le nouveau complexe à gauche deviendra l'ancien complexe à droite. Une dernière isotopie de la forme $(1, \eta, 1)$ fournira le conjoint de Q.

COROLLAIRE 3.9. *L'ensemble des isotopes d'un groupe possède la commutativité demosienne.*

Car ses deux Cayleyens sont isomorphes.

Remarquons pour terminer que l'image d'un système demosien commutatif par une isotopie de la forme $(\xi, \eta = \xi, \zeta)$, appliquée à tous ses groupoïdes, est encore un système commutatif.

4. Loi des keys. DÉFINITION 4.1. *Un système* (E, Φ) *satisfait à la loi demosienne des keys* (à droite) *si la condition*

$$(4.1)\qquad \forall x, y \in E\,,\qquad \forall\varphi_1 \in \Phi\,,\qquad \exists\varphi_2 \in \Phi\,,\qquad (x\varphi_1 y)\varphi_2 y = x$$

est vérifiée. Cette loi apparaît pour la première fois dans Grassmann ([18], p. 37).

THÉORÈME 4.2. *Pour qu'un système* (E, Φ) *de quasigroupes à gauche satisfasse à la loi demosienne des keys* (*à droite*) *il faut et il suffit qu'il contienne, en même temps que tout quasigroupe* $Q = E(\times)$, *son réciproque* $Q' = E(\ominus)$, *défini par* $x \times y = z \rightleftarrows z \ominus y = x$, ([34], déf. 1.2). Pour les keys à gauche il faudrait $x \times y = z \rightleftarrows x \odot z = y$, ([34], déf. 1.5).

Preuve. En effet (4.1) est équivalente à $\forall x, y \in E$, $\forall\varphi_1 \in \Phi$, $\exists\varphi_2 \in \Phi$, $x\varphi_1 y = z \rightleftarrows z\varphi_2 y = x$. On démontre, comme pour 3.7 que, si un quasigroupe à gauche Q est self-réciproque le système dérivé de Q en le soumettant à toutes les isotopies possibles satisfait à la loi demosienne des keys. Enfin une conséquence du Théorème 3.8 est que, pour que l'ensemble des isotopes d'un quasigroupe Q satisfasse à la loi demosienne des keys (à droite) il faut et il suffit que Q soit parastrophique par $(x \times y = z \rightleftarrows x \odot z = y)$ d'un quasigroupe satisfaisant à la condition du Théorème 3.8. Car, soit $a\psi_i b = c \rightleftarrows a\varphi_i c = b$ et $x\psi_1 y = y\psi_2 x = z$; il en résultera $x\varphi_1 z = y$ et $y\varphi_2 z = x$, d'où, par élimination $(x\varphi_1 z)\varphi_2 z = x$. (Cf N°9.3)

5. Demi-symétrie. DÉFINITION 5.1. *Un système* (E, Φ) *de groupoïdes satisfait à la demi-symétrie demosienne si*

$$\forall x, y \in E\,,\qquad \forall\varphi_1 \in \Phi\,,\qquad \exists\varphi_2 \in \Phi\,,\qquad x\varphi_1(y\varphi_2 x) = y$$

([35], p. 153, N°2, équ. 11), ([34], déf. 18.7).

THÉORÈME 5.2. *Pour qu'un système* (E, Φ) *de quasigroupes à droite* ([34], N°1) *admette la demi-symétrie demosienne il faut et il suffit qu'il contienne, avec tout quasigroupe* $E(\varphi)$, *son parastrophique par* $a\varphi b = c \rightleftarrows b = c \oplus a$, ([34], déf. 1.4), *c'est à dire, avec chaque opération, sa division à gauche* ([11], p. 170).

Preuve. Soit $x\varphi_1 u = y$, donc $u = y \oplus x$; alors $x\varphi_1(y \oplus x) = y$, donc $y \oplus x = y\varphi_2 x$, $E(\varphi_2) = E(\oplus)$. Ainsi le système contient, avec tout quasigroupe à droite $E(\varphi_1)$, son parastrophique $E(\oplus)$. La réciproque est évidente.

On parvient au sujet de la demi-symétrie, à des conclusions analogues à celles des N°3 et 4.

6. Inversibilité.[3] DÉFINITION 6.1. ([41], p. 428, dans le cas fini). *Un système* (E, Φ) *admet l'inversibilité demosienne s'il satisfait à*

$$\forall x, y, z \in E\,, \qquad \forall \varphi_1, \varphi_2 \in \Phi\,, \qquad \exists \varphi_3, \varphi_4, \varphi_5, \varphi_6 \in \Phi\,,$$
$$(x\varphi_1 y)\varphi_2 z = (z\varphi_3 y)\varphi_4 x\,, \qquad x\varphi_1(y\varphi_2 z) = z\varphi_5(y\varphi_6 x)\,.$$

Les deux relations sont conjointes. La première est l'extension demosienne de la loi d'Abel-Grassmann ([35], p. 154, N°2, équ. 21).

EXEMPLE 6.2. Sur le corps Q des fractions rationnelles, le système des quasigroupes définis par $x\varphi y = ax + by + c$, $a, b, c \in Q$ est inversible demosien.

7. Associativité. DÉFINITION 7.1. *Un système* (E, Φ) *satisfait à l'associativité demosienne si la condition*

$$\forall x, y, z \in E\,, \qquad \forall \varphi_1, \varphi_2 \in \Phi\,, \qquad \exists \varphi_3, \varphi_4, \varphi_5, \varphi_6 \in \Phi\,,$$
$$(x\varphi_1 y)\varphi_2 z = x\varphi_3(y\varphi_4 z)\,, \qquad x\varphi_1(y\varphi_2 z) = (x\varphi_5 y)\varphi_6 z \tag{7.1}$$

est vérifiée.

Pour approcher de la solution du problème posé par la construction de tels systèmes, on peut d'abord chercher les conditions auxquelles doivent satisfaire quatre groupoïdes pour être solution de (7.1). Cette équation a été étudiée par Evans [13], Belousov [6] et Hosszú [26]. Le premier a montré que, si les $E(\varphi)$ sont isotopes d'un groupoïde fini, avec élément neutre, ce groupoïde est associatif. Le troisième a donné la solution générale de (7.1) dans le cas où les φ sont des fonctions continues, différentiables, strictement monotones ([25], p. 212). Belousov a énoncé et Hosszú [26] a démontré que quatre quasigroupes satisfaisant (7.1) sont isotopes d'un même groupe, théorème que j'ai étendu aux multigroupoïdes [37]. Le théorème suivant donne une solution explicite générale de l'équation demosienne d'associativité lorsque les φ sont des fonctions arbitraires sur un ensemble quelconque, avec fonction inverse uniforme, c'est-à-dire dans le cas des quasigroupes. Cette solution reste valable—mais sans être générale—si les groupoïdes sont quelconques.

THÉORÈME 7.2. *Si E est un ensemble quelconque,* (i) *la solution générale de*

$$(7.2) \qquad \forall x, y, z \in E\ , \qquad (x\varphi_1 y)\varphi_2 z = x\varphi_3(y\varphi_4 z)\ ,$$

où les φ sont des lois de quasigroupes, est

$$(\dot{\mathrm{I}})\begin{cases} x\varphi_1 y = x\xi \cdot y\theta\ , \\ x\varphi_2 y = (x \cdot y\lambda)\zeta^{-1}\ , \\ x\varphi_3 y = (x\xi \cdot y\eta)\zeta^{-1}\ , \\ x\varphi_4 y = (x\theta \cdot y\lambda)\eta^{-1}\ , \end{cases}$$

où $\xi, \eta, \zeta, \theta, \lambda$ sont cinq permutations arbitraires de E et $E(\cdot)$ un groupe défini sur E.

(ii) *Quatre groupoïdes quelconques, isotopes d'un même semigroupe $E(\cdot)$ par les isotopies* ($\dot{\mathrm{I}}$), *sont solution de* (7.2).

Preuve. Soient quatre quasigroupes définis sur le même ensemble E par les lois de composition $\varphi_1, \varphi_2, \varphi_3, \varphi_4$ et satisfaisant à l'identité (7.2). Désignons la translation à droite, relative à l'élément a, dans le quasigroupe $E(\varphi_i)$ par Δ_a^i. Faisons décrire E à l'élément générique x et assignons des valeurs fixes, a et b, à y et z respectivement. L'égalité

$$(7.3) \qquad (x\varphi_1 a)\varphi_2 b = x\varphi_3(a\varphi_4 b)$$

s'écrit

$$(7.4) \qquad \Delta_a^1\Delta_b^2 = \Delta_c^3\ , \qquad (c = a\varphi_4 b)\ .$$

Dans (7.4) on peut maintenant supposer a constant et faire décrire à b tout le champ E; alors c, dans le quasigroupe $E(\varphi_4)$, décrira aussi tout l'ensemble E. Donc Δ_c^3 décrira la totalité des translations à droite du quasigroupe $E(\varphi_3)$. En recommençant le même processus à partir d'une autre valeur de $a \in E$, on devra obtenir, chaque fois, le même ensemble de valeurs de Δ_c^3, sans quoi les translations de $E(\varphi_3)$ ne seraient pas définies univoques. Les ensembles

$$\begin{array}{ccccccc} \Delta_0^1\Delta_0^2, & \Delta_0^1\Delta_1^2, & \Delta_0^1\Delta_2^2, & \Delta_0^1\Delta_3^2, & \cdots, & \Delta_0^1\Delta_i^2, & \cdots \\ \Delta_1^1\Delta_0^2, & \Delta_1^1\Delta_1^2, & \Delta_1^1\Delta_2^2, & \Delta_1^1\Delta_3^2, & \cdots, & \Delta_1^1\Delta_i^2, & \cdots \\ \cdots, & \cdots, & \cdots, & \cdots, & \cdots, & \cdots, & \cdots \\ \Delta_j^1\Delta_0^2, & \Delta_j^1\Delta_1^2, & \Delta_j^1\Delta_2^2, & \Delta_j^1\Delta_3^2, & \cdots, & \Delta_j^1\Delta_i^2, & \cdots \\ \cdots, & \cdots, & \cdots, & \cdots, & \cdots, & \cdots, & \cdots \end{array}$$

où l'emploi des indices inférieurs ne signifie pas que ces ensembles soient dénombrables, seront tous identiques. Considérons le groupe symétrique total $\mathscr{S}_E$, et dans ce groupe les deux complexes

$$A = \Delta_0^1, \Delta_1^1, \Delta_2^1, \Delta_3^1, \cdots, \Delta_j^1, \cdots$$
$$B = \Delta_0^2, \Delta_1^2, \Delta_2^2, \Delta_3^2, \cdots, \Delta_i^2, \cdots.$$

D'après ce qui précède, si l'on multiplie à gauche le second complexe B par un élément quelconque Δ_j^1 du premier A, l'ensemble T des permutations obtenues doit rester le même $\forall \Delta_j^1$. Si l'on introduit les nouveaux complexes $C = P^{-1}A$, $D = BQ^{-1}$, $P \in A$, $Q \in B$, on aura par hypothèse

$$CD = P^{-1}ABQ^{-1} = P^{-1}TQ^{-1} .$$

Or C et D contiennent l'unité $P^{-1}P = QQ^{-1} = 1$ de $\mathscr{S}_E$, donc

$$C1 = C = P^{-1}TQ^{-1} , \qquad 1D = D = P^{-1}TQ^{-1}$$

et

$$P^{-1}TQ^{-1} = C = D = CD = CC .$$

Ainsi C est fermé, associatif (puisque c'est un sous ensemble de $\mathscr{S}_E$) et contient l'identique. C'est donc un semigroupe avec élément neutre, contenu dans $\mathscr{S}_E$. Les permutations Δ_a^0, éléments de C, sont les translations à droite d'un semigroupe isomorphe à C, $G = E(\varphi_0)$, avec $\Delta_a^0 = (x \to x\varphi_0 a)$. Donc

$$\exists S \in \mathscr{S}_E , \qquad x\varphi_1 y = x\Delta_y^1 = (xP)\Delta_{yS}^0 = (xP)\varphi_0(yS) , \tag{7.5}$$

ce qui exprime que $E(\varphi_1)$ est isotope de G par $(\xi = P,\ \eta = S,\ \zeta = 1)$ et $E(\varphi_1)$, étant un quasigroupe, G est donc un groupe. De même

$$\exists T \in \mathscr{S}_E , \qquad x\varphi_2 y = x\Delta_y^2 = (x\Delta_{yT}^0)Q = (x\varphi_0 yT)Q , \tag{7.6}$$

ce qui signifie que $E(\varphi_2)$ est isotope de G par $(\xi = 1,\ \eta = T,\ \zeta = Q^{-1})$. L'égalité (7.4) prend la forme $P\Delta_{aS}^0\Delta_{bT}^0Q = \Delta_c^3$,

$$P\Delta_{aS\varphi_0 bT}^0 Q = \Delta_{a\varphi_4 b}^3 . \tag{7.7}$$

Cela entraîne que $(aS\varphi_0 bT \to a\varphi_4 b)$ soit une permutation de E car, d'abord $a\varphi_4 b$ décrit tout le champ E, ensuite, d'après (7.5), $aS\varphi_0 bT$ décrit aussi tout l'ensemble E, enfin, si avec $aS\varphi_0 bT = a'S\varphi_0 b'T$ on avait $a\varphi_4 b \neq a'\varphi_4 b'$, il en résulterait $\Delta_{a'\varphi_4 b'}^3 = \Delta_{a\varphi_4 b}^3$; deux translations distinctes de $E(\varphi_3)$ coïncideraient et φ_3 ne serait plus une loi de quasigroupe. Le même argument est valable en renversant les rôles de φ_0 et de φ_4. On a donc

$$a\varphi_4 b = (aS\varphi_0 bT)R , \tag{7.8}$$

où R est une permutation de E, et $E(\varphi_4)$ est isotope de $E(\varphi_0)$ par $(\xi = S,\ \eta = T,\ \zeta = R^{-1})$. Maintenant (7.7) s'écrit

$$(7.9)\qquad \varDelta^3_{cR} = P\varDelta^0_c Q\ ,\qquad x\varphi_3 cR = (xP\varphi_0 c)Q\ ,\qquad x\varphi_3 y = (xP\varphi_0 yR^{-1})Q\ ,$$

ce qui exprime que $E(\varphi_3)$ est isotope de G par $(\xi = P,\ \eta = R^{-1},\ \zeta = Q^{-1})$. En remplaçant φ_0 par $(\cdot)$ dans les relations (7.5), (7.6), (7.8), (7.9), elles prennent la forme $(\dot{\mathrm{I}})$ de l'énoncé. On vérifie immédiatement que la condition $(\dot{\mathrm{I}})$ est suffisante, même si les $E(\varphi)$ sont des groupoïdes quelconques, pourvu que $E(\cdot)$ soit associatif, ce qui établit (ii).

Exemple 7.3. Sur l'anneau Z des entiers rationnels tous les groupoïdes $x\varphi_i y = x + y + i$ sont des groupes. Si l'on suppose i etj compris entre deux entiers fixes, l'ensemble obtenu aura l'associativité demosienne avec $(x\varphi_i y)\varphi_j z = x\varphi_{i\pm h}(y\varphi_{j\mp h}z)$.

Exemple 7.4. Soit le semigroupe $x \cdot y = x + y$ sur l'ensemble N^+ des entiers naturels; il ne possède ni inverse, ni élément neutre, mais les quatre isotopes $x\varphi_1 y = ax + y$, $x\varphi_2 y = b(x + y)$, $x\varphi_3 y = b(ax + y + c)$, $x\varphi_4 y = x + y - c$, où $a, b, c \in N^+$, vérifient l'équation (7.2).

Exemple 7.5. Sur l'anneau Z/n (et dans tout champ de Galois), l'ensemble des quasigroupes $x\varphi y = ax + by + c$, $(a, n) = (b, n) = 1$, est associatif demosien (Il est aussi reversible). Cet ensemble contient $n[\varphi(n)]^2$ quasigroupes, $\varphi(n)$ étant l'indicateur d'Euler.

Remarque 7.6. Pour obtenir toutes les solutions de (7.2) il faut faire parcourir à G l'ensemble de tous les groupes $G = E(\cdot)$, que l'on peut construire sur E. Si deux d'entre eux sont isomorphes, $(G' = G_T)$ les formules $(\dot{\mathrm{I}})$ donneront des solutions isomorphes par T. Par conséquent les formes $(\dot{\mathrm{I}}')$

$$(\dot{\mathrm{I}}')\qquad \begin{aligned} x\varphi_1 y &= (x\xi \cdot y\theta)\mu^{-1}\ ,\\ x\varphi_2 y &= (x\mu \cdot y\lambda)\zeta^{-1}\ ,\\ x\varphi_3 y &= (x\xi \cdot y\eta)\zeta^{-1}\ ,\\ x\varphi_4 y &= (x\theta \cdot y\lambda)\eta^{-1}\ ,\end{aligned}$$

ne sont pas plus générales que $(\dot{\mathrm{I}})$.

Théorème 7.7. *Tout système (E, Φ) de groupoïdes satisfaisant aux axiomes,*

(i) *associativité demosienne deux côtés,*
(ii) *élément neutre,*
(iii) *inverse, se réduit à un seul groupe.*

Preuve. Par hypothèse

$$\forall\varphi_1, \varphi_2 \in \Phi \,, \quad \forall x, y, z \in E \,, \quad \exists\varphi_3, \varphi_4 \in \Phi \,, \quad (x\varphi_1 y)\varphi_2 z = x\varphi_3(y\varphi_4 z) \,,$$
$$\forall\varphi_3, \varphi_4 \in \Phi \,, \quad \forall x, y, z \in E \,, \quad \exists\varphi_1, \varphi_2 \in \Phi \,, \quad (x\varphi_1 y)\varphi_2 z = x\varphi_3(y\varphi_4 z) \,.$$

Il existe au moins un élément neutre à gauche, le même pour tous les groupoïdes, $\forall\varphi \in \Phi$, $\forall x \in E$, $\exists e \in E$, $e\varphi x = x$.

Parmi les e il y en a au moins un tel que chaque élément x de E ait un, ou plusieurs, inverses à gauche, dans tous les groupoïdes, ne dépendant que de x

$$\forall\varphi \in \Phi \,, \qquad \forall x \in E \,, \qquad \exists e, \bar{x} \in E \,, \qquad \bar{x}\varphi x = e \,.$$

(A) Soit $e\varphi a = a$, $\bar{x}\varphi x = e$. Multiplions à droite, les deux membres de la seconde égalité par e, $(\bar{x}\phi x)\varphi_1 e = e\varphi_1 e = e = \bar{x}\varphi x$, donc $(\bar{x}\varphi x)\varphi_1 e = \bar{x}\varphi x$. Par l'associativité demosienne,

$$\bar{x}\varphi x = (\bar{x}\varphi x)\varphi_1 e = \bar{x}\varphi_2(x\varphi_3 e) \,.$$

Posons $y\varphi\bar{x} = e$ et multiplions les deux membres de l'égalité précédente par y, à gauche

$$y\varphi_4(\bar{x}\varphi x) = y\varphi_4[\bar{x}\varphi_2(x\varphi_3 e)] \,,$$

appliquant l'associativité, $(y\varphi_7\bar{x})\varphi_8 x = (y\varphi_5\bar{x})\varphi_6(x\varphi_3 e)$, ou $e\varphi_8 x = e\varphi_6(x\varphi_3 e)$, et enfin $x = x\varphi_3 e$. Mais, plus haut, φ_3 a été défini par $(a\varphi b)\varphi_1 c = a\varphi_2(b\varphi_3 c)$. L'associativité demosienne étant supposée bilatère, (i), $\forall\varphi_2, \varphi_3 \in \Phi$, $\exists\varphi_1, \varphi \in \Phi$, donc $x\varphi_3 e = x$ est varie quelque soit φ_3. Ainsi e est élément neutre à droite. Si maintenant e et $é$ sont deux éléments neutres, $e\varphi é = e = é$; donc tous les éléments neutres sont égaux et l'unité est unique et bilatère.

(B) Multiplions les deux membres de $\bar{x}\varphi x = e$ par $\bar{x}$; on a $(\bar{x}\varphi x)\varphi_1\bar{x} = e\varphi_1\bar{x} = \bar{x}$. En appliquant l'associativité, $\bar{x}\varphi_2(x\varphi_3\bar{x}) = \bar{x}$. Posons $x'\varphi\bar{x} = e$ et multiplions les deux membres de la dernière égalité par x', à gauche, $x'\varphi_4[\bar{x}\varphi_2(x\varphi_3\bar{x})] = x'\varphi_4\bar{x}$, ou $(x'\varphi_5\bar{x})\varphi_6(x\varphi_3\bar{x}) = e$, ou $e\varphi_6(x\varphi_3\bar{x}) = e$, enfin $x\varphi_3\bar{x} = e$. Mais φ_3 est défini plus haut par $(a\varphi b)\varphi_1 c = a\varphi_2(b\varphi_3 c)$, où $\forall\varphi_2, \varphi_3 \in \Phi$, $\exists\varphi, \varphi_1 \in \Phi$; donc $\bar{x}$ est aussi inverse à droite de x dans tous les groupoïdes et tout inverse à gauche est aussi inverse à droite.

(C) Soit $x\varphi_1\bar{x} = e$ et $x'\varphi_2 x = e$; on a $(x'\varphi_2 x)\varphi_1\bar{x} = x'\varphi_3(x\varphi_4\bar{x})$, ou $e\varphi_1\bar{x} = x'\varphi_3 e$; $\bar{x} = x'$, finalement $\forall\varphi_1, \varphi_2 \in \Phi$, $x\varphi_1 x' = x'\varphi_2 x = e$, et tout élément est permutable avec son inverse.

(D) Soit $x\varphi_1 y = z$ et y^{-1} l'inverse de y; alors $(x\varphi_1 y)\varphi_2 y^{-1} = x\varphi_3(y\varphi_4 y^{-1}) = x\varphi_3 e = x$. Ainsi $\forall\varphi_2 \in \Phi$, $z\varphi_2 y^{-1} = x$. Le produit de deux éléments quelconques z et y^{-1} étant égal à x, et par conséquent étant le même dans tous les groupoïdes $E(\varphi_2)$ il s'en suit que tous ces groupoïdes coïncident. Le système se réduit à un seul groupoïde, qui est un groupe.

8. Distributivité. DÉFINITION 8.1. *Un système* (E, Φ) *satisfait à la distributivité demosienne à droite* (généralisation de [22]), *si* $\forall x, y, z \in E$, $\forall \varphi_i, \varphi_j \in \Phi$, $\exists \varphi_k, \varphi_m, \varphi_p \in \Phi$, $(i, j, k, m, p = 1, 2, 3, 4, 5)$

$$(8.1) \qquad (x\varphi_1 y)\varphi_2 z = (x\varphi_3 z)\varphi_4(y\varphi_5 z) .$$

(Définition symétrique à gauche). *Le système est self-distributif si* $\varphi_3 = \varphi_5 = \varphi_2$, $\varphi_1 = \varphi_4$ *et si* $\forall \varphi_1, \exists \varphi_2$.

THÉORÈME 8.2. *Si un système* (E, Φ) *de quasigroupes à gauche est distributif demosien à droite,*

(i) *l'ensemble des réciproques de ces quasigroupes est encore distributif à droite,*

(ii) *l'ensemble des conjoints est distributif à gauche.*

Preuve. (i) Posons $x\varphi_3 z = c$, $y\varphi_5 z = d$, $c\varphi_4 d = b$, $x\varphi_1 y = a$, $a\varphi_2 z = b$. Sur le réciproque, en représentant les nouveaux signes opératoires par ψ, on aura $a\psi_1 y = x$, $b\psi_2 z = a$, $c\psi_3 z = x$, $d\psi_5 z = y$, $b\psi_4 d = c$. Egalant les expressions de x, on a $a\psi_1 y = c\psi_3 z$. Remplaçant a, y et c par leur valeur, $(b\psi_2 z)\psi_1(d\psi_5 z) = (b\psi_4 d)\psi_3 z$. Comme, par hypothèse, deux des φ déterminent les trois autres, il en est de même des ψ et le système (E, Ψ) est encore distributif demosien à droite.

(ii) En représentant par θ les opérations conjointes, on aura, sur le conjoint de (E, Φ), $(z\theta_5 y)\theta_4(z\theta_3 x) = z\theta_2(y\theta_1 x)$.

EQUATION FONCTIONNELLE 8.3. Hosszú ([23], p. 160), envisage l'équation fonctionnelle $F[G(x, y), z] = G[F(x, z), F(y, z)]$, où x, y, z appartiennent à un corps. On peut se proposer de trouver deux groupoïdes $E(\times)$ et $E(\cdot)$, définis sur un même ensemble quelconque E, de manière que

$$(8.2) \qquad \forall x, y, z \in E , \qquad (x \times y) \cdot z = (x \cdot z) \times (y \cdot z) .$$

Par exemple, l'un des groupoïdes étant arbitraire, l'équation sera vérifiée si l'autre satisfait à la loi de translation identique $xy = x$ ([35], p. 153, N°2, équ. 9).

THÉORÈME 8.4. *Etant donné un groupoïde arbitraire* $E(\times)$, *on obtient une solution de l'équation* (8.2) *en prenant pour chacune des translations* Δ_z *de* $E(\cdot)$ *un endomorphisme quelconque de* $E(\times)$.

Preuve. Considérons z comme une constante sur $E(\cdot)$ et soit

$$\Delta_z = (x \to f(x)) = (x \to x \cdot z)$$

la translation relative à z. L'équation (8.2) devient $f(x \times y) = f(x) \times f(y)$, ce qui définit un endomorphisme de $E(\times)$; ainsi chaque Δ_z est un endomorphisme de $E(\times)$.

EXEMPLE 8.5. Soit sur $Z/3$ le groupe $x \times y = x + y + 1$. Les séries de produits à droite ([33], p. 87), sont (00121102) et (2), les seuls automorphismes sont l'identique et la transposition (01). En se bornant aux quasigroupes à gauche on pourra prendre $\Delta = 1$ ou (01), ce qui donnera huit solutions $E(\cdot)$.

EXEMPLE 8.6. En prenant pour $E(\times)$ un quasigroupe automorphe par le groupe cyclique ([30], p. 321), on pourra prendre pour $E(\cdot)$ le groupe cyclique et chacun de ses isotopes de la forme $(1, \eta, 1)$, avec η arbitraire. Ainsi en prenant ([30], p. 325, ex. 15), on a $13 \times 7 = 11$; $(13 \times 7) \cdot 4 = 11 + 4 = 0$; $13.4 = 2$; $7.4 = 11$; $2 \times 11 = 0$.

EXEMPLE 8.7. En prenant pour $E(\times)$ un groupoïde "endo" ([32], p. 297, N°3), c'est-à-dire satisfaisant, sur un ensemble E de nombres réels, à $x \times y = z \Rightarrow xm \times ym = zm$, la solution $E(\cdot)$ sera fournie par le semigroupe multiplicatif de E. Ainsi, en prenant ([32], p. 298, Ex. II), on a $5 \times 4 = 2$; $2.3 = 6$ et $5.3 = 1$; $4.3 = 5$; $1 \times 5 = 6$.

COROLLAIRE 8.7. *L'ensemble des quasigroupes à gauche construits sur un même ensemble E satisfait à la distributivité demosienne restreinte*

$$(8.3) \quad \forall x, y, z \in E\,, \quad \forall \varphi_1 \in \Phi\,, \quad \exists \varphi_2 \in \Phi\,, \quad (x\varphi_1 y)\varphi_2 z = (x\varphi_2 z)\varphi_1(y\varphi_2 z)\,.$$

Preuve. On a vu, N°8.4, que, si $E(\varphi_1)$ est un groupoïde quelconque sur E, il existe toujours au moins un groupoïde $E(\varphi_2)$, satisfaisant (8.3) et dont les translations sont des endomorphismes de $E(\varphi_1)$. Si $E(\varphi_1)$ est un quasigroupe à gauche quelconque, les translations pourront être choisies parmi les automorphismes de $E(\varphi_1)$; ce seront alors des permutations de E et $E(\varphi_2)$ sera encore un quasigroupe à gauche.

9. Parastrophies. DÉFINITION 9.1. *Le i-parastrophique d'un système demosien (E, Φ) est le système dérivé du premier en soumettant tous les groupoïdes de (E, Φ) à la même i-parastrophie.* Ainsi le conjoint d'un système est le système formé par les conjoints de tous ses groupoïdes; le réciproque d'un système de quasigroupes à gauche est l'ensemble des réciproques de ces quasigroupes, etc. Si (E, Φ) ne contient que des quasigroupes, il pourra prendre six formes parastrophiques.

THÉORÈME 9.2. *Si (E, Ψ) est le conjoint de (E, Φ), il sera en même temps que lui, commutatif, associatif, inversible bilatère, self-distributif* *bilatère demosien.*

Preuve. Si (E, Φ) a la commutativité demosienne (N°3.1)

$$\forall x, y \in E , \quad \forall \varphi_1 \in \Phi , \quad \exists \varphi_2 \in \Phi , \quad x\varphi_1 y = y\varphi_2 x .$$

En désignant par ψ les opérations conjointes ([34], N°6),

$$\forall x, y \in E , \quad \forall \psi_1 \in \Psi , \quad \exists \psi_2 \in \Psi , \quad x\psi_2 y = y\psi_1 x ,$$

donc (E, Ψ) est encore commutatif demosien.

Si (E, Φ) est associatif demosien (N°7.1)

$$\forall x, y, z \in E , \quad \forall \varphi_1, \varphi_2 \in \Phi , \quad \exists \varphi_3, \varphi_4, \varphi_5, \varphi_6 ,$$
$$(x\varphi_1 y)\varphi_2 z = x\varphi_3(y\varphi_4 z); \quad x\varphi_1(y\varphi_2 z) = (x\varphi_5 y)\varphi_6 z .$$

Donc, sur le système conjoint $(z\psi_4 y)\psi_3 x = z\psi_2(y\psi_1 x)$, et $z\psi_6(y\psi_5 x) = (z\psi_2 y)\psi_1 x$. Ainsi (E, Ψ) est encore associatif demosien.

Le cas de l'inversibilité est analogue; celui de la distributivité résulte de 8.2. On rapprochera ce résultat de ([34], N°14–16). Plus généralement, on verrait que, si ($\dot{\text{I}}$) est une identité self-conjointe ([34], N°19.1), sur un groupoïde, les systèmes (E, Φ) et (E, Ψ) possèdent en même temps la propriété $\dot{\text{I}}$-demosienne.

THÉORÈME 9.3. *Si un système S de quasigroupes satisfait à une identité demosienne, le système dérivé de S en transformant tous les quasigroupes par une même parastrophie satisfait à une identité demosienne dérivée de la première par la même parastrophie.*

Preuve. Supprimons les indices dans les signes opératoires φ_i et ne conservons que les parenthèses, crochets etc, ce qui est évidemment légitime au point de vue formel. Dès lors les calculs qu'il faut effectuer pour trouver ce que devient la relation R demosienne en passant au i-parastrophique, sont précisément ceux que l'on ferait pour passer de la relation R sur un seul quasigroupe à sa i-parastrophique. Soit par exemple un système de quasigroupes à gauche satisfaisant à l'associativité demosienne; on trouve que le systéme réciproque satisfait à la trasitivité demosienne ([24], p. 203, [35], p. 156, N°2, iv, équ. 63). Or la transitivité usuelle est bien réciproque de l'associativité ([35], p. 154, N°2, équ. 25).

Ainsi, *tout système S de quasigroupes à gauche possédant la transitivité demosienne se déduit d'un système associatif S' en remplaçant chaque quasigroupe de S' par son réciproque.* Or S' se compose (ci dessus, Théorème 7.2, [37]), de quasigroupes isomorphes à un groupe G'. Donc S *se compose de quasigroupes isotopes au réciproque d'un* *groupe*. On comparera ce résultat à la solution directe donnée dans [38]. Si G est lui-même un groupe, alors il satisfait à la fois à $(xt)(yt) = xy$ et à l'associativité, donc en faisant $t = y$, on a $(xy)(yy) = xy$, $y^2 = 1$ et tous les les éléments de G sont du second ordre. On obtient le théorème ci-après 22.2. De même *si* (E, Φ) *est un système commutatif demosien de quasigroupes et si l'on remplace chaque quasigroupe par son réciproque, le nouveau système satisfera à la loi demosienne des keys à gauche* ([35], p. 153, N°2, équ. 5). Soit

$$x\varphi_1 y = z \rightleftarrows z\psi_1 y = x \text{ et } y\varphi_2 x = z \rightleftarrows z\psi_2 x = y\,;$$

alors

$$(\forall\varphi_1 \in \Phi\ ,\ \exists\varphi_2 \in \Phi\ ,\ y\varphi_2 x = x\varphi_1 y) \rightleftarrows (\forall\psi_1 \in \Psi\ ,\ \exists\psi_2 \in \Psi\ ,\ z\psi_1(z\psi_2 x) = x)\ .$$

Enfin les conditions exprimant qu'un système satisfait à la commutativité, ou à la loi des keys à gauche démosiennes, doivent être réciproques, au sens des parastrophies. Et en effet, la première (N°3.1) est que le système contienne, avec tout quasigroupe, son conjoint; la seconde (Th. 4.2, in fine), que le système contienne, avec tout quasigroupe, son parastrophique par

$$x \times y = z \leftrightarrows x \odot z = y\ . \tag{9.1}$$

Or si l'on a $xy = z$ et $yx = z$, cela devient, sur les réciproques $zy = x$ et $zx = y$, ce qui est bien la parastrophie (9.1).

COROLLAIRE 9.4. *Si une relation R est self-i-parastrophique et si un système $S(E, \Phi)$ de quasigroupes satisfait à la relation R-demosienne, le système i-parastrophique S' de S satisfera encore à la relation R-demosienne.*

Cela résulte immédiatement du théorème 9.3. Par exemple si un système de quasigroupes satisfait à l'entropie demosienne ([35], p. 155, N°2, équ. 38),

$$\forall x, y, z, u \in E\ , \qquad \forall\varphi_1, \varphi_2, \varphi_3 \in \Phi\ , \qquad \exists\varphi_4, \varphi_5, \varphi_6 \in \Phi\ ,$$
$$(x\varphi_1 y)\varphi_2(z\varphi_3 u) = (x\varphi_4 z)\varphi_5(y\varphi_6 u)\ ,$$

alors tous les parastrophiques de ce système satisferont à l'entropie demosienne, puisque tous les parastrophiques d'un quasigroupe entropique, comme le montre un calcul immédiat, sont eux-mêmes entropiques.

10. Immersion. Etant donné un système demosien (E, Φ) est il possible de le plonger dans un système admettant quelque loi demosienne determinée? La réponse à cette question dépend de la nature de cette loi. Il est clair que tout système peut être immergé dans un système demosien commutatif. D'après le Théorème 7.2, pour qu'un ensemble de quasigroupes puisse être un sous-ensemble d'un système associatif demosien il faut que chaque quasigroupe de cet ensemble puisse être obtenu comme isotope de quelque groupe par $x\xi \cdot y\theta = x\varphi_1 y$. Or ce que l'on sait, par exemple, des quasigroupes du 5° ordre, ([35], N°25) suffit à montrer qu'il existe des quasigroupes qui ne sont isotopes d'aucun groupe. Donc, *on ne peut plonger, en général, un système donné dans un système associatif demosien.* En revanche, il pourra être intéressant d'étudier le nucleus [27] formé par tous les $E(\varphi)$ d'un système demosien (E, Φ) qui satisfont localement à une loi déterminée.

11. Associateur. DÉFINITION 11.1. *L'associateur à gauche d'un système $S(E, \Phi)$ de groupoïdes est un sous-ensemble $A \subseteq E$, satisfaisant aux deux conditions*:

$$\forall a \in A\,, \quad \forall x, y \in E\,, \quad \forall \varphi_1, \varphi_2 \in \Phi\,, \quad \exists \varphi_3, \varphi_4 \in \Phi\,,$$
$$(a\varphi_1 x)\varphi_2 y = a\varphi_3(x\varphi_4 y)\,;$$
$$\forall a, b \in A\,, \quad \forall \varphi_1, \varphi_2 \in \Phi\,, \quad a\varphi_1 b = a\varphi_2 b\,.$$

THÉORÈME 11.2. *L'associateur à gauche d'un système demosien quelconque est un semigroupe,* (sous-groupoïde commun à tous les groupoïdes du système).

Preuve. Soient $a, b \in A$, alors

$$\forall x, y \in E\,, \quad \forall \varphi_1, \varphi_2 \in \Phi\,, \quad \exists \varphi_3, \varphi_4 \in \Phi\,,$$
$$(a\varphi_1 x)\varphi_2 y = a\varphi_3(x\varphi_4 y)\,; \quad (b\varphi_1 x)\varphi_2 y = b\varphi_3(x\varphi_4 y)\,.$$

Considérons le produit $a\varphi b$, $(\varphi \in \Phi)$. On aura

$$\begin{aligned}[(a\varphi b)\varphi_1 x]\varphi_2 y &= [a\varphi_3(b\varphi_4 x)]\varphi_2 y = a\varphi_5[(b\varphi_4 x)\varphi_6 y] = a\varphi_5[b\varphi_7(x\varphi_8 y)] \\ &= (a\varphi_9 b)\varphi_{10}(x\varphi_8 y) = (a\varphi b)\varphi_{10}(x\varphi_8 y)\,.\end{aligned}$$

Donc $a\varphi b \in A$ et A est fermé dans chacun des groupoïdes du système. D'ailleurs

$$(a\varphi_1 b)\varphi_2 c = a\varphi_3(b\varphi_4 c)\,, \quad \forall a, b, c \in A\,, \quad \forall \varphi_1, \varphi_2, \varphi_3\, \varphi_4 \in \Phi;$$

donc, en particulier $(a\varphi b)\varphi c = a\varphi(b\varphi c)$ et A est un semigroupe.

12. Multigroupoïde. DÉFINITION 12.1. *Un multigroupoïde* ([37], [12], p. 183), *est un ensemble, E, muni d'une loi de composition* $(\ast)$, *faisant correspondre à tout couple ordonné $x, y \in E$, un sous-ensemble non vide, $x \ast y = (a, b, c \cdots) \subseteq E$.*

THÉORÈME 12.2. (i) *Tout système demosien S définit un multi-*

groupoïde, M, dans lequel le produit $x * y$ *est déterminé comme étant l'ensemble* $(x\varphi_1 y, x\varphi_2 y, \cdots, x\varphi_i y, \cdots)$ *où* φ_i *décrit* Φ.

(ii) *Pour que M soit associatif il faut et il suffit que le système S ait l'associativité demosienne.*

(iii) *Pour que M soit abélien il faut et il suffit que S ait la commutativité demosienne.*

Preuve. Si M est associatif, on a $(x * y) * z = x * (y * z)$. Donc

(12.1) $\forall i, j, \exists k, m$ et $\forall k, m \exists i, j$, $\quad (x\varphi_i y)\varphi_j z = x\varphi_k(y\varphi_m z)$

et S possède l'associativité demosienne. Réciproquement, si le système S est associatif demosien, (12.1) est vérifiée, donc

$$(x * y) * z \subseteq x * (y * z) \subseteq (x * y) * z \text{ et } (x * y) * z = x * (y * z) .$$

La preuve est analogue pour la commutativité. Plus généralement, le multigroupoïde M satisfera à une identité $(\dot{I})$ en même temps que le système S vérifiera l'identité $(\dot{I})$-démosienne.

CHAPITRE II

INTERACTION DES LOIS DEMOSIENNES

13. **Question.** Lorsqu'un groupoïde G satisfait à une ou plusieurs identités, $I, I', \cdots$, il est bien connu ([25], p. 205, équ. 2, $a(bc) = c(ba)$ entraîne l'entropie; [28], $ab = ba$, $(ab)x = (ax)(bx)$, sur un ensemble ordonné à multiplication monotone, entraîne l'entropie; [35], p. 161, N°8, IV, keys et commutativité, N°35 & 36), que, dans certains cas, il existe une nouvelle identité J, conséquence des $I, I', \cdots$, et qui est encore vérifiée sur G. Si un système (E, Φ) satisfait à diverses lois demosiennes, $I, I', \cdots$, cette situation entraîne-*t*-elle, comme dans le cas d'un seul groupoïde, l'existence d'une nouvelle loi demosienne, J, conséquence des $I, I', \cdots$, qui soit encore vérifiée sur (E, Φ)? Dans l'affirmative, la relation $J = f(I, I', \cdots)$ est-elle la même pour un groupoïde isolé et pour un système (E, Φ)? La réponse à cette question ne peut pas être formulée d'une manière générale. Il peut arriver (N°14.2–14.3) qu'elle dépende de conditions supplémentaires à imposer au système et même que le transfert d'une implication du cas uniforme au cas demosien ne soit pas possible. Ainsi, contrairement à ce qui se passe pour un quasigroupe isolé ([8], p. 112), la self-distributivité demosienne n'entraîne pas l'idempotence, comme le montre l'exemple 7.5. Les quasigroupes $x\varphi y = ax + by + c$ ne sont pas idempotents et ils satisfont néanmoins à l'identité demosienne $\forall\varphi_1, \exists\varphi_2, x\varphi_1(y\varphi_z z) = (x\varphi_1 y)\varphi_2(x\varphi_1 z)$. La solution est $x\varphi_2 y = px + (1 - p)y$.

De même, le théorème de Knaster [28], à savoir qu'un quasigroupe abélien self distributif G, muni d'une relation d'ordre total et d'une multiplication strictement monotone ([14], p. 306) est entropique, ne se laisse pas transmettre au cas des systèmes demosiens. Observons ici que, pour la validité du théorème, la condition que G est un quasigroupe peut être affaiblie et remplacée par cancellabilité de G, condition qui est d'ailleurs entraînée par la stabilité de la relation d'ordre; l'idempotence, puis l'entropie en résultent ensuite, sans considération de continuité. On peut poser cette question:

Le seul systéme demosien sur lequel le théorème de Knaster puisse être transféré se réduit-il à un groupoïde unique?

Dans ce qui suit nous nous bornerons à l'étude de quelques lois particulières.

14. Commutativité-Associativité-Inversibilité.[4] THÉORÈME 14.1. *Si* (E, Φ) *est commutatif demosien, pour qu'il ait l'associativité demosienne il faut et il suffit qu'il soit inversible demosien.* ([41] dans le cas fini).

Preuve. (i) Si le système est associatif et commutatif, on aura

$$\forall x, y, z \in E\ , \qquad \forall \varphi_1, \varphi_2 \in \Phi\ , \qquad (x\varphi_1 y)\varphi_2 z = x\varphi_3(y\varphi_4 z) = (z\varphi_5 y)\varphi_6 x\ ,$$
$$x\varphi_1(y\varphi_2 z) = (x\varphi_7 y)\varphi_8 z = z\varphi_9(y\varphi_{10} x)\ ,$$

où $\varphi_3, \varphi_4, \cdots, \varphi_{10}$ sont des éléments déterminés de Φ. Ainsi (E, Φ) est inversible des deux côtés.

(ii) Réciproquement, si (E, Φ) est inversible et commutatif on aura

$$\forall x, y, z \in E\ , \qquad \forall \varphi_1, \varphi_2 \in \Phi\ , \qquad \exists \varphi_5, \varphi_6, \cdots, \varphi_{10}\ ,$$
$$(x\varphi_1 y)\varphi_2 z = z\varphi_3(y\varphi_4 x) = x\varphi_5(y\varphi_6 z)\ ; \quad x\varphi_1(y\varphi_2 z) = (z\varphi_7 y)\varphi_8 x = (x\varphi_9 y)\varphi_{10} z$$

et le système sera associatif demosien.

THÉORÈME 14.2. *Si un semigroupe inversible* S (au sens du N°6), *satisfait à l'une des conditions suivantes,*

(i) *S est homogène,*
(ii) *S est diagonal,*
(iii) *S est un quasigroupe à droite,*
(iv) *S est idempotent,*
(v) *S a un élément neutre bilatère,*

alors il est abélien.

Preuve. (i) S étant homogène, tout élément x peut être obtenu comme un produit $x = ay$; à son tour, y peut être mis sous la forme $y = bc$, d'où $x = abc$. On a alors, z étant un élément quelconque,

$$xz = abcz = (ab)(c)(z) = (z)(c)(ab) = z(ca)b = b(ca)z$$
$$= (bc)(a)(z) = (z)(a)(bc) = z(abc) = zx\ , \qquad \forall x, z\ .$$

(ii) $\forall x, y$, $xyy = yyx$, ou $x(yy) = (yy)x$. On tire d'abord de là que, dans tout semigroupe inversible (au sens du N°6), *l'ensemble des* (xx), c'est-à-dire la diagonale, *est contenu dans le centre*. Si S est diagonal, cet ensemble coïncide avec S et S est abélien.

(iii) On a $\forall x, y$, $(xy)(xy) = (x)(y)(xy) = (xy)(y)(x)$, cancellant par xy à gauche, $xy = yx$.

(iv) résulte de (ii) car l'idempotence entraîne la diagonalité.

(v) Si u est l'unité, on a $xyu = uyx$, ou $xy = yx$.

Chacune de ces conditions est d'ailleurs une conséquence de la première car tout groupoïde diagonal, ou avec quotient à droite, ou idempotent, ou avec unité bilatère est nécessairement homogène. Elles sont toutes suffisantes, mais aucune n'est nécessaire, comme le montre l'exemple du semigroupe additif, N^+, des entiers naturels, qui n'est pas homogène et qui est tout de même abélien. Quelques-unes des conditions précédentes se laissent transférer aux systèmes demosiens; ce sont (ii), (iv) et (v).

THÉORÈME 14.3. *Si tous les groupoïdes d'un système* (E, Φ) *possédant l'associativité et l'inversibilité demosiennes, satisfont à l'une des conditions suivantes*: (ii) *ils sont diagonaux et ont la même diagonale,* $(\forall i, j,\ x\varphi_i x = x\varphi_j x)$, (iv) *ils sont idempotents,* (v) *ils ont une unité bilatère commune, alors* (E, Φ) *aura la commutativité demosienne.*

Preuve. (ii) Par hypothèse $\forall x, y, z \in E$, $\forall \varphi_1, \varphi_2 \in \Phi$, $\exists \varphi_3, \varphi_4, \cdots \in \Phi$, $(x\varphi_1 y)\varphi_2 z = x\varphi_3(y\varphi_4 z) = z\varphi_5(y\varphi_6 x)$. Si $x = y$, $x\varphi_1 x = x\varphi_6 x = t$ et $t\varphi_2 z = z\varphi_5 t$. Comme tous les groupoïdes sont diagonaux, $\forall t$, $\exists x$, $x\varphi x = t$, donc (E, Φ) a la commutativité demosienne.

(iv) résulte de (ii).

(v) Si u est élément neutre dans tous les groupoïdes, on a $\forall x, y \in E$, $\forall \varphi_1, \varphi_2 \in \Phi$, $\exists \varphi_3, \varphi_4, \cdots \in \Phi$, $(x\varphi_1 y)\varphi_2 u = x\varphi_3(y\varphi_4 u) = u\varphi_5(y\varphi_6 x)$, d'où $x\varphi_1 y = y\varphi_6 x$.

QUESTION 14.4. *Quelle est la condition nécessaire et suffisante pour que, sur un système* (E, Φ), *l'associativité et l'inversibilité demosiennes entraînent la commutativité demosienne*?

15. Distributivité. THÉORÈME 15.1. *Tout système* (E, Φ) *associatif et commutatif demosien, dont tous les groupoïdes sont idempotents, possède la distributivité demosienne.*

Preuve. Par hypothèse $\forall x, y, z \in E$, $\forall \varphi_3, \varphi_4, \varphi_5 \in \Phi$, $\exists \varphi_1, \varphi_2, \varphi_6, \cdots \in \Phi$, $(x\varphi_3 z)\varphi_4(y\varphi_5 z) = x\varphi_6[z\varphi_7(y\varphi_5 z)] = x\varphi_6[z\varphi_7(z\varphi_8 y)] = x\varphi_6[(z\varphi_9 z)\varphi_{10} y] = x\varphi_6(z\varphi_{10} y) = x\varphi_6(y\varphi_{11} z) = (x\varphi_1 y)\varphi_2 z$.

16. Keys, transitivité, eingewandter Produkt. THÉORÈME 16.1. *Sur un système (E, Φ) de quasigroupes à gauche, de ces trois lois:*

(i) *la loi demosienne des keys à droite* ([35], p. 153, N°2, équ. 4) & N°4,

$$(16.1) \qquad \forall z, y \in E\,, \quad \forall \varphi_5 \in \Phi\,, \quad \exists \varphi_6 \in \Phi\,, \quad (z\varphi_5 y)\varphi_6 y = z\,;$$

(ii) *la transitivité demosienne* ([35], p. 154, N°2, équ. 25) & N° 9.3, supra,

$$(16.2) \qquad \forall x, y, z \in E\,, \quad \forall \varphi_1, \varphi_2 \in \Phi\,, \quad \exists \varphi_3, \varphi_4\,, \quad (x\varphi_1 y)\varphi_2(z\varphi_3 y) = x\varphi_4 z\,.$$

(iii) *le "eingewandte Produkt"* ([35], p. 154, N°2, équ. 22)

$$(16.3) \qquad \forall x, y, t \in E\,, \ \forall \varphi_1, \varphi_2 \in \Phi, \ \exists \varphi_3, \varphi_5 \in \Phi, \ (x\varphi_1 y)\varphi_2 t = x\varphi_4(t\varphi_5 y)\,,$$

chacune des deux dernières est entraînée par les deux autres.

Preuve. A. Soit un système (E, Φ) satisfaisant (i) et (ii). La première (N°4), exprime que les quasigroupes du système sont deux à deux réciproques; donc $\forall \varphi_3$, $\exists \varphi_7$, $z\varphi_3 y = t \rightleftarrows t\varphi_7 y = z$. Alors (16.2) devient, $\forall x, y, t \in E$, $\forall \varphi_1, \varphi_2 \in \Phi$, $\exists \varphi_4, \varphi_7 \in \Phi$, $(x\varphi_1 y)\varphi_2 t = x\varphi_4(t\varphi_7 y)$, ce qui est (iii). Ainsi (i) $\cap$ (ii) $\Rightarrow$ (iii).

B. Supposons (i) et (iii) vérifiées. Comme ci-dessus, (i) exprime que $\forall t, y \in E$, $\forall \varphi_3 \in \Phi$, $\exists \varphi_5 \in \Phi$, $z \in E$, $t = z\varphi_3 y \leftrightarrows t\varphi_5 y = z$; Dès lors (iii) devient $\forall x, y, z \in E$, $\forall \varphi_1, \varphi_2 \in \Phi$, $\exists \varphi_3, \varphi_4, \varphi_5, \cdots \in \Phi$, $(x\varphi_1 y)\varphi_2(z\varphi_3 y) = x\varphi_4 z$. Ainsi (i) $\cap$ (iii) $\Rightarrow$ (ii).

17. Entropie. THÉORÈME 17.1 ([25], p. 207, Th. 1, dans le cas d'un seul groupoïde). *Tout système (E, Φ) inversible demosien est entropique* ([35], p. 155, N°2, équ. 38).

Preuve. Par hypothèse $\forall x, y, z \in E$, $\forall \varphi_1, \varphi_2 \in \Phi$, $\exists \varphi_3, \varphi_4 \in \Phi$, $x\varphi_1(y\varphi_2 z) = z\varphi_3(y\varphi_4 x)$. En appliquant itérativement celle-ci $(x\varphi_1 y)\varphi_2(z\varphi_3 t) = t\varphi_4[z\varphi_5(x\varphi_1 y)] = t\varphi_4[y\varphi_6(x\varphi_7 z)] = (x\varphi_7 z)\varphi_8(y\varphi_9 t)$. Les membres extrêmes sont ceux de l'identité d'entropie demosienne. (cf. [26] p. 55.)[1]

18. Commutativité et loi des keys. THÉORÈME 18.1. *Sur un système (E, Φ) de quasigroupes, deux quelconques des lois demosiennes suivantes entraînent la troisième;*

(i) *commutativité,*
(ii) *loi des keys à droite,*
(iii) *loi des keys à gauche.*

Preuve. L'implication (i) $\cap$ (ii) $\Rightarrow$ (iii) a déjà été démontrée ([35], p. 161, N°8.4). Par raison de symétrie, (i) $\cap$ (iii) $\Rightarrow$ (ii) en résulte. Pour établir la dernière, écrivons l'hypothèse sous la forme (ii) $\forall x, y \in E$, $\forall \varphi_1 \in \Phi$, $\exists \varphi_2 \in \Phi$, $(x\varphi_2 y)\varphi_1 y = x$, (iii) $\forall x, y \in E$, $\forall \varphi_2 \in \Phi$, $\exists \varphi_3 \in \Phi$, $x\varphi_3(x\varphi_2 y) = y$. Si l'on choisit arbitrairement y et $z \in E$, puisque les éléments de Φ sont des lois de quasigroupes, l'équation $x\varphi_2 y = z$ a une solution et une seule en x, donc $\forall y, z \in E$, $\forall \varphi_1 \in \Phi$, $\exists \varphi_2 \in \Phi$, $x \in E$, $x\varphi_2 y = z$ et $z\varphi_1 y = x$; $\forall y, z \in E$, $\forall \varphi_1 \in \Phi$, $\exists \varphi_3 \in \Phi$, $x\varphi_3 z = y$. De plus, par hypothèse, $\forall x, y, z \in E$, $\forall \varphi_3 \in \Phi$, $\exists \varphi_4 \in \Phi$, $(x\varphi_3 z)\varphi_4 z = x$. En remplaçant $x\varphi_3 z$ par y, cela devient $y\varphi_4 z = x$, donc $z\varphi_1 y = x = y\varphi_4 z$. Ainsi $\forall y, z \in E$, $\forall \varphi_1 \in \Phi$, $\exists \varphi_4 \in \Phi$, $z\varphi_1 y = y\varphi_4 z$, ce qui est la commutativité demosienne.

QUESTION 18.2. Sur un groupoïde G les identités (α) et (β) entraînent l'identité (γ). Sur l'ensemble (E, Φ) et en particulier sur le système obtenu en soumettant G à toutes les isotopies de quelque groupe d'isotopies, quelle est la condition pour que les identités demosiennes (A), (B) et (C), induites par (α), (β) et (γ) satisfassent à (A) $\cap$ (B) $\Rightarrow$ (C)?

CHAPITRE III

ISOTOPIE. SYSTÈMES (G, K, T)

19. Généralités. QUESTION 19.1. *Un groupoïde G satisfait à une identité* (L); *on soumet G à toutes les isotopies d'un groupe* (d'isotopie); *on obtient un système de groupoïdes. Quelle est la condition pour que ce système satisfasse à l'identité demosienne induite par* (L)?

Comme au Chapitre II, il n'est pas possible de donner à ce problème une solution générale. Parfois, comme dans le cas de la commutativité, la condition se réduit à quelqu' égalité évidente: "Si un groupoïde G est abélien, le système dérivé de G en le soumettant à toutes les isotopies du groupe (X, X, Z), où X et Z sont deux groupes de permutations quelconques de l'ensemble G, est évidemment commutatif demosien". Mais le plus souvent l'existence de la loi demosienne induite dépend de conditions moins évidentes et peut même être hors de cause. Il est un cas où la question peut recevoir une solution complète, c'est celui où G est un groupe, les composantes des isotopies étant des translations de G. De telles isotopies (ξ, η, ζ), $x\xi \times \eta y = xy\zeta$, ne dépendent en réalité que de deux paramètres et peuvent se mettre sous la forme $x \times y = x\xi y\theta$, $\xi, \theta \in G$. Parmi ces isotopies considérons seulement celles qui dérivent de deux sous-groupes $K, T \subseteq G$, avec $\xi \in K$, $\theta \in T$. Les isotopes ainsi définis sont évidemment des quasigroupes. Désignons un tel quasigroupe par $G(\xi, \theta)$.

DÉFINITION 19.2. *Un système (G, K, T), où K et T sont deux sous-groupes du Groupe G, est l'ensemble des quasigroupes $G(\xi, \theta)$, isotopes* *de G par les isotopies* $x \times y = x\xi y\theta$, *où* $\xi \in K$, $\theta \in T$. (Cf. [39].)

THÉORÈME 19.3. *La condition nécessaire et suffisante pour que* $G(\xi, \theta)$ *soit un groupe est que* θ *soit dans le centre de G. L'unité de ce groupe est alors* $u = \xi^{-1}\theta^{-1}$. *Si* $\theta \notin \mathscr{Z}_G$, *le quasigroupe* $G(\xi, \theta)$ *a seulement une unité à droite, u, (et n'est pas associatif.)*

Preuve. La condition d'associativité est $\forall x, y, z \in G$, $x\xi y\theta\xi z\theta = x\xi y\xi z\theta\theta$, ou $\forall z$, $\theta\xi z = \xi z\theta$. Si $z = 1$ (unité de G), $\theta\xi = \xi\theta$. Ainsi θ et ξ sont permutables, donc $\theta\xi z = \xi\theta z = \xi z\theta$, et, en cancellant par ξ, $\forall z$, $\theta z = z\theta$, ce qui exprime que θ est central dans G. Cette condition contient la précédente; elle est visiblement suffisante.

Soit u une unité à droite de $G(\xi, \theta)$, donc $x \times u = x\xi u\theta = x$, et $u = \xi^{-1}\theta^{-1}$. Pour qu'il existe une unité à gauche v, il faut que $v \times x = v\xi x\theta = x$, où v est indépendant de x. En faisant $x = 1$, $v = \theta^{-1}\xi^{-1}$, d'où $\theta^{-1}\xi^{-1}\xi x\theta = x$, $x\theta = \theta x$ et $\theta \in \mathscr{Z}_G$. Ainsi l'unité bilatère n'existe que si θ est central, c'est-a-dire si $G(\xi, \theta)$ est associatif. Si θ n'est pas dans le centre de G, le quasigroupe $G(\xi, \theta)$ a seulement une unité à droite.

20. **Commutativité.** THÉORÈME 20.1. *Pour que le système* (G, K, T) *ait la commutativité demosienne il faut et il suffit que G soit abélien.* (Cf N°3, 7, 8, 9)

Preuve. Que le système ait la commutativité demosienne s'écrit $\forall x, y \in G$, $\forall \xi \in K$, $\theta \in T$, $\exists \xi_1 \in K$, $\theta_1 \in T$, $x\xi y\theta = y\xi_1 x\theta_1$. Si $x=y=1$, $\xi\theta = \xi_1\theta_1$. Si $x = 1$, $\xi y\theta = y\xi_1\theta_1 = y\xi\theta$, d'où, en cancellant par θ, $\xi y = y\xi$. Ainsi tout élément $\xi \in K$ est central. Donc $x\xi y\theta = xy\xi\theta = y\xi_1 x\theta_1 = yx\xi_1\theta_1 = yx\xi\theta$. Cancellant par $\xi\theta$, on a $xy = yx$. La condition est évidemment suffisante. Cette conclusion n'est pas en contradiction avec le Corollaire 3.9, car, ici, le groupe G n'est soumis qu'à une partie de toutes les isotopies possibles.

21. **Associativité.** THÉORÈME 21.1. *Tout système* (G, K, T) *est associatif demosien.*

Preuve. L'associativité demosienne est exprimée par $\forall x, y, z \in G$, $\forall \xi$, $\xi_1 \in K$, θ, $\theta_1 \in T$, $\exists \xi_2$, $\xi_3 \in K$, θ_2, $\theta_3 \in T$.

$$(x\xi y\theta)\xi_1 z\theta_1 = x\xi_2(y\xi_3 z\theta_3)\theta_2 . \tag{21.1}$$

Si $y = z = 1$, en cancellant par x,

$$\xi\theta\xi_1\theta_1 = \xi_2\xi_3\theta_3\theta_2 . \tag{21.2}$$

Si $y = 1$,

$$\xi_3^{-1}\xi_2^{-1}\xi\theta\xi_1 z = z\theta_3\theta_2\theta_1^{-1} .$$

Donc $\theta_3\theta_2\theta_1^{-1}$ est dans le centre de G. Si $z = 1$, $y\theta\xi_1\theta_1\theta_2^{-1}\theta_3^{-1}\xi_3^{-1} = \xi^{-1}\xi_2 y$ et $\xi^{-1}\xi_2 \in \mathscr{Z}_G$ Comme $\mathscr{Z}_G$, K et T sont des sous groupes de G, leurs intersections ne sont jamais vides, Soit $t \in T \cap \mathscr{Z}$ et $k \in K \cap \mathscr{Z}$, arbitraires. On prendra $\theta_3\theta_2\theta_1^{-1} = t$, $\theta_3 = t\theta_1\theta_2^{-1}$, avec θ_2 arbitraire dans T. Ensuite $\xi^{-1}\xi_2 = k$ donne $\xi_2 = \xi k$. Enfin ξ_3 est déterminé par la première condition (21.2), $\xi_3 = \xi_2^{-1}\xi\theta\xi_1\theta_1\theta_2^{-1}\theta_3^{-1} = k^{-1}\theta\xi_1 t^{-1}$. Cela fait, l'égalité (21.1) est sûrement vérifiée car

$$x\xi_2(y\xi_3 z\theta_3)\theta_2 = x\xi k y k^{-1}\theta\xi_1 t^{-1} z t\theta_1\theta_2^{-1}\theta_2 = x\xi y k k^{-1}\theta\xi_1 z t^{-1} t\theta_1 = x\xi y\theta\xi_1 z\theta_1 .$$

Cela est conforme au Théorème 7.2, toutefois, si l'on veut que (21.1) ait d'autres solutions que la solution évidente $\xi = \xi_2$; $\xi_3 = \theta\xi_1$; $\theta_3 = \theta_1\theta_2^{-1}$, il faudra supposer $k, t \neq 1$, c'est-à-dire $T \cap \mathscr{Z}\setminus 1 \neq \phi$ et $K \cap \mathscr{Z}\setminus 1 \neq \phi$.

22. Demi-symétrie, keys, transitivité. LEMME 22.1. *S étant un groupe quelconque et G un de ses sous-groupes, on considère le complexe maximum $A \subseteq S$ satisfaisant à $\forall x \in G$, $\forall \alpha \in A$, $x\alpha x = \alpha$* ([35], p. 153, N°2, équ. 11), [27]. *Alors,* (i) *A est vide ou G est abélien,* (ii) *Si $G \neq (C_2)^n$ et si $\mathscr{Z}$ est le centralisateur de G dans S* ([20], p. 470), *on a $\forall \alpha \in A$, $A = \mathscr{Z}\alpha = \alpha\mathscr{Z}$, et $\mathscr{Z} + A$ est un groupe.*

Preuve. (i) Si $A \neq \phi$, soient $x, y \in G$ et $\alpha \in A$, $x\alpha x = \alpha$, et $y\alpha y = \alpha$, donc $x(y\alpha y)x = \alpha$, ou $(xy)\alpha(yx) = \alpha$. Mais $xy \in G$, donc $(xy)\alpha(xy) = \alpha$. En comparant, $xy = yx$.

(ii) Si $\forall x \in G$, $xz = zx$, ($z \in \mathscr{Z}$), alors $xz\alpha x = zx\alpha x = z\alpha$, donc $z\alpha \in A$. Réciproquement, soient $\alpha, \beta \in A$; alors $\forall x \in G$, $x\beta x = \beta$; mais $x^{-1} \in G$, donc $x^{-1}\alpha x^{-1} = \alpha$, d'où $x\beta x x^{-1}\alpha x^{-1} = \beta\alpha = x\beta\alpha x^{-1}$ ou $(\beta\alpha)x = x(\beta\alpha)$ et $\beta\alpha \in \mathscr{Z}$. D'ailleurs $\alpha \in A \rightleftarrows \alpha^{-1} \in A$ car $x\alpha x = \alpha \rightleftarrows x\alpha x\alpha^{-1} = 1 \rightleftarrows x\alpha^{-1} = \alpha^{-1}x^{-1} \rightleftarrows x\alpha^{-1}x = \alpha^{-1}$. Donc $\beta\alpha^{-1} \in \mathscr{Z}$; $\beta\alpha^{-1} = z$, $\beta = z\alpha$. On a donc $\mathscr{Z}\alpha \subseteq A \subseteq \mathscr{Z}\alpha$, d'où $A = \mathscr{Z}\alpha$. Enfin $\beta = z\alpha \rightleftarrows \beta^{-1} = \alpha^{-1}z^{-1}$, et comme $z^{-1} \in \mathscr{Z}$ puisque $\mathscr{Z}$ est un groupe, $A = \alpha^{-1}\mathscr{Z}$, $A = \alpha\mathscr{Z}$. Les égalités $\mathscr{Z}A = A\mathscr{Z} = A$, $AA \subseteq \mathscr{Z}$ montrent que $A + \mathscr{Z}$ est fermé dans S, et comme A contient, avec tout élément α, son inverse α^{-1}, le complexe $A + \mathscr{Z}$ est un sous-groupe de S. Le centralisateur $\mathscr{Z}$ de G est diviseur normal dans $\mathscr{Z} + A$ et le quotient $(\mathscr{Z} + A)/\mathscr{Z}$ est le groupe du second ordre. Pour construire A il suffit de multiplier $\mathscr{Z}$ par une solution particulière en α. Si $x, y \in G$ et si α satisfait à $x\alpha x = \alpha$, $y\alpha y = \alpha$, on aura, puisque G est alors abélien, $xy\alpha xy = xy\alpha yx = x\alpha x = \alpha$. Par suite, pour que α satisfasse $\forall x \in G$, $x\alpha x = \alpha$, il suffit qu'il satisfasse à cette condition quand x parcourt les générateurs de G. En particulier, si G est cyclique, ($G = C = \{c\}$), on pourra prendre pour α la solution du second ordre $\alpha = (C \to C^{-1})$ et comme dans ce cas G est son propre centralisateur, dans le groupe symétrique $\mathscr{S}_c$, $A = C\alpha$. Exemple $C = \{0123\} = (0123), (02)(13), (0321), (0)(1)(2)(3)$. C est son propre centralisateur dans le groupe symétrique $\mathscr{S}(0, 1, 2, 3)$. En choisissant $\alpha = \begin{pmatrix}0123\\0321\end{pmatrix} = (13)$, on pourra prendre $A = C(13) = (03)(12), (02), (01)(23), (13)$.

Si un élément de A est involutif, tous seront du second ordre car $\alpha(x\alpha x) = \alpha\alpha = (\alpha x)(\alpha x)$. Si $\alpha\alpha = 1$, $(\alpha x)(\alpha x) = 1$.

On pourrait appeler A un "anti-centre". Nous ne hasarderons pas ce néologisme.

THÉORÈME 22.2. *Pour que le système* (G, K, T) *satisfasse*

(i) *à la demi-symétrie demosienne* (N°5),

(ii) *ou à la loi demosienne des keys à droite* (N°4),

(iii) *ou à la transitivité demosienne* (N°9.3), *il faut et il suffit que tous les éléments de G soient involutifs.*

Preuve. (iii) (La démonstration est analogue pour les trois parties; démontrons seulement la dernière) La condition $\forall\varphi_1, \varphi_2 \exists\varphi_3, \varphi_4, (x\varphi_1 y)\varphi_2(z\varphi_3 y) = x\varphi_4 z$ prend la forme

$$m = x\xi_1 y\theta_1\xi_2 z\xi_3 y\theta_3\theta_2 = x\xi_4 z\theta_4 .$$

En faisant $y = z = 1$, on a $\xi_1\theta_1\xi_2\xi_3\theta_3\theta_2 = \xi_4\theta_4$, d'où $\theta_1\xi_2\xi_3 = \xi_1^{-1}\xi_4\theta_4\theta_2^{-1}\theta_3^{-1}$. Si $z = 1$, on a $\xi_1 y\theta_1\xi_2\xi_3 y\theta_3\theta_2 = \xi_4\theta_4$, ou $y\theta_1\xi_2\xi_3 y = \xi_1^{-1}\xi_4\theta_4\theta_2^{-1}\theta_3^{-1} = \theta_1\xi_2\xi_3$, donc $\theta_1\xi_2\xi_3 = \alpha$ tel que $y\alpha y = \alpha$. Par suite (Lemme 22.1), G est abélien et comme $\alpha \in G$, $yy\alpha = \alpha$, $yy = 1$. Tous les éléments de G sont involutifs. Cette condition est suffisante car, si elle est vérifiée, $m = x\xi_1 y\theta_1\xi_2 z\xi_3 y\theta_3\theta_2 = xzyy\xi_1\theta_1\xi_2\xi_3\theta_3\theta_2$. On choisit arbitrairement ξ_1, ξ_2, ξ_3, θ_1, θ_2, θ_3, θ_4 et on détermine ξ_4 par $\xi_1\theta_1\xi_2\xi_3\theta_3\theta_2 = \xi_4\theta_4$; alors $m = xz\xi_4\theta_4 = x\xi_4 z\theta_4$, et la condition du début est vérifiée.

23. Loi de Moufang. THÉORÈME 23.1. *La condition nécessaire et suffisante pour que le système* (G, K, T) *satisfasse à la loi demosienne de Moufang* ([35], p. 154, N°2, équ. 29), *est que les intersections de K et T avec le centre de G ne se réduisent pas à l'identique,* $K \cap \mathscr{Z} \setminus 1 \neq \phi$, $T \cap \mathscr{Z} \setminus 1 \neq \phi$.

Preuve. La condition $[x\varphi_1(y\varphi_2 x)]\varphi_3 z = x\varphi_4[y\varphi_5(x\varphi_6 z)]$ s'écrit $\forall x, y, z \in G$, $\forall\xi_1, \xi_2, \xi_3 \in K$, $\forall\theta_1, \theta_2, \theta_3 \in T$, $\exists\xi_4, \xi_5, \xi_6, \theta_4, \theta_5, \theta_6$,

$$[x\xi_1(y\xi_2 x\theta_2)\theta_1]\xi_3 z\theta_3 = x\xi_4[y\xi_5(x\xi_6 z\theta_6)\theta_5]\theta_4 . \tag{23.1}$$

Elle devient, si $y = z = 1 = x$,

$$\xi_1\xi_2\theta_2\theta_1\xi_3\theta_3 = \xi_4\xi_5\xi_6\theta_6\theta_5\theta_4 . \tag{23.2}$$

En faisant seulement $x = y = 1$, $\xi_1\xi_2\theta_2\theta_1\xi_3 z\theta_3 = \xi_4\xi_5\xi_6 z\theta_6\theta_5\theta_4$, ou

$$\xi_6^{-1}\xi_5^{-1}\xi_4^{-1}\xi_1\xi_2\theta_2\theta_1\xi_3 z = z\theta_6\theta_5\theta_4\theta_3^{-1} .$$

Ainsi $\theta_6\theta_5\theta_4\theta_3^{-1}$ est dans le centre. En faisant dans (23.1) $x = z = 1$, on a $\xi_1 y\xi_2\theta_2\theta_1\xi_3\theta_3 = \xi_4 y\xi_5\xi_6\theta_6\theta_5\theta_4$, ou $y\xi_2\theta_2\theta_1\xi_3\theta_3\theta_4^{-1}\theta_5^{-1}\theta_6^{-1}\xi_6^{-1}\xi_5^{-1} = \xi_1^{-1}\xi_4 y$; donc $\xi_1^{-1}\xi_4$ est central.

Réciproquement, si K et T ont des éléments centraux, soient t, k, k', trois d'entre eux, $t \in T$; $k, k' \in K$. Choisissons arbitrairement ξ_1, ξ_2, ξ_3, θ_1, θ_2, θ_3, θ_4, θ_5 et prenons $\theta_6 = t\theta_3\theta_4^{-1}\theta_5^{-1}$, $\xi_4 = \xi_1 k$, enfin $\xi_5 = \xi_4^{-1}\xi_1\xi_2 k' = k^{-1}\xi_2 k'$. La Condition (23.2) donne alors $\xi_6 = \xi_5^{-1}\xi_4^{-1}\xi_1\xi_2\theta_2\theta_1\xi_3\theta_3\theta_4^{-1}\theta_5^{-1}\theta_6^{-1} = k'^{-1}\xi_2^{-1}kk^{-1}\xi_1^{-1}\xi_1\xi_2\theta_2\theta_1\xi_3\theta_3\theta_4^{-1}\theta_5^{-1}\theta_5\theta_4\theta_3^{-1}t^{-1} = k'^{-1}\theta_2\theta_1\xi_3 t^{-1}$. Si l'on substitue ces valeurs dans (23.1) elle devient $x\xi_1 y\xi_2 x\theta_2\theta_1\xi_3 z\theta_3 = x\xi_1 kyk^{-1}\xi_2 k'xk'^{-1}\theta_2\theta_1\xi_3 t^{-1}zt\theta_3$, ou, en tenant compte de la permutabilité de t, k et k', $y\xi_2 x\theta_2\theta_1\xi_3 z = kk^{-1}k'k'^{-1}t^{-1}ty\xi_2 x\theta_2\theta_1\xi_3 z$, ce qui est une identité. Si l'on ne se contente pas de la solution évidente $\theta_6 = \theta_3\theta_4^{-1}\theta_5^{-1}$, $\xi_4 = \xi_1$, $\xi_5 = \xi_2$, $\xi_6 = \theta_2\theta_1\xi_3$, il faut supposer que les intersections de T et de K avec le centre ne se réduisent pas à l'unité.

24. Remarques. (i) Si l'on soumet tous les groupoïdes d'un système quelconque (E, Φ) à une commune isotopie, le nouveau système ne satisfait plus, en général, aux mêmes identités demosiennes que le premier. (Par exemple les dérivés de deux groupoïdes conjoints par une même isotopie ne sont plus conjoints). A quelles identités obéit le nouveau système?

(ii) M. Schauffler ([41], Th. 2, p. 430) a montré que l'ensemble de tous les quasigroupes finis d'ordre n est associatif demosien seulement pour $n = 2$ et 3. Cette proposition résulte immédiatement du Théorème N°7.2. En effet, d'après 7.2, deux quelconques des quasigroupes du système sont toujours isotopes d'un même groupe. Or on sait ([35], N°25), qu'à partir de $n = 4$ l'ensemble des quasigroupes définis sur un ensemble d'ordre n contient des éléments qui ne coïncident par aucune isotopie.

Chapitre IV

COMPOSITION

25. Produit de deux groupoïdes. L'intérêt des système demosiens serait grandement accru si l'on parvenait à les organiser eux-mêmes en groupoïdes en introduisant quelque loi de composition entre les φ, regardés comme des éléments de Φ. On pourrait alors appliquer à de tels systèmes des isomorphismes, homomorphismes, isotopies, isoméries, parastrophies, y définir des relations d'ordre stables, considérer leurs sous-groupoïdes, leurs groupoïdes quotients, etc. et examiner dans chaque cas la nature nouvelle des identités demosiennes obéies par le système, en elles-mêmes et dans leurs connexions avec les identités initiales et les constructions affectuées dans le système primitif. Deux lois de composition entre groupoïdes ont déjà èté introduites par l'auteur: le *produit à droite*, ([36], N°12), le *produit suivant un groupoïde de base donné* $G(\circ)$, ([31], p. 230, N°4). Mais on peut en imaginer d'autres, en particulier sur les systèmes construits par isotopie. Le chapitre se termine par la définition du produit direct de deux systèmes demosiens.

26. Produit à droite. DÈFINITION 26.1. (i) *Etant donnès deux groupoïdes, définis sur le même ensemble E par les lois de composition* φ_1 *et* φ_2, *leur produit à droite* $\varphi_1 \nabla \varphi_2 = \varphi$ *est le groupoïde défini sur E par la relation,* $\forall x, y \in E$, $x\varphi y = (x\varphi_1 y)\varphi_2 y$, ([36], N°12), [27].

(ii) *Le semigroupe défini par* $xy = y$ *s'appelle le semigroupe à translation identique, ou de Thierrin.*

Ce semigroupe, qui satisfait à la loi de translation identique ([35], p. 153, N°2, équ. 9, [36], N°3, ex. IV), a été étudié par Thierrin ([44], p. 178), sous le nom d'*anti-semigroupe.*

THÉORÈME 26.2. (i) *L'ensemble de groupoïdes, S, engendré par un système* (E, Φ) *au moyen du produit à droite, est un semigroupe par rapport à ce produit. Le semigroupe* $xy = x$ *est unité, le semigroupe de Thierrin est idéal nul à droite dans tous les cas et zéro à gauche si tous les* $\varphi_i \in \Phi$ *sont idempotents.*

(ii) *Pour que S soit un groupe il faut et il suffit,*

(a) *qu'il contienne le semigroupe* $xy = x$,

(b) *que ses générateurs soient des quasigroupes à gauche,*

(c) *que* (E, Φ) *satisfasse à la loi demosienne des keys* (N°4). *Si E est fini la condition* (b) *entraîne les deux autres.*

Preuve. (i) Par définition $\forall x, y \in E$, $(x\varphi_1 y)\varphi_2 y = x\varphi_3 y \rightleftarrows \varphi_3 = \varphi_1 \nabla \varphi_2$, $(x\varphi_3 y)\varphi_4 y = x\varphi_5 y \rightleftarrows \varphi_5 = (\varphi_1 \nabla \varphi_2) \nabla \varphi_4$, $(x\varphi_2 y)\varphi_4 y = x\varphi_6 y \rightleftarrows \varphi_6 = \varphi_2 \nabla \varphi_4$, $(u\varphi_1 y)\varphi_6 y = u\varphi_7 y \rightleftarrows \varphi_7 = \varphi_1 \nabla \varphi_6 = \varphi_1 \nabla (\varphi_2 \nabla \varphi_4)$. On tire de la 3^{me} équation, en mettant $u\varphi_1 y$ à la place de x, $[(u\varphi_1 y)\varphi_2 y]\varphi_4 y = (u\varphi_1 y)\varphi_6 y = u\varphi_7 y$; donc $(\varphi_1 \nabla \varphi_2) \nabla \varphi_4 = \varphi_7 = \varphi_1 \nabla (\varphi_2 \nabla \varphi_4)$.

Le semigroupe de translation identique, $x\varphi y = x$, est visiblement unité de S car $(x\varphi_1 y)\varphi y = x\varphi_1 y = (x\varphi y)\varphi_1 y$, donc $\varphi_1 \nabla \varphi = \varphi \nabla \varphi_1 = \varphi_1$. Le semigroupe de Thierrin, $x\varphi_0 y = y$, est idéal nul à droite car $(x\varphi_i y)\varphi_0 y = y = x\varphi_0 y$, donc $\varphi_i \nabla \varphi_0 = \varphi_0$. Si tous les $\varphi_i \in \Phi$ sont des groupoïdes idempotents, $x\varphi_i x = x$, il en est de même de tous les $\varphi_i \in \{\Phi\}$ et alors $(x\varphi_0 y)\varphi_i y = y \rightleftarrows y\varphi_i y = y = x\varphi_0 y$, ou $\varphi_0 \nabla \varphi_i = \varphi_0$. Le zéro φ_0 et l'unité φ sont conjoints ([35], p. 155, N°2, ii).

(ii) Pour que S soit un groupe il faut et il suffit qu'il contienne l'unité, φ, et, avec tout groupoïde φ_i, son inverse φ_j, $(\varphi_i \nabla \varphi_j = \varphi)$. Si

S est un groupe et φ_i la loi de composition d'un de ses groupoïdes, l'inverse de φ_i sera φ_j, défini par

$$\text{(26.1)} \qquad \varphi_i \nabla \varphi_j = \varphi \ , \quad \text{ou} \quad (x\varphi_i y)\varphi_j y = x\varphi y = x \ .$$

Si $x\varphi_i a = b$, on aura donc $b\varphi_j a = x$; cela signifie que l'équation $x\varphi_i a = b$, $\forall \varphi_i \in \{\Phi\}$, a une solution et une seule en x, $x = b\varphi_j a$. Tous les φ_i, et en particulier $\varphi_i \in \Phi$, sont donc des lois de quasigroupes à gauche. Enfin la relation (26.1) exprime que S satisfait à la loi demosienne des keys à droite (déf. 4.1). Ainsi, les conditions (a), (b), et (c) sont remplies. Réciproquement, supposons (a), (b), (c) vérifiées. Tous les générateurs, $\varphi_i \in \Phi$, de S sont des quasigroupes à gauche et il en sera évidemment de même de tous les φ_j engendrés par Φ au moyen du produit à droite; d'autre part Φ contient, en même temps que tout groupoïde, son inverse car, tout quasigroupe à gauche a un réciproque ([34], déf. 1.2), $x\varphi_i y = z \rightleftarrows z\varphi_j y = x$. Ensuite, en vertu de (c), on a $\forall \varphi_i \in \Phi$, $\exists \varphi_j \in \Phi$, $(x\varphi_i y)\varphi_j y = x$. Ainsi, φ_j, inverse de φ_i, est défini (b) et appartient à Φ, (c). Dès lors, soient φ_i et φ_k deux groupoïdes de Φ, φ_j et φ_t leurs inverses respectifs; alors,

$$(\varphi_i \nabla \varphi_k) \nabla (\varphi_t \nabla \varphi_j) = \varphi_i \nabla (\varphi_k \nabla \varphi_t) \nabla \varphi_j = \varphi_i \nabla \varphi \nabla \varphi_j = \varphi_i \nabla \varphi_j = \varphi \ .$$

Donc les produits $\varphi_i \nabla \varphi_k$ et $\varphi_t \nabla \varphi_j$ sont encore inverses. A toute étape de la construction de $\{\Phi\}$, la partie déjà engendrée contiendra toujours, avec tout groupoïde, son inverse et, d'après la remarque précédente, il sera possible de maintenir cette situation jusqu' au bout en faisant, en même temps que le produit de deux groupoïdes de cette partie déjà construite, celui de leurs inverses. De sorte que S, contenant lui aussi, avec tout groupoïde, son inverse, est un groupe.

Si l'on postule seulement la condition (b), soit φ_i un élément quelconque de $\{\Phi\}$. L'application $\Delta_a = (x \to x\varphi_i a)$ est une permutation de l'ensemble E sur lequel est construit chaque quasigroupe. Les translations Δ_a calculées pour les puissances successives de φ_i : φ_i, $\varphi_i^2 = \varphi_i \nabla \varphi_i$, $\varphi_i^3 = \varphi_i^2 \nabla \varphi_i$, $\cdots$ seront $\Delta_a = (x \to x\varphi_i a)$, $[x \to (x\varphi_i a)\varphi_i a] = (\Delta_a)^2$, $[x \to (x\varphi_i^2 a)\varphi_i a] = (\Delta_a)^3$, $\cdots$ Ce sont les puissances successives de la première. Si E est fini, il existera un entier positif n, tel que $(\Delta_a)^n$ soit identique quel que soit a. Donc $x\varphi_i^n y = x = x\varphi y$ et $\{\Phi\}$ contiendra le semigroupe unité. Enfin l'égalité $\varphi_i^{n-1} \nabla \varphi_i = \varphi$ montre que tout φ_i aura un inverse; donc S sera un groupe. Le raisonnement n'est plus valable si E n'est pas fini. Alors S peut ne pas contenir l'unité. Par exemple si S est engendré par le groupe cyclique $x\varphi_1 y = x + y$ sur l'anneau Z des entiers rationnels, $x\varphi_i y = x + yi$; il n'y a aucune valeur de l'entier positif i pour laquelle $x\varphi_i y$ se réduise à x. Si l'on adjoint à l'ensemble Φ le semigroupe unité $x\varphi y = x$, néanmoins, cet ensemble ne contiendra pas les réciproques de φ_i, $(x\varphi_k y = x - yi)$.

REMARQUE 26.3. Dans tout ce chapitre, certains vocables sont inséparables du contexte. Ainsi quand on dira qu'un groupoïde φ_i est idempotent, cela peut signifier que tous ses éléments sont idempotents ($x\varphi_i x = x$), ou que l'élément φ_i du semigroupe Φ est idempotent par rapport au produit à droite (∇), c'est-à-dire $\varphi_i \nabla \varphi_i = \varphi_i$, ou $(x\varphi_i y)\varphi_i y = x\varphi_i y$. De même il est très différent de dire que le groupoïde $E(\varphi_i)$ satisfait à la loi des keys, $(x\varphi_i y)\varphi_i y = x$, ou que le système (E, Φ) possède la loi des keys demosienne, $(x\varphi_1 y)\varphi_2 y = x$. Dans le premier cas on pourrait parler de la loi des keys en soi, dans le second de la loi demosienne.

27. Système associatif. THÉORÈME 27.1. *Si (E, Φ) est un système associatif demosien, dont chaque élément est un groupoïde idempotent en soi, alors (E, Φ) est fermé par rapport au produit à droite* (∇). *Conclusion analogue si le système a l'inversibilité demosienne.*

Preuve. Par définition du produit (∇) de deux groupoïdes, $(x\varphi_1 y)\varphi_2 y = x\varphi_3 y \rightleftarrows \varphi_1 \nabla \varphi_2 = \varphi_3$. Comme (E, Φ) a l'associativité demosienne, $\forall \varphi_1, \varphi_2 \in \Phi$, $\exists \varphi_4, \varphi_5 \in \Phi$, $(x\varphi_1 y)\varphi_2 y = x\varphi_4(y\varphi_5 y)$. Mais puisque tous les groupoïdes de (E, Φ) sont idempotents, $\forall \varphi_5$, $y\varphi_5 y = y$, d'où $(x\varphi_1 y)\varphi_2 y = x\varphi_4 y$, et $\varphi_3 = \varphi_4$. Donc $\varphi_1 \nabla \varphi_2 \in \Phi$. De plus $\varphi_1 \nabla \varphi_2$ est évidemment idempotent en soi. Ainsi Φ est fermé par rapport au produit (∇).

L'exemple suivant est susceptible d'applications pratiques.

EXEMPLE 27.2. Sur un corps quelconque, K, le système des quasigroupes $x\varphi_a y = ax + (1 - a)y$, $(a, x, y \in K)$, ([8], p. 112) possède l'associativité demosienne et il est fermé par rapport à (∇). Il est isomorphe au groupe multiplicatif du corps.

28. Systèmes unipotents, idempotent, nilpotents. THÉORÈME 28.1. *Etant donné un système (E, Φ) engendré par le produit à droite* (∇), (i) *pourqu'un groupoïde $\varphi_i \in \Phi$ soit unipotent par rapport à ce produit* ($\varphi_i \nabla \varphi_i = \varphi$, $x\varphi y = x$), *il faut et il suffit qu'il satisfasse à la loi des keys à droite* (en soi). *Toutes les translations à droite Δ_y sont alors, dans φ_i, des involutions.* (ii) *Pour que φ_i soit idempotent par rapport à* (∇),($\varphi_i \nabla \varphi_i = \varphi_i$), *il faut et il suffit que les éléments de E qui ont une unité à gauche dans $E(\varphi_i)$,* ($u\varphi_i x = x, \forall x$) *soient idempotents et que, dans chaque translation $\Delta_y = (x \rightarrow x\varphi_i y)$, tout élément x soit sa propre projection, ou ait comme image un élément se projetant sur lui-même.* (iii) *Pour que φ_i soit nilpotent d'index n, il faut et il suffit qu'il soit idempotent en soi* ($x\varphi_i x = x$) *et que chaque translation Δ_y définisse sur E une relation d'ordre partiel* ($a \succ b \rightleftarrows a\varphi_i b = b$) *ayant pour diagramme de Hasse* ([21], II, N°17, p. 102, [7], p. 6) *un arbre issu de b y et dont les chaînes maximales aient pour longueur n.*

Preuve. (i) Soit $x\varphi_i y = a$ et $a\varphi_i y = b$, donc $(x\varphi_i y)\varphi_i y = b$. L'unité, $x\varphi y = x$ est le semigroupe conjoint de celui de Thierrin (N°26.1). Pour que $\varphi_i \nabla \varphi_i = \varphi$ il faut et il suffit que $b = x$, ou $(x\varphi_i y)\varphi_i y = x$. C'est la loi des keys à droite (en soi). La translation Δ_y est alors $\Delta_y = (x \to a)$ et comme $a\varphi_i y = x$, on a aussi $\Delta_y = (a \to x)$; donc Δ est du second ordre.

(ii) Pour que φ_i soit idempotent il faut et il suffit que $\varphi_i \nabla \varphi_i = \varphi_i$ ou $(x\varphi_i y)\varphi_i y = x\varphi_i y$. Si l'élément a possède une unité gauche x, dans φ_i, $x\varphi_i a = a$, la condition ci-dessus devient, en faisant $y = a$, $a\varphi_i a = a$. L'élément a est donc idempotent dans $E(\varphi_i)$. En particulier, les *zéroïdes* ([10], p. 118, [45], p. 87) à gauche de $E(\varphi_i)$ seront idempotents. D'autre part $x\varphi_i y = a \Rightarrow a\varphi_i y = a$; donc la translation $\Delta_y = (x \to x\varphi_i y)$ projette chaque élément, soit sur lui-même, soit sur un élément qui est sa propre image par Δ. La condition est visiblement suffisante.

(iii) Supposons φ_i nilpotent d'index $n \neq 1$; soit $x\varphi_i y = a$. Il y a au moins un $x \in E$ pour lequel $a \neq y$, sans quoi on aurait $x\varphi_i y = y$, $\forall x$, $\varphi_i = \varphi_0$ et φ_i serait nul. Il y a de même au moins un b tel que $a\varphi_i y = b \neq y$, $b\varphi_i y = c \neq y, \cdots$, etc. enfin $k\varphi_i y = y$. Il y a donc une chaîne unique $x, a, b, \cdots, y$, définie par les puissances successives de la translation Δ_y, reliant un élément arbitraire de E à y, et dont la longueur est inférieure ou égale à n, le maximum étant atteint au moins une fois. Partant de l'élément a défini ci-dessus, et puisque φ_i est nilpotent d'index n, on devra avoir $[(a\varphi_i y)\varphi_i y \cdots]\varphi_i y = y$. Mais, d'après le choix de a, l'expression entre crochets est égale à $k\varphi_i y$, ou y, d'où $y\varphi_i y = y$, $\forall y$. Ainsi φ_i est idempotent en soi. La réciproque est évidente.

29. Système associatif fermé. THÉORÈME 29.1. *Si $S = (E, \Phi)$ est un système de groupoïdes idempotents ($x\varphi_i x = x$), avec élément neutre à gauche commun, u, et satisfaisant à l'associativité demosienne, alors, tout sous-système $S' = (E, \Phi')$, $\Phi' \subseteq \Phi$, fermé par rapport au produit à droite (∇), possède aussi l'associativité demosienne.*

Preuve. Par hypothèse $\forall \varphi_1, \varphi_2 \in \Phi'$, $\exists \varphi_3, \varphi_4 \in \Phi$, $(x\varphi_1 y)\varphi_2 z = x\varphi_3(y\varphi_4 z)$. Si $z = y$, on a $(x\varphi_1 y)\varphi_2 y = x\varphi_3(y\varphi_4 y)$. Le premier membre est le produit *à* droite $\varphi_1 \nabla \varphi_2$ et, dans le second, $y\varphi_4 y = y$ en vertu de l'idempotence, donc $\forall x, y \in E$, $x(\varphi_1 \nabla \varphi_2)y = x\varphi_3 y$, ou $\varphi_3 = \varphi_1 \nabla \varphi_2$, et puisque Φ' est fermé par rapport à (∇), $\varphi_3 \in \Phi'$. Si $x = u$, l'hypothèse devient $(u\varphi_1 y)\varphi_2 z = y\varphi_2 z = u\varphi_3(y\varphi_4 z) = y\varphi_4 z$, donc $\varphi_4 = \varphi_2 \in \Phi'$. Puisque φ_3 et φ_4 sont dans Φ', le système (E, Φ') est associatif demosien. On aurait un théorème analogue avec la transitivité.

REMARQUE 29.2. La réciproque n'est pas vraie. Ainsi, le système proposé en exemple au N° 27.2 admet le sous-système engendré par les puissances de φ_a, $x\varphi_a^n y = a^n x + (1 - a^n)y$. Ce soussystème est associatif demosien; φ_a^n est idempotent, mais n'a pas d'unité à gauche.

QUESTION 29.3. *On munit un système (E, Φ), satisfaisant à l'associativité demosienne, d'une loi de composition (L) entre les groupoïdes de ce système. Le système (E, Φ) engendre, au moyen de cette loi, un système plus large (E, Φ'), $\Phi' \supseteq \Phi$. Quelles sont les lois (L) pour lesquelles (E, Φ') possède encore l'associativité demosienne?*

30. **Produit suivant un groupoide** $(\ast)$. DEFINITION 30.1. Si $E(\varphi_1)$ et $E(\varphi_2)$ sont deux groupoïdes quelconques sur un ensemble commun E, et $E(\ast)$ un groupoïde fixe (*fondamental*) donné sur E, on *appelle produit des groupoïdes $E(\varphi_1)$ par $E(\varphi_2)$, selon le groupoïde de base $E(\ast)$*, le groupoïde $E(\varphi_3)$ défini par la relation $x\varphi_3 y = (x\varphi_1 y) \ast (x\varphi_2 y)$, symboliquement $\varphi_3 = \varphi_1 \circ \varphi_2$.

Cette définition a été donnée à l'occasion des groupoïdes orthogonaux ([**31**], N°4, p. 230), mais elle est générale quels que soient les composants φ_1, φ_2. Un ensemble demosien est *fermé* par rapport au produit $(\circ)$ selon un groupoïde fondamental $E(\ast)$—appartenant ou non à Φ—si $\forall \varphi_1, \varphi_2 \in \Phi$, $\varphi_1 \circ \varphi_2 \in \Phi$. L'ensemble $(E, \{\Phi\})$ engendré par un systéme demosien (E, Φ) au moyen du produit $(\circ)$ est évidemment fermé par rapport à ce produit.

EXEMPLE 30.2. Sur l'anneau Z des entiers rationnels le système demosien dont les éléments sont les groupoïdes $x\varphi y = ax + by + c$, $Z \ni a, b, c =$ Constantes, est fermé par rapport à chacun de ces groupoïdes.

THÉORÈME 30.3. *Tout système (E, Φ) dont chaque groupoïde est idempotent en soi $(x\varphi x = x)$, et qui possède l'associativité et la commutativité demosiennes, est fermé par rapport à chacun de ses groupoïdes.*

Preuve. Par hypothèse, le groupoïde fondamental, φ_i, appartient au système et si $\varphi = \varphi_1 \circ \varphi_2$ on aura

$$\forall \varphi_1, \varphi_2, \varphi_i = \Phi \;, \qquad \exists \varphi_3, \varphi_4, \cdots, \varphi_{10} \in \Phi \;.$$
$$\begin{aligned} x\varphi x &= (x\varphi_1 y)\varphi_i(x\varphi_2 y) = x\varphi_3[y\varphi_4(x\varphi_2 y)] = x\varphi_3[y\varphi_4(y\varphi_5 x)] \\ &= x\varphi_3[(y\varphi_6 y)\varphi_7 x] = x\varphi_3(y\varphi_7 x) = x\varphi_3(x\varphi_8 y) \\ &= (x\varphi_9 x)\varphi_{10} y = x\varphi_{10} y \;. \end{aligned}$$

Donc $$\varphi = \varphi_{10} \in \Phi.$$

La réciproque n'est pas vraie, comme le montre l'exemple ci dessus, qui est fermé sans que ses groupoïdes soient idempotents en eux-mêmes.

THÉORÈME 30.4. *Si $G(\ast)$ est un groupoïde quelconque, appartenant ou non à (E, Φ), et fixé sur l'ensemble E, le système $(E, \{\Phi\})$, engendré* *par le système* (E, Φ) *sous la loi de composition* $\varphi_1 \circ \varphi_2 = \varphi_3 \rightleftarrows \forall x, y, z \in E$, $(x\varphi_1 y) \ast (x\varphi_2 y) = x\varphi_3 y$, *est un groupoïde* $\Sigma(\circ)$ *par rapport à l'opération* $(\circ)$, *ayant pour éléments les* φ_i, *et* Σ *est* (i) *un quasigroupe,* (ii) *un semigroupe,* (iii) *un groupoïde abélien, en même temps que* $G(\ast)$.

Preuve. (i) Si $G(\ast)$ est un quasigroupe, à tous $x\varphi_2 y$ et $x\varphi_3 y$ donnés correspond un $z \in E$, et un seul, tel que $z \ast (x\varphi_2 y) = x\varphi_3 y$. Maintenant, z étant défini univoque, $\forall x, y \in E$, la fonction $z = x\varphi_1 y$ est déterminée. Ainsi, l'équation $\varphi \circ \varphi_2 = \varphi_3$ a une solution unique φ_1 en φ, $\forall \varphi_2, \varphi_3$ et la loi de composition $(\circ)$ satisfait à l'axiome du quotient à gauche. Le même argument est valable à droite et $\Sigma(\circ)$ est un quasigroupe.

(ii) Si $G(\ast)$ est associatif, $\forall x, y, z \in E$, $\forall \varphi_1, \varphi_2, \varphi_3 \in \{\Phi\}$, $[(x\varphi_1 y) \ast (x\varphi_2 y)] \ast (x\varphi_3 y) = (x\varphi_1 y) \ast [(x\varphi_2 y) \ast (x\varphi_3 y)]$, ou $(\varphi_1 \circ \varphi_2) \circ \varphi_3 = \varphi_1 \circ (\varphi_2 \circ \varphi_3)$. Donc $\Sigma(\circ)$ est un semigroupe.

(iii) Si $G(\ast)$ est commutatif, $\forall x, y \in E$, $\forall \varphi_1, \varphi_2 \in \{\Phi\}$, $(x\varphi_1 y) \ast (x\varphi_2 y) = (x\varphi_2 y) \ast (x\varphi_1 y)$, ou $\varphi_1 \circ \varphi_2 = \varphi_2 \circ \varphi_1$, et $\Sigma(\circ)$ est abélien. Les réciproques sont évidentes (Cf [31], p. 233, N°8).

THÉORÈME 30.4.1. *Si chaque groupoïde d'un système associatif et inversible demosien* (E, Φ) *est idempotent en soi,* $(x\varphi_i x = x)$, *alors le système* $(E, \{\Phi\})$, *engendré par* (E, Φ) *sous la loi de composition* $\varphi_1 \circ \varphi_2 = \varphi_3 \rightleftarrows (x\varphi_1 y) \ast (x\varphi_2 y) = x\varphi_3 y$, $\Phi \ni \ast$, *fixé, possède encore l'associativité l'inversibilité demosiennes, l'idempotence en soi.*

Preuve. (i) Démontrons l'idempotence par induction. Supposons qu'à une étape donnée de la construction de $(E, \{\Phi\})$ tous les groupoïdes déjà engendrés soient idempotents; alors, $\forall x \in E$, $x\varphi_i x = x$, $x\varphi_j x = x$, où $E(\varphi_i)$ et $E(\varphi_j)$ sont des groupoïdes déjà construits. Le produit $\varphi_i \circ \varphi_j = \varphi_k$ est défini par $x\varphi_k y = (x\varphi_i y) \ast (x\varphi_j y)$. Si $x = y$, $x\varphi_k x = (x\varphi_i x) \ast (x\varphi_j x) = x \ast x = x$; donc φ_k est idempotent.

(ii) Il résulte du N°14.3, iv que (E, Φ) a la commutativité demosienne. Le même argument que ci-dessus montre que cette commutativité se transfère à $(E, \{\Phi\})$. En effet, avec les mêmes notations que plus haut, soit $x\varphi_i y = y\varphi_1 x$, $x\varphi_j y = y\varphi_2 x$, $x\varphi_k y = (x\varphi_i y) \ast (x\varphi_j y) = (y\varphi_1 x) \ast (y\varphi_2 x) = y\varphi_3 x$, avec $\varphi_3 = \varphi_1 \circ \varphi_2$.

(iii) La même induction s'applique à l'inversibilité demosienne. On a successivement

$$\begin{aligned} x\varphi_k y &= (x\varphi_i y) \ast (x\varphi_j y) = (y\varphi_3 x) \ast (x\varphi_j y) = y\varphi_4[x\varphi_5(x\varphi_j y)] \\ &= y\varphi_4[(x\varphi_6 x)\varphi_7 y] = y\varphi_4(x\varphi_7 y) = y\varphi_4(y\varphi_8 x) \\ &= (y\varphi_9 y)\varphi_{10} x = y\varphi_{10} x \,. \end{aligned}$$

Donc $(x\varphi_k y)\varphi_{11} z = (y\varphi_{10} x)\varphi_{11} z = (x\varphi_{12} y)\varphi_{11} z = x\varphi_{13}(y\varphi_{14} z)$. La démonstration serait analogue à partir de $(x\varphi_{11}y)\varphi_k z$; par suite, φ_k satisfait à l'associativité demosienne et, en vertu du N° 14.1, le système est aussi inversible demosien.

THÉORÈME 30.5. *Sur un système (E, Φ) associatif inversible demosien, fermé par rapport au semigroupe (avec unité) $\ast \in \Phi$, l'application $(\varphi_i \to \varphi_2 \circ \varphi_i \circ \varphi_1) \rightleftarrows [x\varphi_i y \to (x\varphi_2 y) \ast (x\varphi_i y) \ast (x\varphi_1 y)]$, φ_1, φ_2 fixés, où $(x\varphi_1 y) \ast (x\varphi_2 y) = u$, unité de $E(\ast)$, $\forall x, y \in E$, est un endomorphisme.*

Preuve. Soit $\varphi_i \circ \varphi_j = \varphi_k$, ou $(x\varphi_i y) \ast (x\varphi_j y) = x\varphi_k y$.

$$\begin{aligned}(x\varphi_2 y) \ast (x\varphi_i y) \ast (x\varphi_1 y) \ast (x\varphi_2 y) \ast (x\varphi_j y) \ast (x\varphi_1 y) \\ = (x\varphi_2 y) \ast (x\varphi_i y) \ast u \ast (x\varphi_j y) \ast (x\varphi_1 y) \\ = (x\varphi_2 y) \ast (x\varphi_i y) \ast (x\varphi_j y) \ast (x\varphi_1 y) \\ = (x\varphi_2 y) \ast (x\varphi_k y) \ast (x\varphi_1 y)\ .\end{aligned}$$

Donc $(\varphi_2 \circ \varphi_i \circ \varphi_1) \circ (\varphi_2 \circ \varphi_j \circ \varphi_1) = \varphi_2 \circ \varphi_k \circ \varphi_1$. On peut aussi concevoir des endomorphismes s'exerçant par application de E dans lui-même $(x \to x')$, dans chaque groupoïde. Alors, si I_i est le semigroupe d'endomorphisme de $E(\varphi_i)$, celui du système sera évidemment $\cap I_i$.

REMARQUES 30.6. Soient α, β, m, p des constantes quelconques choisies parmi les éléments du E. On peut définir des classes—analogues aux classes d'isonomie ([31], p. 236, N°16)—par la condition que deux groupoïdes $E(\varphi_i)$ et $E(\varphi_j)$ du système (E, Φ) soient dans la même classe K_m si $\alpha\varphi_i\beta = \alpha\varphi_j\beta = m$. Ici ces classes sont nécessairement disjointes et, à chaque classe, correspond univoquement un élément $m \in E$. Une telle partition de (E, Φ) est régulière par rapport au produit selon un groupoïde $G(\ast)$, car soit $\varphi_3 = \varphi_1 \circ \varphi_2 \rightleftarrows \forall x, y \in E$, $(x\varphi_1 y) \ast (x\varphi_2 y) = x\varphi_3 y$. Si $\varphi_1 \in K_m$ et $\varphi_2 \in K_p$, on aura $\alpha\varphi_1\beta = m$, $\alpha\varphi_2\beta = p$, et $\alpha\varphi_3\beta = (\alpha\varphi_1\beta) \ast (\alpha\varphi_2\beta) = m \ast p = r$. L'élément r ne dépend que de m et de p, donc $K_m \circ K_p = K_{m \ast p}$. L'application $(\varphi_i \to m) \rightleftarrows \varphi_i \in K_m$ est un homomorphisme de (E, Φ) sur $G(\ast)$. A chaque choix de (α, β), c'est-à-dire à chaque élément de EE correspond un tel homomorphisme. L'ensemble de tous ces homomorphismes, quand (α, β) décrit EE, muni de la loi de composition $(\alpha, \beta) \circ (\alpha', \beta') = (\alpha \ast \alpha', \beta \ast \beta')$, est isomorphe au carré direct de $G(\ast)$.

On peut imaginer d'autres moyens d'organiser un système demosien en groupoïde. Bornons-nous, pour terminer, à indiquer la construction du produit de deux systèmes demosiens.

31. Produit direct. DÉFINITION 31.1. Soient (E, Φ) et (E', Φ') deux systèmes demosiens, EE' et $\Phi\Phi'$ les ensembles produits. Le produit direct de ces deux systèmes est le système $(EE', \Phi\Phi')$ défini $\forall x, y \in E$,

$x', y' \in E'$, $\varphi_i, \varphi_j \in \Phi$, $\varphi'_i, \varphi'_j \in \varphi'$, et dont les groupoïdes ont pour loi de composition, sur l'ensemble EE', les $[\varphi_i, \varphi'_j] \in \Phi\Phi'$, tels que $(x, x')[\varphi_i, \varphi'_j](y, y') = (x\varphi_i y, x'\varphi'_j y')$. Comme cas particulier, les ensembles peuvent coïncider: $E = E'$, ou $\Phi = \Phi'$. Dans le premier cas $(x, y)[\varphi_i, \varphi'_j](z, t) = (x\varphi_i z, y\varphi'_j t)$. Dans le second cas, on peut prendre $(x, x')\varphi(y, y') = [(x\varphi y), (x'\varphi y')]$, sans modifier φ, ou bien $(x, x')[\varphi_i, \varphi_j](y, y') = (x\varphi_i y, x'\varphi_j y')$ en passant de Φ à $\Phi\Phi$.

THÉORÈME 31.2. *Si deux systèmes satisfont à une même loi demosienne, leur produit direct satisfait aussi à cette loi.*

Bornons nous à exposer les calculs dans le cas de l'associativitè demosienne.

$$\begin{aligned}\{(x, x')[\varphi_i, \varphi'_i](y, y')\}[\varphi_j, \varphi'_j](z, z') &= (x\varphi_i y, x'\varphi'_i y')[\varphi_j, \varphi'_j](z, z')\\ &= [(x\varphi_i y)\varphi_j z, (x'\varphi'_i y')\varphi'_j z'] = [x\varphi_k(y\varphi_m z), x'\varphi'_k(y'\varphi'_m z')]\\ &= (x, x')[\varphi_k, \varphi'_k]((y\varphi_m z), (y'\varphi'_m z'))\\ &= (x, x')[\varphi_k, \varphi'_k]\{(y, y')[\varphi_m, \varphi'_m](z, z')\} .\end{aligned}$$

EXEMPLE 31.3. Le système $E = (0, 1)$, $\Phi = (\times, \ast)$, défini par $0 \times 0 = 1 \times 1 = 0$, $0 \times 1 = 1 \times 0 = 1$, $0 \ast 0 = 1 \ast 1 = 1$, $0 \ast 1 = 1 \ast 0 = 0$, est associatif demosien. Si l'on fait son carré direct, on obtient un système associatif demosien de quatre groupes isomorphes au groupe carré de Klein.

Terminologie

Anticentre	22.1
Associateur d'un système (E, Φ)	11
Associativité demosienne	7
Complexe relatif aux translations à droite	3.6
Conjoint	3
Demi-symétrie	5
Demosien	2
Distributivité demosienne	8
Egales (Expressions)	2
Ensemble demosien fermé	30
Equation	2
Expression sur (E, Φ)	2
Identité	2
Inversibilité demosienne	6
Keys	4
Multigroupoïde	12
Multistructuré	2
Parastrophie	9

Produit à droite de deux groupoïdes	26.1
Produit direct de deux (E, Φ)	31
Produit suivant un groupoïde fondamental	25; 30.1
Semigroupe à translation identique	26
Semigroupe de Thierrin	26.1
Sous-système demosien	3.3
Système (G, K, T)	19.2
Transitivité	9.3
Zéroïde	28

NOTES

1. *Cf. Un theorèm plus général—Entropie demosienne de multigroupoïdes et de quasigroupes*—Ann. Soc. Sci. Bruxelles. **73** (1959), 302-309.
2. *Il faut entendre ici l'isomorphisme de deux complexes comme appliquant les sous-ensembles* $\Sigma_j \Delta_i^{-1} \Delta_j$ (i=constante) *les uns sur les autres.*
8. *Ne pas confondre cette notion, due à Schanffler, avec le concept usuel d'inverse.*
4. *Au sens de Schauffler.*

RÉFÉRENCES

1. Janos Aczél, *Some general methods in the theory of functional equations in one variable; new applications of functional equations,* Uspehi Mat. Nauk. (N.S.) **11** (1956), N°3, (69), 3–68.

2. ———, *Remarques algébriques sur la solution donnée par M. Fréchet à l'équation de Kolmogoroff,* Publ. Math. Debrecen, **4** (1955), 33–42.

3. ———, *Ueber Additions-und Substraktionstheoreme,* Publ. Math. Debrecen, **4** (1956), 325–333.

4. J. Aczél et Miklós Hosszú, *On transformations with several parameters and operations in multidimensional spaces,* Acta Math. Acad. Sci. Hungaricae, **7** (1956), 327–338.

5. Adrian A. Albert, *Quasigroups,* Trans. Amer. Math. Soc., **54** (1943), 507–519.

6. Belousov, *Associativnye Sistemy* Kvazigrupp, Uspehi Math. Nauk, **13** (1958), fas. 3, (81). 243.

7. Garrett Birkhoff, *Lattice theory,* Amer. Math. Soc. Colloqu. Publ., New-York, **25** (1948).

8. C. Burstin et W. Mayer, *Distributive Gruppen,* Journ. reine angew. Math., **160** (1929), 111–130.

9. R. Cacciopoli, *L'equazione funzionale* $f(x+y) = F[f(x), f(y)]$, Giorn. Mat. Battaglini (3), **66** (1928), 69–74.

10. A. H. Clifford et D. D. Miller, *Semi-groups having zeroïd elements,* Amer. J. Math., **70** (1948), 117–125.

11. P. M. Cohn, *Embeddings in semigroups with one-sided division,* J. London Math. Soc., **31** (1956), 169–181.

12. Karel Drbohlav, *Gruppenartige Multigruppen,* Чехословацкий Математический Журнал, Т. 7 (82), (1957), Прага, 183–190.

13. Trevor Evans, *A note on the associative law,* Journ. London Math. Soc., **25** (1950), 196–201.

14. Ladislaus Fuchs, *On mean systems,* Acta. Math. Ac. Sci. Hungaricae, **1** (1950), 303–320.

15. Michel Ghermănescu, *Sur un système d'équations fonctionnelles,* Acad. Republ. Pop. Romíne. Bul. Sti. Sect. Sti. Mat. Fiz., **5** (1953), 575–582.

16. Michel Ghermănescu, *Sur quelques équations fonctionnelles à deux variables*, Ibid., **7** (1955), 963–975.
17. S. Golab, *Zum distributiven Gesetz der reellen Zahlen*, Studia Math., **15** (1956), 353–358.
18. Hermann Grassmann, *Gesammelte Werke*, Leipzig (1894), I, 1.
19. Béla Gyires, *Ueber die Grenzwerte von Matrizen, die den Cauchyschen Funktionalgleichung genügen*, Acta Univers. Debrecen, **1** (1954), 136–144.
20. Philip Hall, *On a theorem of Frobenius*, Proc. London Math. Soc., (2) **40** (1936), 468–501.
21. Helmut Hasse, *Höhere Algebra*, II, Berlin (1951).
21bis. ———, *Jnvariante Kennzeichung Galoischer Körper mit vorgegebener Galoisgruppe*, J. reine angew. Math, **187** (1950), 14–43.
22. Miklós Hosszú, *On the functional equation of auto-distributivity*, Publ. Math. Debrecen, **3** (1953), 83–86.
23. ———, *On the functional equation of distributivity*, Acta Math. Acad. Sci. Hungaricae, **4** (1953), 159–167.
24. ———, *On the functional equation of transitivity*, Acta Scientiarum Math., **15** (1954), 203–208.
25. ———, *Some functional equations related with the associative law*, Publ. Math. Debrecen, **3** (1954), 205–214.
26. ———, Lettre à l'auteur du 10 Juillet 1958, et *Belouszov* egy *tételéröl és annak néhány alkalmazásáról*. M. Tud. Akad. Mat. Fiz. Osztály Közleményei, **9** (1959), 51–56.
27. Andor Kertész et Albert Sade, *On some mappings of groupoids into themselves; nuclei and their $\langle\lambda, \mu\rangle$ isotopes*, Publ. Math. Debrecen, **6** (1959), 000–000.
28. Bronislaw Knaster, *Sur une équivalence pour les fonctions*, Colloquium Math., **2** (1949), 1–4.
29. Svetozar Kurepa, *On some functional equations*, Glasnik Mat. Fiz. Astr. Drustwo Mat. Fiz. Hrvatske, Ser. II, **11** 1956), 3–5.
30. Albert Sade, *Groupoïdes automorphes par le groupe cyclique*, Canadian J. Math., **9** (1957), 321–335.
31. ———, *Groupoïdes orthogonaux*, Publ. Math. Debrecen, **5** (1958), 229–240.
32. ———, *Groupoïdes automorphes par le groupe géométrique et quasigroupes "endo"*, Canadian J. Math., **10** (1958), 294–320.
33. ———, *Quelques remarques sur l'isomorphisme et l'automorphisme des quasigroupes*, Abhand. Math. Semin. Univ. Hamburg, **22** (1958), 84–91.
34. ——— , *Quasigroupes parastrophiques. Expressions et identités*, Math. Nachrichten, **20**, (1959), 73-106.
35. ———, *Quasigroupes obéissant à certains lois*, Istanbul üniversitesi Fen Fakültesi Mecmuasi, Ser. A, **22** (1959), 151–184.
36. ———, *Groupoïdes en relation associative et semigroupes mutuellement associatifs*, Publ. Math. Debrecen, **7** (1961).
37. ———, *Système demosien associatif de multigroupoïdes avec un scalaire non singulier*. Ann. Soc. Sci. Bruxelles. **73** (1959), 231-234.
38. ———, *Solution générale de l'équation de transitivité sur un ensemble quelconque avec division bilatère*,Bull. Acad. Roy. Belgique. Classe des Sci., 5e S. **45** (1959), 451–455.
39. ———, *Multigroupoïde sur un système demosien de groupoïdes*, Colloquium Math. Vol. VIII, fasc. 1 (1960).
40. Rudolf Schauffler, *Ueber die Bildung von Codewörtern*, Arch. elektr. Uebertragung, **10** (1956), 303–314.
41. ———, *Die Assoziativität im Ganzen, besonders bei Quasigruppen*, Math. Zeitsch., **67** (1957), 428–435.
42. H. F. Scherk, *Analytisch-combinatorisch Zätze*, Journ. reine angew. Math., **11** (1834), 226–240.

43. Th. Skolem, *The abundance of arithmeticfunctions satisfying some simple functional equations*, Kongelige Norske Vid. Sels. forhandlinger, **29** (1956), 47–53.

44. Gabriel Thierrin, *Sur une classe de demi-groupes inversifs*, C. R. Acad. Sci. Paris, **234** (1952), 177–179.

45. ———, *Demi-groupes inversés et rectangulaires*, Bulletin Acad. Roy. Belgique. Classe des Sci. 5e S., **41** (1955), 83–92.

ON NORMAL NUMBERS

WOLFGANG SCHMIDT

1. Introduction. A real number ξ, $0 \leqq \xi < 1$, is said to the *normal in the scale of* r (or *to base* r), if in $\xi = 0 \cdot a_1 a_2 \cdots$ expanded in the scale of r[1] every combination of digits occurs with the proper frequency. If $b_1 b_2 \cdots b_k$ is any combination of digits, and Z_N the number of indices i in $1 \leqq i \leqq N$ having

$$b_1 = a_i, \cdots, b_k = a_{i+k-1} ,$$

then the condition is that

$$\lim_{N \to \infty} Z_N N^{-1} = r^{-k} . \tag{1}$$

A number ξ is called *simply normal* in the scale of r if (1) holds for $k = 1$. A number is said to be *absolutely normal* if it is normal to every base r. It is well-known (see, for example, [6], Theorem 8.11) that almost every number ξ is absolutely normal.

We write $r \sim s$, if there exist integers n, m with $r^n = s^m$. Otherwise, we put $r \not\sim s$.

In this paper we solve the following problem. *Under what conditions on* r, s *is every number* ξ *which is normal to base* r *also normal to base* s? The answer is given by

THEOREM 1. A *Assume* $r \sim s$. *Then any number normal to base* r *is normal to base* s.

B If $r \not\sim s$, *then the set of numbers* ξ *which are normal to base* r *but not even simply normal to base* s *has the power of the continuum.*

The A-part of the Theorem is rather trivial, but I shall sketch a proof of it, since I could not find one in the literature.

Next, let I be an interval of length $|I|$ contained in the unit-interval $U = [0, 1]$. We write $M_N(\xi, r, I)$ for the number of indices i in $1 \leqq i \leqq N$ such that the fractional part $\{r^i \xi\}$ of $r^i \xi$ lies I. A sequence $\xi, r\xi, r^2\xi, \cdots$ *has uniform distribution modulo* 1 if

$$R_N(\xi, r, I) = M_N(\xi, r, I) - N|I| = o(N)$$

for any I. It was proved by Wall [8] (the most accessible proof in [6], Theorem 8.15) that ξ is normal to base r if and only if $\xi, r\xi, r^2\xi, \cdots$ has uniform distribution modulo 1.

Write $T_{s,t}$, where $1 < t < s$, for the following mapping in U: If $\xi = 0 \cdot a_1 a_2 \cdots$ in the scale of t, then $T_{s,t}\xi = 0 \cdot a_1 a_2 \cdots$ in the scale of s.

Received June 2, 1959.

[1] In case of ambiguity we take the representation with an infinity of a_i less then $r - 1$. But this does not affect the property of ξ to be normal or not.

THEOREM 2. *Assume $r \not\sim s$. Then there exists a constant $\alpha_1 = \alpha_1(r, s, t) > 0$ such that for almost every ξ there exists a $N_0(\xi)$ with*

$$R_N(T_{s,t}\xi, r, I) \leqq N^{1-\alpha_1} \tag{2}$$

for every $N \geqq N_0(\xi)$ and any I.

Thus $T_{s,t}\xi$ is normal to base r for almost all ξ. Since $T_{s,t}\xi$ is not simply normal to base s part B of Theorem 1 follows. It does not follow immediately for $s = 2$, but instead of $T_{2,t}$, which does not exist, we may take $T_{4,t}$.

We can interpret our results as follows. Write $C_{s,t}$ for the image set $T_{s,t}U$ of the unit-interval U under the mapping $T_{s,t}$. $C_{s,t}$ is essentially a Cantor set. In $C_{s,t}$ we define a measure $\mu_{s,t}$ by

$$\int_{C_{s,t}} f(\xi)d\mu_{s,t} = \int_0^1 f(T_{s,t}\xi)d\xi \ , \tag{3}$$

where $f(\xi)$ is any real-valued function such that the integral on the right hand side of (3) exists. Then it follows from Theorem 2 that *with respect to $\mu_{s,t}$ almost every ξ in $C_{s,t}$ is normal in the scale of r.*

Throughout this paper, lower case italics stand for integers. $\alpha_1 = \alpha_1(r, s, t), \alpha_2, \alpha_3, \cdots$ will be positive constants depending on some or all the variables r, s, t.

1. The case $r \sim s$. First, it follows almost from the definition that any number normal to base s^n is normal to base s.

Next, assume ξ is normal to base r, we shall show it is normal in the scale of r^m. If $\xi = 0 \cdot a_1 a_2 \cdots$ in the scale of r, $b_1 \cdots b_{mk}$ is any combination of mk digits and $Z_N^{(1)}$ is the number of indices i in $1 \leqq i \leqq N$ with $i \equiv 1 \pmod m$ satisfying

$$b_1 = a_i, \cdots, b_{mk} = a_{i+mk-1} \ ,$$

then it was shown in [7] and in [3] that

$$\lim_{N\to\infty} Z_N^{(1)} N^{-1} = r^{-mk} m^{-1}$$

and hence

$$\lim_{N\to\infty} Z_{mN}^{(1)} N^{-1} = (r^m)^{-k} \ .$$

Thus ξ is normal to base r^m.

Combining the above remarks we obtain the A-part of Theorem 1.

2. The measure $\mu_{s,t}$. We define *numbers of order h* to be the number $0 \cdot a_1 \cdots a_h$ with $0 \leqq a_i < t$ in the scale of s. There are t^h numbers of order h, we denote them in ascending order by $\theta_1^{(h)}, \cdots, \theta_{t^h}^{(h)}$.

LEMMA 1. *Let $f(\xi)$ be a step-function, having a finite number of steps. Then*

$$\int_{C_{s,t}} f(\xi)d\mu_{s,t} = \int_0^1 f(T_{s,t}\xi)d\xi = \lim_{h\to\infty} t^{-h}\sum_{k=1}^{t^h} f(\theta_k^{(h)}) .$$

The integrals and the limit exist and are finite.

Proof. It will be sufficient to prove the lemma for $f(\xi) = \{\xi, \gamma\}$, where $0 \leqq \gamma \leqq 1$ and

$$\{\xi, \gamma\} = \begin{cases} 1, & \text{if } \{\xi\} < \gamma \\ 0 & \text{otherwise.} \end{cases}$$

$\xi_k^{(h)} = \int_0^1 \{T_{s,t}\xi, \theta_k^{(h)}\}d\xi$ is the least upper bound of numbers ξ having $T_{s,t}\xi \leqq \theta_k^{(h)}$. Thus if $\theta_k^{(h)} = 0\cdot a_1 \cdots a_h$ in the scale of s, then $\xi_k^{(h)} = 0\cdot a_1 \cdots a_h$ in the scale of t and therefore $\xi_k^{(h)} = (k-1)t^{-h}$.

Hence if $\theta_k^{(h)} \leqq \gamma \leqq \theta_{k+1}^{(h)}$, or if $\theta_k^{(h)} \leqq \gamma$ with $k = t^h$, then

$$\int_0^1 \{T_{s,t}\xi, \gamma\}d\xi = kt^{-h} - \varepsilon ,$$

where $0 \leqq \varepsilon \leqq t^{-h}$. We can rewrite this in the form

$$\int_0^1 \{T_{s,t}\xi, \gamma\}d\xi = t^{-h}\sum_{k=1}^{t^h} \{\theta_k^{(h)}, \gamma\} - \varepsilon ,$$

and Lemma 1 follows.

Particularly, for

$$\mu(\gamma, x) = \int_0^1 \{xT_{s,t}\xi, \gamma\}d\xi$$

$$\mu(\gamma, x, y) = \int_0^1 \{xT_{s,t}\xi, \gamma\}\{yT_{s,t}\xi, \gamma\}d\xi$$

we have

$$\mu(\gamma, x) = \lim_{h\to\infty} t^{-h}\sum_{k=1}^{t^h} \{x\theta_k^{(h)}, \gamma\} , \tag{4}$$

$$\mu(\gamma, x, y) = \lim_{h\to\infty} t^{-h}\sum_{k=1}^{t^h} \{x\theta_k^{(h)}, \gamma\}\{y\theta_k^{(h)}, \gamma\} . \tag{5}$$

3. **Exponential sums.** Write $e(\xi)$ for $e^{2\pi i\xi}$. There exist ([5], pp. 91–92, 99) for any γ, $0 \leqq \gamma \leqq 1$, and any $\eta > 0$ functions $f_1(\xi), f_2(\xi)$ periodic in ξ with period 1, such that $f_1(\xi) \leqq \{\xi, \gamma\} \leqq f_2(\xi)$, having Fourier expansions

$$f_1(\xi) = \gamma - \eta + \sum_u{}' A_u^{(1)}e(u\xi)$$

$$f_2(\xi) = \gamma + \eta + \sum_u{}' A_u^{(2)} e(u\xi) ,$$

where the summation is over all $u \neq 0$ and $A_u^{(i)}$ is majorized by

$$| A_u | \leqq \frac{1}{u^2\eta} . \tag{6}$$

Applying this to (5) we obtain

$$\mu(\gamma, x, y) \leqq (\gamma + \eta)^2 + \overline{\lim_{h\to\infty}}\, t^{-h} \sum_{\substack{u,v \\ \neq 0,0}}{}' \left| A_u^{(2)} \right| \left| A_v^{(2)} \right| \left| \sum_{k=1}^{t^h} e((ux + vy)\theta_k^{(h)}) \right| ,$$

where we put $A_0^{(2)} = \gamma + \eta$ and take the sum over all pairs u, v of numbers not both being zero. Since

$$\left| t^{-h} \sum_{k=1}^{t^h} e((ux + vy)\theta_k^{(h)}) \right| \leqq 1 ,$$

and since the double sum over u, v is uniformly convergent in h, we may change the order of limit and summation and obtain

$$\mu(\gamma, x, y) \leqq (\gamma + \eta)^2 + \sum_{u,v}{}' | A_u^{(2)} || A_v^{(2)} | \overline{\lim_{h\to\infty}}\, t^{-h} \left| \sum_{k=1}^{t^h} e((ux + vy)\theta_k^{(h)}) \right| .$$

The numbers $\theta_k^{(h)}$ are the numbers

$$\frac{a_1}{s} + \frac{a_2}{s^2} + \cdots + \frac{a_h}{s^h} ,$$

where $0 \leqq a_i < t$. Hence

$$\sum_{k=1}^{t^h} e(w\theta_k^{(h)}) = \prod_{j=1}^{h} \left(1 + e\left(\frac{w}{s^j}\right) + e\left(\frac{2w}{s^j}\right) + \cdots + e\left(\frac{(t-1)w}{s^j}\right)\right) .$$

If we keep w fixed, and if j is large, then

$$\left| \left(1 + e\left(\frac{w}{s^j}\right) + \cdots + e\left(\frac{(t-1)w}{s^j}\right)\right) t^{-1} - 1 \right| < \frac{t\,|w|}{s^i} .$$

Therefore

$$\Pi\,(s, t\,; w) = \prod_{j=1}^{\infty} \left| \left(1 + e\left(\frac{w}{s^j}\right) + \cdots + e\left(\frac{(t-1)w}{s^i}\right)\right) t^{-1} \right| \tag{7}$$

exists and

$$\mu(\gamma, x, y) \leqq (\gamma + \eta)^2 + \sum_{u,v}{}' | A_u^{(2)} | | A_v^{(2)} | \,\Pi(s, t\,; ux + vy) . \tag{8}$$

The next three sections will be devoted to finding bounds for sums like

$$\sum_{N_1 < n,m \leqq N_2} \Pi(s, t\,; ur^n + vr^m) .$$

4. Two lemmas on digits.

LEMMA 2. *Write $w = c_g \cdots c_2 c_1$ in the scale of s. Assume there are at least z pairs of digits $c_{i+1}c_i$ with*

$$(9) \qquad 1 \leqq c_{i+1}c_i \leqq s^2 - 2 .$$

(Here $c_{i+1}c_i = sc_{i+1} + c_i$). Then

$$\Pi(s, t; w) \leqq \alpha_2^z ,$$

where $\alpha_2 = \alpha_2(s, t)$, $0 < \alpha_2 < 1$.

Proof. There are at least z numbers i having

$$\frac{1}{s^2} \leqq \left\{\frac{w}{s^i}\right\} \leqq 1 - \frac{1}{s^2} .$$

For such an i we have

$$\left|1 + e\left(\frac{w}{s^i}\right) + \cdots + e\left(\frac{(t-1)w}{s^i}\right)\right| \leqq \left|1 + e\left(\frac{1}{s^2}\right)\right| + t - 2 = t\alpha_2$$

and the Lemma is proved.

There exists an $\alpha_3(s)$, $0 < \alpha_3 < 1/4$, such that

$$\frac{(s^2-2)^{\alpha_3} 2^{1/2-\alpha_3}}{(2\alpha_3)^{\alpha_3}(1-2\alpha_3)^{1/2-\alpha_3}} < 2^{3/4} .$$

LEMMA 3. *If k is large, $k > \alpha_4(s)$, then the number of combinations of digits $c_k c_{k-1} \cdots c_1$ in the scale of s with less than $\alpha_3(s)k$ indices i satisfying* (9) *is not greater than $2^{(3/4)k}$.*

Proof. It will be sufficient to show that the number of combinations with less than $\alpha_3(s)k$ indices i satisfying both (9) and $i \equiv 1 \pmod 2$ is not greater than $2^{(3/4)k}$. We first assume k is even. There exist

$$\binom{\frac{k}{2}}{l}(s^2-2)^l 2^{k/2-l}$$

combinations $c_k \cdots c_1$ with exactly l indices i having both (9) and $i \equiv 1 \pmod 2$. Hence the number of combinations with less than $\alpha_3(s)k$ indices i satisfying (9) and $i \equiv 1 \pmod 2$ does not exceed

$$k\binom{\frac{k}{2}}{[\alpha_3 k]}(s^2-2)^{[\alpha_3 k]} 2^{(k/2)-[\alpha_3 k]} .$$

Using Stirling's formula for the binomial coefficient we obtain for large enough k the upper bound

$$\alpha_5(s)k\frac{(s^2-2)^{\alpha_3 k}2^{((1/2)-\alpha_3)k}}{(2\alpha_3)^{\alpha_3 k}(1-2\alpha_3)^{((1/2)-\alpha_3)k}} < 2^{(3/4)k} .$$

Actually, the expression on the left hand side is $< 2^{\alpha_6 k}$, where $\alpha_6 < 3/4$. This permits us to extend the result to odd k.

5. The order of r modulo p^k as a function of k.

LEMMA 4. *Assume p is a prime with $p \nmid r$. Then the order $o(r, p^k)$, of r modulo p^k satisfies*

$$o(r, p^k) \geqq \alpha_7(r, p)p^k .$$

COROLLARY. *Let n run through a residue system modulo p^k. Then at most $\alpha_8(r, p)$ of the numbers r^n will fall into the same residue class modulo p^k.*

Proof. Write

$$g = g(p) = \begin{cases} p-1 , & \text{if } p \text{ is odd} \\ 2 , & \text{if } p = 2. \end{cases}$$

There exists an $\alpha_9 = \alpha_9(r, p)$ such that

$$r^g \equiv 1 + qp^{\alpha_9 - 1} \pmod{p^{\alpha_9}} , \tag{10}$$

where $q \not\equiv 0 \pmod p$. We have necessarily $\alpha_9 > 1$ and even $\alpha_9 > 2$ if $p = 2$. If follows from (10) by standard methods (see, for instance, [4], § 5.5) that

$$r^{gp^e} \equiv 1 + qp^{\alpha_9 - 1 + e} \pmod{p^{\alpha_9 + e}}$$

for any $e \geqq 0$. Thus for $k \geqq \alpha_9$ we have

$$r^{gp^{k-\alpha_9}} \equiv 1 + qp^{k-1} \pmod{p^k}$$

and

$$o(r, p^k) \geqq gp^{k-\alpha_9} = \alpha_7(r, p)p^k .$$

Assume $r \not\sim s$. Write

$$r = p_1^{d_1}p_2^{d_2} \cdots p_h^{d_h}$$
$$s = p_1^{e_1}p_2^{e_2} \cdots p_h^{e_h} ,$$

where we may assume that never both $d_i = 0$, $e_i = 0$. We also may assume that the primes $p_1, \cdots, p_h$ are ordered in such a way that

$$\frac{e_1}{d_1} \geqq \frac{e_2}{d_2} \geqq \cdots \geqq \frac{e_h}{d_h},$$

where we put $(e_i/d_i) = +\infty$ if $d_i = 0$. Since $r \not\sim s$, we have

$$r_1 = \frac{r^{e_1}}{s^{d_1}} > 1 .$$

From now on, $p = p_1(r, s)$ is the prime defined above. We have $p \mid s$ but $p \nmid r_1$. For any $x \neq 0$, $y > 1$ we define two new numbers x_y and x'_y by $x = x_y x'_y$, where x_y is a power of y and $y \nmid x'_y$.

LEMMA 5. A. *Assume $r \not\sim s$, $v \neq 0$. Let m run through a system $K(s^k)$ of non-negative representatives modulo s^k. Then at most*

$$\alpha_{10}(r, s)\left(\frac{s}{2}\right)^k v_p$$

of the numbers

$$v(r^m)'_s$$

are in the same residue class modulo s^k.

B. *Assume $r \not\sim s$, furthermore $p \nmid r$. Suppose $u \neq 0$, $v \neq 0$, n are fixed. Then, if m runs through $K(s^k)$, at most*

$$\alpha_{11}(r, s)\left(\frac{s}{2}\right)^k v_p$$

of the numbers

$$ur^n + vr^m$$

will fall into the same residue class modulo s^k.

Proof. A. Write $m = m_1 e_1 + m_2$, $0 \leqq m_2 < e_1$. Then $r^m = r^{m_1 e_1 + m_2} = s^{m_1 d_1} r_1^{m_1} r^{m_2}$ and $v(r^m)'_s = v r_1^{m_1} (r^{m_2})'_s$. The equation

$$r_1^{m_1} \equiv a \pmod{p^k}$$

has for fixed a at most $e_1 \alpha_8(r_1, p)$ solutions in $m = m_1 e_1 + m_2$, if m runs through a system $K(p^k)$ of residues modulo p^k. This follows from the corollary of Lemma 4. The equation

$$av(r^{m_2})'_s \equiv b \pmod{p^k}$$

has for fixed b, m_2 at most

$$\text{g.c.d.}(v(r^{m_2})'_s, p^k) \leqq v_p r^{m_2}$$

solutions in a. Hence the number of solutions of

$$vr_1^{m_1}(r^{m_2})'_s \equiv b \pmod{p^k}$$

in $m = m_1e_1 + m_2 \in K(p^k)$ does not exceed

$$e_1\alpha_8 v_p(1 + r + \cdots + r^{e_1-1}) = \alpha_{10}(r, s)v_p \ .$$

But this implies that the number of solutions of

$$vr_1^{m_1}(r^{m_2})'_s \equiv b \pmod{s^k}$$

in $m = m_1e_1 + m_2 \in K(s^k)$ is not greater than

$$\alpha_{10}(r, s)v_p\left(\frac{s}{p}\right)^k \leqq \alpha_{10}(r, s)\left(\frac{s}{2}\right)^k v_p \ .$$

B. The equation

$$ur^n + vr^m \equiv b \pmod{p^k}$$

has according to the corollary of Lemma 4 at most

$$\alpha_8(r, p)v_p$$

solutions in $m \in K(p^k)$. The result follows as before.

The following conjecture seems related to our results: *Assume $r \not\sim s$. Then for any ε and k almost all the numbers $r, r^2, \cdots$ are (ε, k)-normal to the base s in the sense of* Besicovitch [1]; *that is, the number of $n \leqq N$ for which r^n is not (ε, k)-normal is $o(N)$ as $N \to \infty$ for fixed ε and k.*

6. Bounds for exponential sums.

LEMMA 6. A. *Let r, s, v be as in Lemma* 5A. *Then*

$$\sum_{m \in K(s^k)} \Pi(s, t\,;\, vr^m) \leqq \alpha_{12} v_p s^{(1-\alpha_{13})k}$$

B. *Let r, s, u, v, n be as in Lemma* 5B. *Then*

$$\sum_{m \in K(s^k)} \Pi(s, t\,;\, ur^n + vr^m) \leqq \alpha_{14} v_p s^{(1-\alpha_{15})k} \ .$$

Proof. A. Write $v(r^m)'_s = c_g \cdots c_k \cdots c_1$ in the scale of s. Lemma 5A implies that any digit combination $c_k c_{k-1} \cdots c_1$ will occur at most $\alpha_{10}(r, s)(s/2)^k v_p$ times. According to Lemma 3, there are for large k not more than $2^{(3/4)k}$ digit-combinations $c_k \cdots c_1$ with less than $\alpha_3 k$ indices i satisfying (9). Thus of all the numbers $v(r^m)'_s$, $m \in K(s^k)$, and hence of all the numbers vr^m there will be at most

$$\alpha_{10}(r, s)(s/2)^k v_p 2^{(3/4)k} = \alpha_{10}(r, s)v_p(s/2^{1/4})^k = \alpha_{10}(r, s)v_p s^{(1-\alpha_{16})k}$$

having less than $\alpha_3 k$ digits c_i in their expansion in the scale of s satisfying (9). Thus Lemma 2 yields

$$\Pi(s, t; vr^m) \leqq \alpha_2^{k\alpha_3}$$

for all but at most

$$\alpha_{10}(r, s)v_p s^{(1-\alpha_{16})k}$$

numbers $m \in K(s^k)$. This gives

$$\sum_{m \in K(s^k)} \Pi(s, t; vr^m) \leqq s^k \alpha_2^{k\alpha_3} + \alpha_{10} v_p s^{(1-\alpha_{16})k} \leqq \alpha_{12} v_p s^{(1-\alpha_{13})k} .$$

B is proved similarly, using Lemma 5B.

LEMMA 7. A. *Assume* $r \not\sim s$, $v \neq 0$. *Then*

(11) $$\sum_{N_1 < n \leqq N_2} \Pi(s, t; vr^m) \leqq \alpha_{17}(N_2 - N_1)^{1-\alpha_{18}} v_p .$$

B. Assume $r \not\sim s$, $u \neq 0$, $v \neq 0$. Then

(12) $$\sum_{N_1 < n, m \leqq N_2} \Pi(s, t; ur^n + vr^m) \leqq \alpha_{19}(N_2 - N_1)^{2-\alpha_{20}} \max(u_p, v_p) .$$

Proof. A. There exists a k having $s^{2k} \leqq N_2 - N_1 < s^{2(k+1)}$, hence there exists a w satisfying $s^k w \leqq N_2 - N_1 < s^k(w+1)$, where $s^k \leqq w < s^{k+2}$. Thus if m runs from N_1 to N_2, then m runs through w systems $K(s^k)$ of residue classes modulo s^k and at most s^k other numbers. Hence by Lemma 6A

$$\sum_{N_1 < m \leqq N_2} \Pi(s, t; vr^m) \leqq w\, \alpha_{12} v_p s^{(1-\alpha_{13})k} + s^k \leqq \alpha_{17}(N_2 - N_1)^{1-\alpha_{18}} v_p .$$

B. If $p \nmid r$, then we proceed as in part A. We first take the sum over m and use Lemma 6B.

If p/r, then our argument is as follows. Consider, for example, the part of the sum with $n \leqq m$. Changing the notation in n, m, we see that this part of the sum (12) equals

$$\sum_{n=0}^{N_2 - N_1 - 1} \sum_{m=N_1+1}^{N_2 - n} \Pi(s, t; (ur^n + v)r^m) .$$

Except for possibly one exceptional n we have $(ur^n)_p \neq v_p$ and therefore $(ur^n + v)_p \leqq v_p \leqq \max(u_p, v_p)$. If n is not exceptional, then the already proved Lemma 7A can be applied to the inner sum and we obtain the bound

$$\alpha_{17}(N_2 - N_1 - n)^{1-\alpha_{18}} \max(u_p, v_p) .$$

Taking the sum over n we obtain (12).

7. A fundamental lemma. Generalizing $M_N(\xi, r, I)$ we write ${}_{N_1}M_{N_2}(\xi, r, I)$ for the number of indices i in $N_1 < i \leqq N_2$ such that $\{r^i\xi\}$ lies in I. We put

$$ {}_{N_1}R_{N_2}(\xi, r, I) = {}_{N_1}M_{N_2}(\xi, r, I) - (N_2 - N_1)|I| . $$

Fundamental lemma. *Assume $r \not\sim s$. Then*

$$ \int_0^1 {}_{N_1}R^2_{N_2}(T_{s,t}\xi, r, I)d\xi \leqq \alpha_{21}(N_2 - N_1)^{2-\alpha_{22}} . $$

Proof. It is enough to prove this for intervals of the type $I = [0, \gamma)$. Then

$$ {}_{N_1}M_{N_2}(\xi, r, I) = \sum_{N_1 < n \leqq N_2} \{r^n\xi, \gamma\} $$

and

$$ \int_0^1 {}_{N_1}M_{N_2}(T_{s,t}\xi, r, I)d\xi = \sum_{N_1 < n \leqq N_2} \mu(\gamma, r^n) \tag{13} $$

$$ \int_0^1 {}_{N_1}M^2_{N_2}(T_{s,t}\xi, r, I)\,d\xi = \sum_{N_1 < n, m \leqq N_2} \mu(\gamma, r^n, r^m) . \tag{14} $$

Now we combine (8) and Lemma 7. We obtain, together with (6),

$$ \begin{aligned} \sum_{N_1 < n, m \leqq N_2} \mu(\gamma, r^n, r^m) &\leqq (\gamma + \eta)^2(N_2 - N_1)^2 \\ &+ 2(\gamma + \eta) \sum_{v \neq 0} \frac{v_p}{\eta v^2} \alpha_{17}(N_2 - N_1)^{2-\alpha_{18}} \\ &+ \sum_{u \neq 0} \sum_{v \neq 0} \frac{\max(u_p, v_p)}{\eta u^2 \eta v^2} \alpha_{19}(N_2 - N_1)^{2-\alpha_{20}} . \end{aligned} $$

Since the sums

$$ \sum_{v \neq 0} \frac{v_p}{v^2} , \qquad \sum_{u \neq 0} \sum_{v \neq 0} \frac{\max(u_p, v_p)}{u^2 v^2} $$

are convergent, and since η was arbitrary, we have

$$ \sum_{N_1 < n, m \leqq N_2} \mu(\gamma, r^n, r^m) - (N_2 - N_1)^2\gamma^2 \leqq \alpha_{23}(N_2 - N_1)^{2-\alpha_{24}} . $$

In the same fashion we can prove

$$ \left| \sum_{N_1 < n, m \leqq N_2} \mu(\gamma, r^n, r^m) - (N_2 - N_1)^2\gamma^2 \right| \leqq \alpha_{23}(N_2 - N_1)^{1-\alpha_{24}} $$

$$ \left| \sum_{N_1 < n \leqq N_2} \mu(\gamma, r^n) - (N_2 - N_1)\gamma \right| \leqq \alpha_{25}(N_2 - N_1)^{1-\alpha_{26}} . $$

These two inequalities, together with (13) and (14), give the Fundamental Lemma.

8. Proof of the theorems. Once the Fundamental Lemma is shown, we can prove Theorem 2 by the standard method developed in [2].

By J_B, $B > 0$, we denote the set of intervals $[\beta, \gamma)$, $0 \leqq \beta < \gamma < 1$ of the type $\beta = a2^{-b}$, $\gamma = (a+1)2^{-b}$, where $0 \leqq b \leqq \alpha_{22}B/2$. By P_B we denote the set of all pairs of integers N_1, N_2 having $0 \leqq N_1 < N_2 \leqq 2^B$ of the type $N_1 = a2^b$, $N_2 = (a+1)2^b$ for integers a and $b \geqq 0$.

LEMMA 8. *Assume $r \not\sim s$. Then*

$$\sum_{(N_1,N_2)\in P_B} \sum_{I\in J_B} \int_0^1 {}_{N_1}R^2_{N_2}(T_{s,t}\xi, r, I)d\xi \leqq \alpha_{27} 2^{2B(1-\alpha_{28})} .$$

Proof. Because of the Fundamental Lemma the left hand side is not greater than

$$\alpha_{21} 2^{\alpha_{22}B/2+1} \Sigma ,$$

where $2^{\alpha_{22}B/2+1}$ is an upper bound for the number of intervals in J_B and

$$\Sigma = \sum_{(N_1,N_2)\in P_B} (N_2 - N_1)^{2-\alpha_{22}} . \tag{15}$$

In (15) each value of $N_2 - N_1 = 2^b$ occurs 2^{B-b} times, so that

$$\Sigma = \sum_{b=0}^{B} 2^{B-b+b(2-\alpha_{22})} \leqq \alpha_{29} 2^{2B(1-\alpha_{22}/2)} .$$

Hence Lemma 8 is true with $\alpha_{28} = \alpha_{22}/4$.

LEMMA 9. *For large B there exists a set E_B of measure not greater than $2^{-\alpha_{30}B}$ such that*

$$R_N(T_{s,t}\xi, r, I) \leqq 2^{B(1-\alpha_{31})} \tag{16}$$

for all I, $N \leqq 2^B$ and all ξ in $[0, 1)$ but not in E_B.

Proof. We define E_B to be the set consisting of all ξ in $[0, 1)$ for which it is not true that

$$\sum_{(N_1,N_2)\in P_B} \sum_{I\in J_B} {}_{N_1}R^2_{N_2}(T_{s,t}\xi, r, I) \leqq 2^{2B(1-\alpha_{28}/2)} . \tag{17}$$

Lemma 8 assures that the measure of E_B does not exceed

$$\alpha_{27} 2^{-2B\alpha_{28}/2} < 2^{-\alpha_{30}B}$$

for large B. We have to show that (16) is a consequence of (17).

We first assume I to be of the type $I = [0, \gamma)$, $\gamma = a2^{-b}$, where $0 \leqq b \leqq \alpha_{22}B/2$. Then the interval $[0, \gamma)$, is the sum of at most $b < B$ intervals I, $I \in J_B$, as may be seen by expressing a in the binary scale.

Expressing N in the binary scale we see that the interval $[0, N)$ can be expressed as a union of at most B intervals $[N_1, N_2)$, where the pair $N_1, N_2 \in P_B$. Hence we can write $R_N(T_{s,t}\xi, r, I)$ as a sum of ${}_{N_1}R_{N_2}(T_{s,t}\xi, r, I)$ over at most B^2 sets N_1, N_2, I, where $N_1, N_2 \in P_B$, $I \in J_B$:

$$R_N(T_{s,t}\xi, r, I) = \Sigma_{N_1}R_{N_2}(T_{s,t}\xi, r, I) .$$

Hence by (17) and Cauchy's inequality,

$$R_N^2(T_{s,t}\xi, r, I) \leqq B^2 2^{2B(1-\alpha_{28}/2)} < 2^{2B(1-\alpha_{32})}$$

for large B.

Next, let $I = [0, \gamma)$ be of the type $a2^{-b} \leqq \gamma \leqq (a+1)2^{-b}$, where $\alpha_{22}B/4 < b \leqq \alpha_{22}B/2$. Then

$$\begin{aligned} &| R_N(T_{s,t}\xi, r, [0, \gamma)) | = | M_N(T_{s,t}\xi, r, [0, \gamma)) - \gamma N | \\ &\leqq | R_N(T_{s,t}\xi, r, [0, (a+1)2^{-b})) | + | R_N(T_{s,t}\xi, r, [0, a2^{-b})) | + 2^{-b}N \\ &\leqq 2 \cdot 2^{B(1-\alpha_{32})} + 2^{(1-\alpha_{22}/4)B} < 2^{B(1-\alpha_{33})} . \end{aligned}$$

The Lemma now follows from

$$| R_N(\ , \ , [\beta, \gamma)) | \leqq | R_N(\ , \ , [0, \beta)) | + | R_N(\ , \ , [0, \gamma)) | .$$

Proof of Theorem 2. Since $\Sigma 2^{-\alpha_{30}B}$ is convergent, there exists for almost all ξ a $B_0 = B_0(\xi)$ such that $\xi \notin E_B$ for $B \geqq B_0$. If $N \geqq 2^{B_0}$, then we can find a $B \geqq B_0$ satisfying $2^{B-1} < N \leqq 2^B$ and Lemma 9 yields

$$R_N(T_{s,t}\xi, r, I) < 2^{B(1-\alpha_{31})} < 2N^{1-\alpha_{31}} < N^{1-\alpha_1}$$

for large enough N.

REFERENCES

1. A. S. Besicovitch, *The asymptotic distribution of numerals in the decimal representation of squares of natural numbers*, Math. Zeit. **39** (1934), 146–156.
2. J. W. S. Cassels, *Some metrical theorems in Diophantine approximation III*, Proc. Camb. Phil. Soc. **46** (1951), 219–225.
3. ———, *On a paper of Niven and Zuckerman*, Pacific J. Math, **2** (1952), 555–557.
4. H. Hasse, *Vorlesungen über Zahlentheorie*, Springer Grundlehren LIX (1950).
5. J. F. Koksma, *Diophantische Approximationen*, Ergebnisse der Math. u. ihrer Grenzgebiete 4, **4** (1936).
6. I. Niven, *Irrational numbers*, The Carus Math. Monographs **11** (1956).
7. I. Niven and H. S. Zuckerman, *On the definition of normal numbers*, Pacific J. Math. **1** (1951), 103–110.
8. D. D. Wall, *Normal numbers*, Ph. D. thesis (1949), University of California, Berkeley.

MONTANA STATE UNIVERSITY

THE METRIZATION OF STATISTICAL METRIC SPACES

B. SCHWEIZER, A. SKLAR AND E. THORP

In a previous paper on statistical metric spaces [3] it was shown that a statistical metric induces a natural topology for the space on which it is defined and that with this topology a large class of statistical metric (briefly, *SM*) spaces are Hausdorff spaces.

In this paper we show that this result (Theorem 7.2 of [3]) can be considerably generalized. In addition, as an immediate corollary of this generalization, we prove that with the given topology a large number of *SM* spaces are metrizable, i.e., that in numerous instances the existence of a statistical metric implies the existence of an ordinary metric.[1]

THEOREM 1.[2] *Let $(S, \mathscr{F})$ be a statistical metric space, $\mathscr{U}$ the two-parameter collection of subsets of $S \times S$ defined by*

$$\mathscr{U} = \{U(\varepsilon, \lambda); \varepsilon > 0, \ \lambda > 0\} ,$$

where

$$U(\varepsilon, \lambda) = \{(p, q); \ p, q \text{ in } S \text{ and } F_{pq}(\varepsilon) > 1 - \lambda\} ,$$

and T a two-place function from $[0, 1] \times [0, 1]$ to $[0, 1]$ satisfying $T(c, d) \geqq T(a, b)$ for $c \geqq a$, $d \geqq b$ and $\sup_{x<1} T(x, x) = 1$. Suppose further that for all p, q, r in S and for all real numbers x, y, the Menger triangle inequality.

$$F_{pr}(x + y) \geqq T(F_{pq}(x), F_{qr}(y)) \tag{1}$$

is satisfied. Then $\mathscr{U}$ is the basis for a Hausdorff uniformity on $S \times S$.

Proof. We verify that the $U(\varepsilon, \lambda)$ satisfy the axioms for a basis for a Hausdorff (or separated) uniformity as given in [2; p. 174–180] (or in [1; II, §1, $n^\circ 1$]).

(a) Let $\Delta = \{(p, p); p \in S\}$ and $U(\varepsilon, \lambda)$ be given. Since for any $p \in S$, $F_{pp}(\varepsilon) = 1$, it follows that $(p, p) \in U(\varepsilon, \lambda)$. Thus $\Delta \subset U(\varepsilon, \lambda)$.

(b) Since $F_{pq} = F_{qp}$, $U(\varepsilon, \lambda)$ is symmetric.

(c) Let $U(\varepsilon, \lambda)$ be given. We have to show that there is a $W \in \mathscr{U}$ such that $W \circ W \subset U$. To this end, choose $\varepsilon' = \varepsilon/2$ and λ' so small that $T(1 - \lambda', 1 - \lambda') > 1 - \lambda$. Suppose now that (p, q) and (q, r) belong to

Received June 12, 1959.

[1] These considerations have led to the study of *SM* spaces which are not metrizable as well as to other natural topologies for *SM* spaces, questions which will be investigated in a subsequent paper.

[2] The terminology and notation are as in [3].

$W(\varepsilon', \lambda')$ so that $F_{pq}(\varepsilon') > 1 - \lambda'$ and $F_{qr}(\varepsilon') > 1 - \lambda'$. Then, by (1),

$$F_{pr}(\varepsilon) \geqq T(F_{pq}(\varepsilon'), F_{qr}(\varepsilon')) \geqq T(1 - \lambda', 1 - \lambda') > 1 - \lambda .$$

Thus $(p, r) \in U(\varepsilon, \lambda)$. But this means that $W \circ W \subset U$.

(d) The intersection of $U(\varepsilon, \lambda)$ and $U(\varepsilon', \lambda')$ contains a member of $\mathscr{U}$, namely $U(\min(\varepsilon, \varepsilon'), \min(\lambda, \lambda'))$.

Thus $\mathscr{U}$ is the basis for a uniformity on $S \times S$.

(e) If p and q are distinct, there exists an $\varepsilon > 0$ such that $F_{pq}(\varepsilon) \neq 1$ and hence ε_0, λ_0 such that $F_{pq}(\varepsilon_0) = 1 - \lambda_0$. Consequently (p, q) is not in $U(\varepsilon_0, \lambda_0)$ and the uniformity generated by $\mathscr{U}$ is separated, i.e., Hausdorff.

Note that the theorem is true in particular for all Menger spaces in which $\sup_{x<1} T(x, x) = 1$. However, it is true as well for many *SM* spaces which are not Menger spaces.

COROLLARY. *If* $(S, \mathscr{F})$ *is an SM space satisfying the hypotheses of Theorem* 1, *then the sets of the form* $N_p(\varepsilon, \lambda) = \{q; F_{pq}(\varepsilon) > 1 - \lambda\}$ *are the neighborhood basis for a Hausdorff topology on* S.

Proof. These sets are a neighborhood basis for the uniform topology on S derived from $\mathscr{U}$.

THEOREM 2. *If an SM space satisfies the hypotheses of Theorem* 1, *then the topology determined by the sets* $N_p(\varepsilon, \lambda)$ *is metrizable.*

Proof. Let $\{(\varepsilon_n, \lambda_n)\}$ be a sequence that converges to $(0, 0)$. Then the collection $\{U(\varepsilon_n, \lambda_n)\}$ is a countable base for $\mathscr{U}$. The conclusion now follows from [2; p. 186].

Theorem 2 may be restated as follows: Under the hypotheses of Theorem 1, there exist numbers $\delta(p, q)$ which are determined by the distance distribution functions F_{pq} in such a manner that the function δ is an ordinary metric on S. Loosely speaking, if the statistical distances have certain properties, then certain numerical quantities associated with them have the properties of an ordinary distance. In a given particular case such quantities might be the means, medians, modes, etc.. For example, most of the particular spaces studied in [3] satisfy the hypotheses of Theorem 2, hence are metrizable. Indeed, as was shown in [3], in a simple space, the means (when they exist), medians, and modes (if unique) of the statistical distances each form metric spaces; and similarly, in a normal space, the means of the F_{pq} form a (generally discrete) metric space. What Theorem 2 now tells us is that in many (though not all!) *SM* spaces we can expect results of this general nature to hold.

REFERENCES

1. N. Bourbaki, *Topologie générale,* Actualités Sci. Ind. 858–1142 (second edition).
2. J. L. Kelley, *General Topology*, Princeton, 1955.
3. B. Schweizer and A. Sklar, *Statistical metric spaces*, Pacific J. Math., **10** (1960), 313-334.

UNIVERSITY OF CALIFORNIA AT LOS ANGELES
ILLINOIS INSTITUTE OF TECHNOLOGY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

ON UNIQUENESS QUESTIONS FOR HYPERBOLIC DIFFERENTIAL EQUATIONS

JOHN P. SHANAHAN

1. Statement of results. This note is concerned with the existence, uniqueness, and successive approximations for solutions of the initial value problem

$$z_{xy} = f(x, y, z, p, q),\ z(x, 0) = \sigma(x),\ z(0, y) = \tau(y)\ ,$$

where $\sigma(0) = \tau(0) = z_0$, on a rectangle $R: 0 \leqq x \leqq a,\ 0 \leqq y \leqq b$. By a solution is meant a continuous function having partial derivatives almost everywhere and satisfying the integral equation

$$(1)\qquad z(x, y) = \sigma(x) + \tau(y) - z_0 + \int_0^x\int_0^y f(s, t, z(s, t), z_x(s, t), z_y(s, t))dsdt.$$

Actually it will be clear from the conditions imposed on σ, τ and f that any solution of (1) is uniformly Lipschitz continuous. Let D be the five-dimensional set $D = \{(x, y, z, p, q,) : (x, y) \in R$ and z, p, q arbitrary$\}$. Let $f(x, y, z, p, q)$ be defined and continuous on D, such that $|f(x, y, z, p, q,)| < N =$ const. for $(x, y, z, p, q) \in D$. Let $\sigma(x)$, $\tau(y)$ be defined and uniformly Lipschitz continuous on $0 \leqq x \leqq a$, $0 \leqq y \leqq b$, respectively (so that $|\sigma(x) - \sigma(\bar{x})| \leqq K|x - \bar{x}|$, $|\tau(y) - \tau(\bar{y})| \leqq K|y - \bar{y}|$ for some constant K) and let $\sigma(0) = \tau(0) = z_0$. In addition, for $(x, y) \in R$ and arbitrary $z, p, q, \bar{z}, \bar{p}, \bar{q}$ assume that

$$(2)\qquad |f(x, y, z, p, q,) - f(x, y, \bar{z}, \bar{p}, \bar{q})| \leqq \varphi(x, y, |z - \bar{z}|, |p - \bar{p}|, |q - \bar{q}|)\ ,$$

where $\varphi(x, y, z, p, q)$ is a continuous, non-negative function defined for $(x, y) \in R$ and non-negative z, p, q, non-decreasing in each of the variables z, p, q, and with the property that for every (α, β), where $0 < \alpha \leqq a$, $0 < \beta \leqq b$, the only solution of

$$(3)\qquad z(x, y) = \int_0^x\int_0^y \varphi(s, t, z(s, t), z_x(s, t), z_y(s, t))dsdt$$

in the rectangle $R_{\alpha\beta}: 0 \leqq x \leqq \alpha,\ 0 \leqq y \leqq \beta$ is $z \equiv 0$.

THEOREM (*). *Under the above assumptions on σ, τ, f and φ,* (1) *possesses one and only one solution on R. This solution is the uniform limit of the successive approximations defined by*

Received June 25, 1959. This research was supported by the United States Air Force through the Air Force Office of Scientific Research of the Air Research and Development Command, under contract No. AF 18 (603)–41. Reproduction in whole or in part is permitted for any purpose of the United States Government.

$$(4_0) \qquad z_0(x, y) = \sigma(x) + \tau(y) - z_0$$

and, for $n = 1, 2, 3, \cdots$, *by*

$$(4_n) \quad z_n(x, y) = z_0(x, y) + \int_0^x \int_0^y f(x, y, z_{n-1}(s, t), z_{n-1\,x}(s, t), z_{n-1\,y}(s, t))\,ds dt\,.$$

The existence assertion of (*) neither implies nor is implied by that in Hartman-Wintner [3] and its generalizations due to Conti, Szmydt, Ciliberto, Kisynski (for references, see [6] and [2]). The uniqueness assertion of (*) can be considered as a crude analogue of Kamke's uniqueness theorem (cf. [5], p. 139) in the theory of ordinary differential equations. Finally, the assertion concerning the convergence of successive approximations is an analogue of a result on ordinary differential equations (cf. Viswanatham [8] and references there to van Kampen, to Wintner and to Dieudonne, and Coddington and Levinson [1]).

A theorem similar to (*), in which f and φ do not depend on p, q is proved by Guglielmino [2]. The proof of (*) below will be a generalization of that of [2]. A uniqueness theorem for (1) involving a majorant function of the form $\varphi(z, p, q) = \varphi(|z| + |p| + |q|)$ is given in [6]. (After the completion of this manuscript, I learned[1] of a paper "On the existence theorem of Caratheodory for ordinary and hyperbolic differential equations" by W. Walter, written at about the same time, which contains a theorem in the direction of the uniqueness assertion of (*). Walter's assumptions, however, are somewhat different.)

REMARK. It will be clear from the proofs that (*) *remains valid if* f, z, p, q, σ, τ *are n-vectors* (say, with the norm $|z| = \sum_{k=1}^n |z^k|$ or $|z| = \max(|z^1|, \cdots, |z^n|)$ if $z = (z^1, \cdots, z^n)$). Of course φ will still be a function of 5 variables, (not of $(3n + 2)$ variables as f is).

A theorem suggested by Nagumo's uniqueness theorem (cf. [5], p. 97) for ordinary differential equations is the following:

THEOREM (**). *Let* $f(x, y, z, p, q)$ *be defined, continuous and bounded on D, and satisfy, for* $xy > 0$ *and arbitrary* $z, p, q, \bar{z}, \bar{p}, \bar{q}$,

$$(5) \quad |f(x, y, z, p, q,) - f(x, y, \bar{z}, \bar{p}, \bar{q}) \leqq c_1(x, y)|z - \bar{z}|/xy + c(x, y)|p - \bar{p}|/y + c_3(x, y)|q - \bar{q}|/x\,,$$

where $c_i(x, y)$, $i = 1, 2, 3$, *are non-negative, continuous functions such that*

$$c_1 + c_2 + c_3 \equiv 1\,.$$

Let $\sigma(x)$, $\tau(y)$ *be as in* (*), *and, in addition, let*

[1] Added in proof, 4 April 1960. Since this paper was accepted for publication, the following related articles have appeared: W. L. Walter, *Ueber die Differentialgleichung* $u_{xy}=f(x, y, u, u_x, u_y)$, I and II, Math. Zeit., **71** (1959), 308–324 and 436–453; my attention has also been called to the paper of J. B. Diaz and W. L. Walter, *On uniqueness theorems for ordinary differential equations and for partial differentiale equatitions of hyperbolic type*, to appear in Trans. A.M.S..

$$(6)\qquad \sigma_x(+0) = \lim_{x\to+0} \sigma_x(x),\ \tau_y(+0) = \lim_{y\to+0} \tau_y(y)$$

exist. Then (1) *has at most one solution* $z = z(x, y)$. *Furthermore, if*

$$(6*)\qquad c_1(0,0) > 0\ ,$$

then a solution exist and is the uniform limit of the successive approximations (4).

In (6), x[or y] tends to $+0$ through the set of values on which σ_x [or τ_y] exists.

Nagumo's theorem follows from Kamke's (with $\varphi(x, y) = y/x$). However (**) does not follow from (*) because $\varphi(x, y, z, p, q)$ is assumed continuous on $x = 0$ and on $y = 0$.

REMARK 1. (**) is valid if f, z, p, q, σ, τ are n-vectors (say $z = (z^1, \cdots, z^n)$ and either $|z| = \sum_{k=1}^n |z^k|$ or $|z| = \max(|z^1|, \cdots, |z^n|)$).

REMARK 2. A modification of an example of Perron [7] in the theory of ordinary differential equations will show that (**) is false if $c_1 = \text{const.} > 1$, $c_2 \equiv c_3 \equiv 0$ (so that f does not depend on p, q). Also, a modification of an example of Haviland [4] shows that successive approximations need not converge if $c_1 = \text{const.} > 1$, $c_2 = c_3 \equiv 0$.

The proof of (*) will be given in §§ 2-4 below; that of (**) in §§ 5-6; finally, the proof of the last remark will be indicated in § 7.

The results above answer some questions suggested by Professor P. Hartman. I also wish the acknowledge helpful discussions with him.

2. Proof of (*). Preliminaries. In the proof of (*) below, there is no loss of generality in supposing that φ is bounded, say $0 \leqq \varphi(x, y\ z, p, q,) \leqq 2N$ on D. For otherwise φ can be replaced by $\bar\varphi$, where $\bar\varphi(x, y, z, p, q)$ equals $\varphi(x, y, z, p, q)$ or $2N$ according as $\varphi(x, y, z, p, q)$ does not or does exceed $2N$. It is clear that $\bar\varphi$ is continuous and non-decreasing in each of the variables z, p, q. Furthermore, the only solution $z(x, y)$ of

$$(3')\qquad z(x, y) = \int_0^x\int_0^y \bar\varphi(s, t, x(s, t), z_x(s, t,), z_y(s, t))dsdt$$

on any rectangle $R_{\alpha\beta} : 0 \leqq x \leqq \alpha(\leqq a),\ 0 \leqq y \leqq \beta(\leqq b)$ is $z \equiv 0$.

In order to see this, note that $\varphi(x, y, 0, 0, 0) \equiv 0$ because $z = 0$ is a solution of (3). Hence there exists an $\varepsilon > 0$ such that $0 \leqq \varphi(x, y, z, p, q) \leqq 2N$ if $|z|, |p|, |q| < \varepsilon$. Suppose that $z(x, y) \not\equiv 0$ is a solution of (3′) on $R_{\alpha\beta}$. Let $d, 0 \leqq d \leqq (\alpha^2 + \beta^2)^{\frac{1}{2}}$, be the largest value of r for which $z(x, y) \equiv 0$ in the intersection S_r of $x^2 + y^2 \leqq r^2$ and $R_{\alpha\beta}$. If U is any neighborhood of S_d (relative to $R_{\alpha\beta}$), there exists a rectangle $R_{\gamma\delta}$ in U on which $z \not\equiv 0$. Since $z \equiv 0$ on S_d, it is clear that if U is "sufficiently small", then, on U (hence on $R_{\gamma\delta}$), $|z| < \varepsilon$ and, almost everywhere, $|z_x| + |z_y| < \varepsilon$. But then $z \not\equiv 0$ is a solution of (3) on $R_{\gamma\delta}$. Since this is impossible, the only solution of (3′) on $R_{\alpha\beta}$ is $z \equiv 0$.

It will be convenient to have the following notation. R_1 denotes a subset (not always the same) of R of the from $E \times [0, b]$, where E is a (Lebesgue) measurable subset of $[0, a]$ with means $E = a$. Similary, R_2 is a subset (not always the same) of the form $[0, a] \times E$, where E is a measurable subset of $[0, b]$ and means $E = b$. Partial derivatives z_x, z_y of a function z will be denoted by p, q.

3. **Lemma for (*).**

The proof of (*) will depend on the following lemma.

LEMMA 1. *Let $\alpha(x, y)$, $\beta(x, y)$ $\gamma(x, y)$ be non-negative, measurable functions defined on R, R_1, R_2, respectively, such that α is continuous, β is uniformly Lipschitz continuous with respect to y and γ is uniformly Lipschitz continuous with respect to x, In addition, let*

$$(7) \qquad \alpha(x, y) \leqq \int_0^x \int_0^y \varphi(s, t, \alpha(s, t), \beta(s, t), \gamma(s,t))dsdt \ ,$$

$$(8) \qquad \beta(x, y) \leqq \int_0^y \varphi(s, t, \alpha(x, t), \beta(x, t), \gamma(x, t))dt \ ,$$

$$(9) \qquad \gamma(x, y) \leqq \int_0^x \varphi(s, y, \alpha(s, y), \beta(s, y), \gamma(s, y))\, ds \ ,$$

where φ satisfies the conditions of (*) and is bounded. Then $\alpha \equiv \beta \equiv \gamma \equiv 0$.

Note that the Lipschitz continuity of β [or α] with respect to y [or x] is assumed to be uniform with respect to x and y.

The proof of the lemma below follows a suggestion made by R. Sacksteder. My original proof, which will be omitted, depended on two results. The first result is an existence theorem for

$$(10) \qquad z(x, y) = \psi(x, y) + \int_0^x \int_0^y \varphi(s, t, z(s, t), p(s, t), q(s, t))dsdt \ ,$$

where ψ is a non-negative, uniformly Lipschitz continuous function which is non-decreasing in x and in y. This existence theorem is proved by using the successive approximations $z_0 = \psi(x, y)$ and

$$(11) \qquad z_n(x, y) = z_0(x, y) + \int_0^x \int_0^y \varphi(s, t, z_{n-1}, p_{n-1}, q_{n-1})dsdt$$

which satisfy

$$(12) \qquad z_n \leqq z_{n+1}, p_n \leqq p_{n+1}, q_n \leqq q_{n+1} \, .$$

The second result is the fact that if ψ is replaced by another function $\bar{\psi}$ with similar properties and, almost everywhere,

$$(13) \qquad \psi \leqq \bar{\psi}, \ \psi_x \leqq \bar{\psi}_x, \ \psi_y \leqq \bar{\psi}_y \ ,$$

then the corresponding solution $\bar{z}$ satisfies

$$(14) \qquad z \leqq \bar{z}, \ p \leqq \bar{p}, \ q \leqq \bar{q} \ .$$

Proof. Define sequences of successive approximations as follows: Let

(15) $$z_0(x, y) = \alpha(x, y),\ u_0(x, y) = \beta(x, y),\ v_0(x, y) = \gamma(x, y)$$

and, for $n \geqq 1$,

(16) $$z_n(x, y) = \int_0^x\int_0^y \varphi(s, t, z_{n-1}(s, t), u_{n-1}(s, t), v_{n-1}(s, t))\,dsdt\ ,$$

(17) $$u_n(x, y) = \int_0^y \varphi(x, t, z_{n-1}(x, t), u_{n-1}(x, t), v_{n-1}(x, t))\,dt,$$

(18) $$v_n(x, y) = \int_0^x \varphi(s, y, z_{n-1}(s, y), u_{n-1}(s, y), v_{n-1}(s, y))ds\ .$$

The functions z_n, u_n, v_n are defined on sets R, R_1, R_2, respectively, which can be taken independent of n. The inequalities (7), (8), (9) give the case $n = 0$ of

(19) $$z_n \leqq z_{n+1},\ u_n \leqq u_{n+1},\ v_n \leqq v_{n+1}\ .$$

The cases $n > 0$ of these inequalities follow by induction by virtue of the monotony of φ.

The boundedness of φ implies the uniform boundedness of the functions z_n, u_n, v_n. Hence, as $n \to \infty$

(20) $$z = \lim z_n,\ u = \lim u_n,\ v = \lim v_n\ ,$$

exist on R, R_1, R_2, respectively. It is clear from (15) and (19), (20) that

(21) $$0 \leqq \alpha \leqq z,\ 0 \leqq \beta \leqq u,\ 0 \leqq \gamma \leqq v\ .$$

Lebesgue's theorem on term-by-term integration under bounded convergence implies

(22) $$z(x, y) = \int_0^x\int_0^y \varphi(s, t, z(s, t), u(s, t), v(s, t))\,dsdt\ ,$$

(23) $$u(x, y) = \int_0^y \varphi(x, t, z(x, t), u(x, t), v(x, t))\,dt\ ,$$

(24) $$v(x, z) = \int_0^x \varphi(s, y\, z(s, y), u(s, y), v(s, y))\,ds\ .$$

It is clear that $z_y = u$, $z_y = v$ almost everywhere. Thus the assumption on φ concerning (3) shows that $z \equiv u \equiv v \equiv 0$. Lemma 1 follows from (21).

4. Proof of (*). (i). Let $z(x, y)$ be a solution of (1). There exist functions $u(x, y)$, $v(x, y)$ defined on sets R_1, R_2, respectively, such that

(25) $$z(x, y) = \sigma(x) + \tau(y) - z_0 + \int_0^x\int_0^y f(s, t, z(s, t), u(s, t), v(s, t)\,dsdt\ ,$$

(26) $$u(x, y) = \sigma_x(x) + \int_0^y f(x, t, z(x, t), u(x, t), v(x, t))\,dt\ ,$$

$$(27)\qquad v(x, y) = \tau_y(y) + \int_0^x f(s, y, z(s, y), z_x(s, y), z_y(s, y))ds ,$$

and the relations $u = z_x$ and $v = z_y$ hold almost everywhere. In order to see this, note that almost everywhere on R,

$$z_x(x, y) = \sigma_x(x) + \int_0^y f(x, t, z(x, t), z_x(x, t), z_y(x, t))dt ,$$

$$z_y(x, y) = \sigma_y(y) + \int_0^x f(s, y, z(s, y), z_x(s, y), z_y(s, y))ds ,$$

The expressions on the right side of these equations are defined for (x, y) on sets R_1, R_2, respectively. Define $u(x, y), v(x, y)$ to be these expressions on R_1, R_2. In particular $z_x = u$ and $z_y = v$ almost everywhere. Hence (26), (27) hold on (possibly different) sets R_1, R_2. Clearly (25) is valid for all (x, y) on R.

(ii). *Uniqueness in* (*). Suppose that (1) possesses two solutions $z = z_1(x, y), z_2(x, y)$ on R. Let $u_1(x, y), v_1(x, y)$ and $u_2(x, y), v_2(x, y)$ be the functions associated with z_1, z_2 by (i). Let $\alpha = |z_1 - z_2|$, $\beta = |u_1 - u_2|$, $\gamma = |v_1 - v_2|$. If the relations (25) for $z = z_1, z_2$ are subtracted, it is seen that the inequality (2) for f implies (7). Similarly (26), (27) imply (8), (9) respectively.

The functions α, β, γ satisfy the assumptions of Lemma 1. Hence the uniqueness assertion in (*) follows from Lemma 1.

(iii). *Existence and successive approximations.* Let $z_0(x, y), z_1(x, y)$, $\cdots$ be the successive approximations defined by (4). Corresponding to each $z_n(x, y)$, it is possible to introduce functions $u_n(x, y)$, $v_n(x, y)$ defined on sets R_1, R_2, respectively, and satisfying $u_0 = \sigma_x(x)$, $v_0 = \tau_y(y)$,

$$(28_n)\qquad z_n(x, y) = \sigma(x) + \tau(y) - z_0 + \int_0^x\int_0^y f(s, t, z_{n-1}(s, t), u_{n-1}(s, t), v_{n-1}(s, t))dsdt ,$$

$$(29_n)\qquad u_n(x, y) = \sigma_x(x) + \int_0^y f(x, t, z_{n-1}(x, t), u_{n-1}(x, t), v_{n-1}(x, t))dt ,$$

$$(30_n)\qquad v_n(x, y) = \tau_y(y) + \int_0^x f(s, y, z_{n-1}(x, t), u_{n-1}(s, y), v_{n-1}(x, t))ds .$$

The sets R_1, R_2 can be assumed to be independent of n.

Let $Z_{mn} = |z_m - z_n|$, $U_{mn} = |u_m - u_n|$, $V_{mn} = |v_m - v_n|$ and

$$(31)\qquad \alpha_k(x, y) = \operatorname*{l.u.b.}_{m,n\geqq k} Z_{mn}, \quad \beta_k(x, y) = \operatorname*{l.u.b.}_{m,n\geqq k} U_{mn}, \quad \gamma_k(x, y) = \operatorname*{l.u.b.}_{m,n\geqq k} V_{mn} .$$

It is clear that Z_{mn}, U_{mn}, V_{mn} are uniformly Lipschitz continuous with respect to $(x, y), x, y$, respectively, and that a corresponding statement holds for $\alpha_k, \beta_k, \gamma_k$.

By subtracting the relation (28_n) from (28_{n-1}) and using the inequality (2) for f, it is seen that

$$Z_{mm}(x, z) \leqq \int_0^x\int_0^y \varphi(s, t, Z_{m-1\ n-1}(s, t), U_{m-1\ n-1}(s, t), V_{m-1\ n-1}(s, t))\,dsdt\ .$$

Thus, if $m, n \geqq k$, the monotony of φ shows that

$$Z_{mn}(x, y) \leqq \int_0^x\int_0^y \varphi(s, t, \alpha_{k-1}(s, t), \beta_{k-1}(s, t), \gamma_{k-1}(s, t))\,dsdt\ .$$

Hence

$$\alpha_k(x, y) \leqq \int_0^x\int_0^y \varphi(s, t, \alpha_{k-1}(s, t), \beta_{k-1}(s, t), \gamma_{k-1}(s, t))\,dsdt\ .$$

Similarly

$$\beta_k(x, y) \leqq \int_0^y \varphi(x, t, \alpha_{k-1}(x, t), \beta_{k-1}(x, t), \gamma_{k-1}(x, t))\,dt\ ,$$

$$\gamma_k(x, y) \leqq \int_0^x \varphi(s, y, \alpha_{k-1}(s, y), \beta_{k-1}(s, y), \gamma_{k-1}(s, y))\,ds\ .$$

By (31), the sequences $\{\alpha_k(x, y)\}$, $\{\beta_k(x, y)\}$, $\{\gamma_k(x, y)\}$ are non-increasing (and non-negative). Let $\alpha(x, y)$, $\beta(x, y)$, $\gamma(x, y)$ denote the respective limits of these sequence, The Lipschitz continuity of α_k, β_k, γ_k is preserved under the limiting process. Lebesgue's theorem on term-by-term integration under bounded convergence gives the inequalities (7), (8), (9). Hence Lemma 1 shows that $\alpha \equiv 0, \beta \equiv 0, \gamma \equiv 0$ on R, R_1, R_2, respectively. This implies the existence of the functions $z = \lim z_n$, $u = \lim u_n$, $v = \lim v_n$ on R_1, R_2, as $n \to \infty$, satisfying (25), (26), (27). It is clear that the limit function $z(x, y)$ is a solution of (1).

Finally, the equicontinuity of the functions $z_n(x, y)$ (implied by their uniform Lipschitz continuity) shows that $z(x, z)$ is the *uniform* limit of the $z_n(x, y)$. This proves (*).

5. Lemma for (**). The proof of (**) will depend on the following lemma:

LEMMA 2. *Let $\alpha(x, y)$, $\beta(x,y)$, $\gamma(x, y)$ be non-negative, measurable functions defined on R, R_1, R_2, respectively, so that α is continuous, β is uniformly Lipschitz continuous with respect to y and γ is uniformly Lipschitz continuous with respect to x. Furthermore, assume that*

(32) $$\alpha(x, y)/xy \to 0 \text{ as } 0 < xy \to 0$$

and that, uniformly with respect to x and y, respectively,

(33) $$\beta(x, y)/y \to 0 \text{ as } y \to 0 \text{ and } \gamma(x, y)/x \to 0 \text{ as } x \to 0\ .$$

Finally, suppose that

(34) $$\alpha(x, y) \leqq \int_0^x\int_0^y \{c_1(s, t)\alpha(s, t)/st + c_2(s, t)\beta(s, t)/t + c_3(s, t)\gamma(s, t)/s\}\,dsdt\ ,$$

$$\beta(x, y) \leqq \int_0^y \{c_1(x, t)\alpha(x, t)/xt + c_2(x, t)\beta(x, t)/t + c_3(x, t)\gamma(x, t)/x\}\,dt\,, \tag{35}$$

$$\gamma(x, y) \leqq \int_0^x \{c_1(s, y)\alpha(s, y)/sy + c_2(s, y)\beta(s, y)/y + c_3(s, y)\gamma(s, y)/s\}\,ds\,, \tag{36}$$

where c_1, c_2, c_3 are as in the first part of (**). *Then* $\alpha \equiv \beta \equiv \gamma \equiv 0$.

Proof. By (32), if $\alpha(x, y)/xy$ is defined as 0 when $xy = 0$, it becomes a continuous function on R. Hence, it assumes its maximum M_1 at some point $(x^1, y^1) \in R$. Let $M_2 = \text{l.u.b. } \beta(x, y)/y$ and $M_3 = \text{l.u.b. } \gamma(x, y)/x$ for $(x, y) \in R$.

Note that there exist numbers M_{jk}, where $j, k = 1, 2, 3$, satisfying

$$M_{jk} \geqq 0 \text{ and } \sum_{k=1}^{3} M_{jk} = 1 \qquad \text{for } j = 1, 2, 3\,, \tag{37}$$

and

$$M_j \leqq \sum_{k=1}^{3} M_{jk} M_k\,. \tag{38_j}$$

If $M_1 \neq 0$, then $M_1 = \alpha(x^1, y^1)/x^1y^1$ holds for some point (x^1, y^1) of R with $x^1y^1 > 0$. In this case, (38_1) follows from (34) with $(x, y) = (x^1, y^1)$ if

$$M_{1k} = (x^1y^1)^{-1} \int_0^{x^1}\int_0^{y^1} c_k(s, t)\,ds dt\,. \tag{39}$$

If $M_1 = 0$, let $M_{1k} = c_k(0, 0)$.

In order to obtain (38_2), let (x_j, y_j), where $j = 1, 2, \cdots$, be points of R such that $\lim (x_j, y_j) = (x^2, y^2)$ exists, $\lim \beta(x_j, y_j)/y_j = M_2$ and $\lim \beta(x_j, y) = \beta(y)$ exists uniformly for $0 \leqq y \leqq b$. Then (35) leads to (38_2) with

$$M_{2k} = (y^2)^{-1} \int_0^{y^2} c_k(x^2, t)\,dt \text{ or } M_{2k} = c_k(x^2, 0) \tag{40}$$

according as $y^2 > 0$ or $y^2 = 0$. A relation of the type (38_3) is obtained similarly.

Let $M_J = \max(M_1, M_2, M_3)$. Suppose, if possible, that $M_J > 0$. Assume, for the moment, that $M_J > M_j$ if $j \neq J$. Then, by (37) and (38_J), $M_{JJ} = 1$ and $M_{Jk} = 0$ for $k \neq J$. But the derivation of (38_J) can then be modified to obtain $M_J < M_J$. For example, if $J = 1$, then $c_1(s, t) \equiv 1$ and $c_2(s, t) = c_3(s, t) = 0$ in (34) when $(x, y) = (x^1, y^1)$, while $\alpha(s, t)/st$ is nearly zero for small st, so that one obtains $M_1 < M_1$. Or if $J = 2$, then $y^2 > 0$ and $c_1(x^2, t) = 1, c_2(x^2, t) = c_3(x^2, t) = 0$ for $0 \leqq t \leqq y^2$, while the relations

$$\beta(y) \leqq \int_0^y \beta(t)\,dt/t, \qquad \beta(y^2)/y^2 = M_2$$

give $M_2 < M_2$ since $\beta(t)/t$ is nearly 0 for small t by the uniformity of the first limit relation in (33).

Similar arguments show that if two or three of the numbers M_1, M_2, M_3 are equal to $M_J > 0$, one is led to a contradiction. Hence $M_J = 0$. This proves the lemma.

6. Proof of (**). (i). *Uniqueness* in (**). Let $z = z_1(x, y)$, $z_2(x, y)$ be two solutions of (1) on R. Let $u_1(x, y)$, $v_1(x, y)$ and $u_2(x, y)$, $v_2(x, y)$ be the functions associated with them as in the proof of (*). Let $\alpha = |z_1 - z_2|$, $\beta = |u_1 - u_2|$, $\gamma = |v_1 - v_2|$. It will be verified that, as x (or y) $\to 0$, then, except for sets of measure zero,

$$\alpha(x, y), \beta(x, y), \gamma(x, y) \to 0 . \tag{41}$$

Consider the case $x \to 0$. The assertions (41) concerning α and γ are clear. In order to verify assertion (41) for the function β, it will first be shown that if $z = z(x, y)$ is any solution of (1) (say, $z = z_1$ or $z = z_2$) and if $u(x, y)$ $v(x, y)$ are its associated functions, then

$$\lim u(x, y) = \rho(y), \text{ as } x \to 0, \text{ exists uniformly in } y . \tag{42}$$

To see this, let x_j, where $j = 1, 2, 3, \cdots$ be a sequence of x values such that $\lim x_j = 0$ and $\lim u(x_j, y) = \rho(y)$ exists uniformly as $j \to \infty$. Putting $x = x_j$ in (26) and letting $j \to \infty$, it is seen that

$$\rho(y) = \sigma_x(+0) + \int_0^y f(0, t, \tau(t), \rho(t), \tau_y(t))\,dt . \tag{43}$$

We note that $\rho(y)$ is continuous. Furthermore, $\rho(y)$ does not depend on the sequence $x_1, x_2, \cdots$. Suppose that another sequence leads to a different limit $\bar{\rho}(y) \not\equiv \rho(y)$. By substituting $\bar{\rho}$ for ρ in (43), and subtracting, we get

$$|\bar{\rho}(y) - \rho(y)| \leqq \int_0^y |f(0, t, \tau(t), \bar{\rho}(t), \tau_y(t)) - f(0, t, \tau(t), \rho(t), \tau_y(t))|\,dt . \tag{44}$$

Since f, ρ, $\bar{\rho}$ are continuous and $\rho(0) = \bar{\rho}(0) = \sigma_x(+0)$, the integrand of (44) can be made small by making y small. Hence

$$|\bar{\rho}(y) - \rho(y)|/y \to 0, \text{ as } y \to 0 . \tag{45}$$

By relation (5),

$$|\bar{\rho}(y) - \rho(y)|/y \leqq y^{-1} \int_0^y c_2(0, t)|\bar{\rho}(t) - \rho(t)|\,dt/t ,$$

Using (45) as before, this leads to a contradiction. Hence $\bar{\rho} \equiv \rho$. Therefore every sequence, for which the limit in (42) exists, leads to the same limit. Hence (42) holds.

If $\lim u_1(x, y) = \rho_1(y)$ and $\lim u_2(x, y) = \rho_2(y)$, as $x \to 0$, we can repeat the above argument and obtain $\rho_1 \equiv \rho_2$. This completes the verification of (41).

We now verify assumptions (32) and (33) of Lemma 2. Consider, for example, the assertion

$$\beta(x, y)/y \to 0 \text{ as } y \to 0 . \tag{46}$$

By putting $u = u_1, u_2$ in (26) and subtracting we get

$$\beta(x, y) \leqq \int_0^y |f(x, t, z_1(x, t), u_1(x, t), v_1(x, t)) - f(x, t, z_2(x, t), u_2(x, t), v_2(x, t))| dt . \tag{47}$$

Now the integrand of (47) can be made small, by making y small, and using (41). This proves (46). The other limits in (32) and (33) are verified similarly. The other assumptions of Lemma 2 are quite straightforward. Therefore $\alpha \equiv \beta \equiv \gamma \equiv 0$. This proves "uniqueness".

(ii). *Existence and successive approximations in* (**). Let $z_0(x, y)$, $z_1(x, y), \cdots$, be the successive approximations defined by (4). Corresponding to $z_n(x, y)$ it is possible to introduce, as in the proof of (*), functions $u_n(x, y)$, $v_n(x, y)$ defined on sets R_1, R_2 (independent of n) and satisfying $u_0 = \sigma_x(x)$, $v_0 = \tau_y(y)$, (28_n), (29_n) and (30_n). Let Z_{mn}, U_{mn}, V_{mn} be defined as in the existence proof (*) above. It will be verified that, given ε, there exists a $\delta(\varepsilon)$ and an $N(\varepsilon)$, such that

$$Z_{mn}(x, y),\ U_{mn}(x, y),\ V_{mn}(x, y) < \varepsilon \tag{48}$$

for $x < \delta(\varepsilon)$ and for all $m, n > N(\varepsilon)$. A similar statement will be seen to hold when x is replaced by y. The assertion (48) concerning Z_{mn} and V_{mn} is clear. In order to verify (48) for the function U_{mn} it will first be shown that

$$\lim u_n(x, y) = h_n(y), \text{ as } x \to 0, \text{ exists uniformly in } y \text{ and } n . \tag{49}$$

It is easily verified, by induction, that $h_n(y)$ exists uniformly in y for fixed n, where

$$h_n(y) = \sigma_x(+0) + \int_0^y f(0, t, \tau(t), h_{n-1}(y), \tau_y(t)) dt . \tag{50_n}$$

To see the uniformity in n, define

$$\bar{z}_n(x, y) = z_n(x, y) - \sigma(x) - \tau(y) + z_0;\ \bar{u}_n(y, y) = u_n(y, y) - \sigma_x(y); \quad \bar{v}_n(x, y) = v_n(x, y) - \tau_y(y) ; \tag{51_n}$$

$$g(x, y, z, p, q) = f(x, y, z + \sigma(x) + \tau(y) - z_0, p + \sigma_x(x), q + \tau_y(y)) . \tag{52}$$

For $\bar{u}_n$ we define $\bar{h}_n$ corresponding to h. Clearly g satisfies a condition analogous to (5), $\bar{u}_0(x, y) = \bar{h}_0(y) \equiv 0$, and

$$\bar{u}_n(x, y) = \int_0^y g(x, t, \bar{z}_{n-1}(x, t), \bar{u}_{n-1}(x, t), \bar{v}_{n-1}(x, t)) dt,\ n \geqq 1 \tag{53_n}$$

$$\bar{h}_n(y) = \int_0^y g(0, t, 0, \bar{h}_{n-1}(t), 0) dt,\ n \geqq 1 . \tag{54_n}$$

To prove (49) it suffices to verify that

(55) $\lim \bar{u}_n(x, y) = \bar{h}_n(y)$, as $x \to 0$, exists uniformly in y and n.

By subtracting (54_n) from (53_n), it is seen that

$$|\bar{u}_n(x, y) - \bar{h}_n(y)| \leqq \int_0^y \{|g_1 - g_2| + |g_2 - g_3|\} dt \tag{56}$$

where $g_1 = g(x, t, \bar{z}_{n-1}(x, t), \bar{u}_{n-1}(x, t), \bar{v}_{n-1}(x, t))$, $g_2 = g(0, t, 0, \bar{u}_{n-1}(x, t), 0)$ and $g_3 = g(0, t, 0, \bar{h}_{n-1}(t), 0)$. We note that, given $\varepsilon > 0$, there exists a $\delta(\varepsilon)$ such that $|g_1 - g_2| < \varepsilon$ if $x < \delta$ for all y and n. Hence, noting (5),

$$|\bar{u}_n(x, y) - \bar{h}_n(z)| \leqq \int_0^y \{\varepsilon + t^{-1} c_2(0, t)|\bar{u}_{n-1}(x, t) - \bar{h}_{n-1}(t)|\} dt . \tag{57_n}$$

By continuity, because of (6∗), $c_2(0, t) < 1$ for small $t > 0$. Hence there exists a number θ, $0 < \theta < 1$, such that

$$\int_0^y c_2(0, t)\, dt \leqq \theta y \text{ for } 0 < y \leqq b .$$

A simple induction shows that

$$|\bar{u}_n(x, y) - \bar{h}_n(y)| \leqq (1 - \theta^n)\, \varepsilon\, y/(1 - \theta) \leqq b\, \varepsilon\, /(1 - \theta) . \tag{58}$$

This proves (55). Hence (49) is established.

Next we note that $h_n(y)$, $n = 0, 1, 2, \cdots$, are the successive approximations for the initial value problem

$$dw/dt = F(t, w),\ w(0) = \sigma_x(+ 0) , \tag{59}$$

where $F(t, w) = f(0, t, \tau(t), w, \tau_y(t))$ is bounded, measurable and continuous in w (for almost all fixed t). By (5),

$$|F(t, w) - F(t, \bar{w})| \leqq |w - \bar{w}|/t . \tag{60}$$

Note that the existence of $\tau_y(+ 0)$ implies that $F(t, w) \to F(0, w) = f(0, 0, \tau(0), w, \tau_y(+ 0))$ as $t \to + 0$. The proof of the main theorem in [8] shows that these successive approximations converge uniformly, (60) being Nagumo's uniqueness condition (cf. [5], p. 97). Hence

(61) $\lim h_n(y) = h(y)$, exists uniformly in y as $n \to \infty$.

Now (61) and (49) together give (48) for $U_{mn}(x, y)$. Hence (48) is established.

By an argument similar to that used in verifying (46) it is seen that, given $\varepsilon > 0$, there exists $\delta(\varepsilon)$ such that

$$\begin{aligned} &(xy)^{-1} Z_{mn}(x, y) < \varepsilon \text{ for } xy < \delta(\varepsilon) \text{ and for } m, n > N(\varepsilon) \\ &x^{-1} U_{mn}(x, y) < \varepsilon \text{ for } x < \delta(\varepsilon) \text{ and for } m, n > N(\varepsilon) \\ &y^{-1} V_{mn}(x, y) < \varepsilon \text{ for } y < \delta(\varepsilon) \text{ and for } m, n > N(\varepsilon) . \end{aligned} \tag{52}$$

Now defining $\alpha_k, \beta_k, \gamma_k$ as in (31), we note that we can substitute them for Z_{mn}, U_{mn}, V_{mn}, respectively, in (62) changing $m, n > N(\varepsilon)$ to $k > N(\varepsilon)$. Proceeding as in the analogous section of the proof of theorem (*), we conclude that α, β, γ, satisfy (34), (35) and (36), also (32) and (33). Therefore, by Lemma 2, the successive approximations converge uniformly to a solution of (1).

7. Counter-examples. (a). Let $a = b = 1, 1 + \varepsilon = \delta^2$, $\varepsilon > 0, \delta > 1$. Let $f(x, y, z, p, q)$ be independent of p, q and defined by

$$f(x,y,z,p,q) = \begin{cases} 0 & \text{if } (x, y) \in R, z \leqq 0 , \\ (1+\varepsilon)z/xy & \text{if } (x, y) \in R, 0 < z < (xy)^\delta , \\ (1+\varepsilon)(xy)^{\delta-1} & \text{if } (x, y) \in R, (xy)^\delta \leqq z . \end{cases}$$

Then $f(x, y, z, p, q)$ is continuous and satifies (5) for $c_1(x, y) = 1 + \varepsilon$, (and $c_2 = c_3 \equiv 0$). Let $\sigma(x) = \tau(y) \equiv 0$. Then (1) has an infinity of solutions, namely, $z = c(xy)^\delta$, where $0 < c < 1$.

(b). Let $a = b = 1$, $R^0 = \{(x, y) : 0 < x, \ y \leqq 1\}$, $1 + \varepsilon = \delta^2$, $\varepsilon > 0$, $\delta > 0$ and

$$f(x,y,z,p,q) = \begin{cases} 0 & \text{if } x = 0, y = 0 , \\ (xy)^{\delta-1} & \text{if } (x, y) \in R^0, z < 0 , \\ (xy)^{\delta-1} - (1+\varepsilon)z/xy & \text{if } (x, y) \in R^0, 0 \leqq z \leqq (xy)^0 , \\ -\varepsilon(xy)^{\delta-1} & \text{if } (x, y) \in R^0, (xy)^\delta < z . \end{cases}$$

Then $f(x, y, z, p, q)$ satisfies the same relation (5) as in example (a). However, in (4), $z_{2n} = 0$, $z_{2n+1} = (xy)^\delta/\delta^2$, so that successive approximations (4) do not converge.

References

1. E. A. Coddington and N. Levinson, *Uniqueness and the convergence of successive approximations*, J. Indian Math. Soc,. **16** (1952), 75–81.
2. F. Guglielmino, *Sulla risoluzione del problema de Darboux per I'equazione* $s=f(x, y, z,)$, Bollettino della Unione Matematica Italiana, **13** (1958), 308–318.
3. P, Hartman and A. Wintner, *On hyperbolic partial differential equations*, Amer. J. Math. **74** (1952), 832–864.
4. E. K. Haviland, *A note on the convergence of the successive approximations to the solution of an ordinary differential equation*, Amer. J. Math. **54** (1932), 632–634.
5. E. Kamke, *Differentialgleichungen reeler Funktionen*, Leipzig (1930).
6. J. Kisynski, *Sur I'existence et l'unicité des solutions des problemes classiques relatifs á l'equation* $s=F(x, y, z, p, q,)$, Annales Mariae Curie-Sklodowska, **11** (1957), 73–112.
7. O. Perron, *Eine hinreichende Bedingung für die Unität der Lösung von Differentialgleichungen erster Ordnnug*, Mathematische Zeitschrift, **28** (1928), 216–219.
8. B. Viswanatham, *The general uniqueness theorem and successive approximations*, J.Indian Math. Soc., **16** (1952), 69–74.

THE JOHNS HOPKINS UNIVERSITY

SEQUENCES OF COVERINGS

A. H. Stone

1. Introduction. The metrisable spaces S for which S' (the set of limit points of S) is compact, can be characterized as those uniformisable spaces for which the finest uniformity (compatible with the topology) is metrisable (see [5], [1], where further characterizations are given). B. T. Levshenko has shown [4] that they also coincide with the regular spaces in which every point-finite covering[1] can be refined by one of a fixed sequence of point-finite coverings, and that "point-finite" can be replaced throughout by "star-finite" or "locally finite". We shall extend these results (Theorem 2) and obtain an analogue for uniform spaces (Theorem 3). The proofs depend on a criterion for metrisability (Theorem 1) which may be of independent interest since, though not really new in content, it is particularly simple in form.

NOTATION. If $\mathscr{U}$ is a covering of a space S, and $A \subset S$, the star $St(A, \mathscr{U})$ of A in $\mathscr{U}$ is $\bigcup \{U \mid U \in \mathscr{U}, A \cap U \neq \phi\}$. When A is a 1-point set (x), we abbreviate $St((x), \mathscr{U})$ to $St(x, \mathscr{U})$. The covering by the sets $St(U, \mathscr{U})$, $U \in \mathscr{U}$, is denoted by $St(\mathscr{U})$. A covering $\mathscr{U}$ will be called "almost discrete" if only finitely many pairs U, V of sets of $\mathscr{U}$ intersect; such a covering is clearly star-finite (in fact star-bounded) and so locally finite.

2. Metrisation criterion.

THEOREM 1. *A necessary and sufficient condition that a T_0 space S be metrisable is that S have a sequence of coverings $\mathscr{U}_n$, $n = 1, 2, \cdots$, such that, for each $x \in S$, the stars $St(G, \mathscr{U}_n)$ of the open sets $G \ni x$ form a basis for the neighborhoods of x.*

The condition is trivially necessary. To prove it sufficient, we observe first that S is developable—i.e., the stars $St(x, \mathscr{U}_n)$ form a basis for the neighborhoods of each $x \in S$. It follows that S is T_1; for if x, y are distinct points of S, one of them, say x, has a neighborhood $St(x, \mathscr{U}_n)$ not containing y, and then $St(y, \mathscr{U}_n)$ does not contain x. We next show that S is collectionwise normal (see [2]). We may assume that $\mathscr{U}_{n+1}$ refines $\mathscr{U}_n$ (by replacing each $\mathscr{U}_n$ by the "intersection" of the coverings $\mathscr{U}_1, \cdots, \mathscr{U}_n$). Let A_λ $(\lambda \in \Lambda)$ be a discrete collection of closed subsets of S, and for each n and λ put

$$H_{n\lambda} = \bigcup \{U \mid U \in \mathscr{U}_n,\ St(U, \mathscr{U}_n) \text{ meets } A_\lambda \text{ but not } A_\mu \text{ if } \mu \neq \lambda\}\ .$$

Received June 8, 1959.

[1] Throughout this paper, "covering" means "open covering."

Let $P_{n\lambda} = \bigcup \{H_{m\mu} \mid m \leqq n, \mu \neq \lambda\}$, $K_{n\lambda} = H_{n\lambda} - \overline{P}_{n\lambda}$, $H_\lambda = \bigcup \{H_{n\lambda} \mid n = 1, 2, \cdots\}$, $K_\lambda = \bigcup \{K_{n\lambda} \mid n = 1, 2, \cdots\}$; these sets are all open. It is easy to verify that $K_\lambda \cap K_\mu = \phi$ if $\lambda \neq \mu$, that $A_\lambda \subset H_\lambda$, and that $A_\lambda \cap \overline{P}_{n\lambda} = \phi$; hence $A_\lambda \subset K_\lambda$ where the sets K_λ are disjoint and open, as required.

As Bing has proved [2, Th. 10] that every developable collectionwise normal T_1 space is metrisable, the theorem follows. Alternatively Theorem 1 could be deduced from a general theorem of Nagata [6], or from a theorem of F. B. Jones [3].

3. THEOREM 2. *The following statements about a regular T_1 space S are equivalent*:

(1) *S is metrisable and S' is compact,*

(2) *S has a sequence of coverings $\mathscr{G}_n$ $(n = 1, 2, \cdots)$ such that each finite covering of S is refined by some $\mathscr{G}_n$,*

(3) *S has a sequence of almost discrete coverings $\mathscr{G}_n$ $(n = 1, 2, \cdots)$ such that each covering of S is refined by some $\mathscr{G}_n$.*

The implication (3) $\rightarrow$ (2) is trivial. To prove (2) $\rightarrow$ (1), we first show that, assuming (2), S is metrisable. Given $x \in U$ where U is open in S, there is an open set V such that $x \in V$ and $\overline{V} \subset U$. The finite covering $\mathscr{F} = \{V, U - (x), S - \overline{V}\}$ of S has a refinement $\mathscr{G}_n$, and $x \in$ some $G^0 \in \mathscr{G}_n$; then $G^0 \subset V$, the only set of $\mathscr{F}$ which contains x. If $G^1 \in \mathscr{G}_n$ and meets G^0, it follows that $G^1 \subset V \cup (U - (x)) = U$. Thus $St(G^0, \mathscr{G}_n) \subset U$, so Theorem 1 applies and S is metrisable. Let ρ be a metric for S; we construct another, σ, for which each $\mathscr{G}_n$ is uniform. We do this by successively constructing coverings $\mathscr{U}_1, \mathscr{U}_2, \cdots$, such that $St(\mathscr{U}_{n+1})$ refines $\mathscr{U}_n$, $\mathscr{U}_n$ refines $\mathscr{G}_n$, and $\mathscr{U}_n$ consists of sets of ρ-diameters $< 1/n$. By [7, p. 51] there is a corresponding pseudo-metric σ for which each $\mathscr{U}_n$, and so each $\mathscr{G}_n$, is uniform; and as $\sigma(x, y) = 0$ implies $\rho(x, y) = 0$ here, σ is a metric. Condition (2) shows that every finite covering of S is uniform in the metric σ; it follows ([5]; see also [1, Th. 1, (4) $\rightarrow$ (3)] that S' is compact (and *every* covering of S is uniform).

Finally, (1) $\rightarrow$ (3) by the argument in [4], which we sketch for completeness. For each $n = 1, 2, \cdots$, cover S' by a finite system of open sets G_{ni} $(i = 1, 2, \cdots, k_n)$ of diameters $< 1/n$, all meeting S', and adjoin the 1-point sets (x) for each $x \in S - \bigcup \{G_{ni} \mid i = 1, \cdots, k_n\}$ to produce an almost discrete covering $\mathscr{G}_n$ of S. It is easy to see that every covering $\mathscr{U}$ of S is refined by $\mathscr{G}_n$ when n is large enough.

REMARK. To require that S be separable, in (1), would be equivalent to requiring that the coverings $\mathscr{G}_n$ be countable, in (2) and (3).

THEOREM 3. *The following statements about a completely regular T_1 space S are equivalent*:

(1) *S is metrisable,*

(2) *S has a uniformity in which every finite uniform covering is refined by some member of a fixed sequence of (not necessarily uniform) coverings* $\mathscr{G}_n$ *of S,*

(3) *S has a uniformity in which every uniform covering is refined by some member of a fixed sequence of locally finite uniform coverings* $\mathscr{G}_n$ *of S.*

To prove (1) → (3), we use the fact that S is paracompact to take $\mathscr{G}_n$ = a locally finite refinement of the covering of S by "spheres" of radius $1/n$. As (3) → (2) trivially, it remains to deduce (1) from (2). Given a neighborhood N of $x \in S$, there exists a uniform covering $\mathscr{U}$ such that $St(x, \mathscr{U}) \subset N$, and there exist uniform coverings $\mathscr{V}, \mathscr{W}$ such that $St(\mathscr{V})$ refines $\mathscr{U}$ and $St(\mathscr{W})$ refines $\mathscr{V}$. Let $x \in W_0 \in \mathscr{W}$ and $St(W_0, \mathscr{W}) \subset V \in \mathscr{V}$. Write $X = St(W_0, \mathscr{W})$, $Y = \bigcup \{W \mid W \in \mathscr{W}, x \notin W, W \text{ meets } V\}$, $Z = \bigcup \{W \mid W \in \mathscr{W}, W \cap V = \phi\}$. Then $\mathscr{F} = \{X, Y, Z\}$, being refined by $\mathscr{W}$, is a uniform covering of S. Some $\mathscr{G}_n$ refines $\mathscr{F}$; say $x \in G^0 \in \mathscr{G}_n$. Because $X \cap Z = \phi$, it follows by an argument similar to one used in the proof of Theorem 2 that $St(G^0, \mathscr{G}_n) \subset X \cup Y \subset St(V, \mathscr{W}) \subset St(V, \mathscr{V}) \subset St(x, \mathscr{U}) \subset N$; hence S is metrisable, by Theorem 1.

REMARK. The uniformities in (2) and (3) of Theorem 3 will be different in general; that in (3) will be metrisable, while that in (2) need not be. By Theorem 2, not every uniformity on S can arise in (2) or (3) (unless S' is compact), but I have not found any satisfactory description of those which do.

REFERENCES

1. M. Atsuji, *Uniform continuity of continuous functions of metric spaces*, Pacific J. Math. **8** (1958), 11–16.
2. R. H. Bing, *Metrization of topological spaces*, Canadian J. Math. **3** (1951), 175–186.
3. F. Burton Jones, *R. L. Moore's Axiom 1' and metrisation*, Proc. Amer. Math. Soc. **9** (1958), 487.
4. B. T. Levshenko, *On the concept of compactness and point-finite coverings*, Mat. Sbornik. **42**(84) (1957), 479–484.
5. J. Nagata, *On the uniform topology of bicompactifications*, J. Inst. Pol. Osaka City Univ **1** (1950), 28–38.
6. J. Nagata, *A contribution to the theory of metrisation*, ibid. **8** (1957), 185–192.
7. J. W. Tukey, *Convergence and uniformity in topology*, Princeton 1940.

UNIVERSITY OF MANCHESTER,
MANCHESTER, ENGLAND

PROJECTIONS ONTO THE SUBSPACE OF COMPACT OPERATORS

E. O. THORP

Introduction. The purpose of this paper is to establish the following theorem.

THEOREM. *Suppose U and V are Banach spaces and that there are bounded projections P_1 from U onto X and P_2 from V onto Y. Then there are no bounded projections from the space of bounded operators on U into V onto the closed subspace of compact operators, in the following cases:*

1. *X is isomorphic* [1] *to ℓ^p, $1 \leq p < \infty$; Y is isomorphic to ℓ^q, $1 \leq p \leq q \leq \infty$ or c_0 or c.*
2. *X is isomorphic to c_0; Y is isomorphic to ℓ^∞, c_0 or c.*
3. *X is isomorphic to c; Y is isomorphic to ℓ^∞.*

NOTATION. If X and Y are Banach spaces, $[X, Y]$ is the set of bounded linear operators from X into Y. ℓ^∞ is the set of bounded sequences with the sup norm.

A space X is said to have a countable basis if there is a countable subset of elements of X, called a basis, such that each $x \in X$ is uniquely expressible as

$$x = \sum_{i=1}^{\infty} \xi_i \varphi_i$$

in the sense that

$$\lim_{n\to\infty} \| x - \sum_{i=1}^{n} \xi_i \varphi_i \| = 0 .$$

If X and Y are spaces with countable bases (φ_i) and (ψ_i) respectively and A is a bounded linear transformation from X into Y, then A can be represented by an infinite matrix (a_{ij}), with

$$A\varphi_j = \sum_{i=j}^{\infty} a_{ij} \psi_i$$

[2]. In what follows, the basis used for ℓ^p will be given by $\varphi_j = (0, 0, \cdots, 0, 1, 0, 0, \cdots)$ where there is a 1 in the jth place and 0 elsewhere. Similarly for ψ_i. The matrix representations of operators will all be with respect to these bases.

Received April 29, 1959. The author thanks Professor Angus Taylor for proposing this problem and thanks both him and Professor Richard Arens for helpful discussions.

Proof of the theorem. The details of the proof are given below only for $X = \ell^p, 1 \le p < \infty$, and $Y = \ell^q, 1 \le p \le q < \infty$. The proof for the remaining pairs is similar and is indicated in a remark at the end.

DEFINITION. Let E be the function on $[\ell^p, \ell^q], 1 \le p \le q < \infty$, which sends an operator whose matrix is (a_{ij}) into the operator whose matrix is $(a_{ij}\delta_{ij})$, i.e. the non-diagonal matrix elements are replaced by zero and the diagonal elements are unaltered.

LEMMA 1. *E is a projection with $\| E \| = 1$, range the diagonal operators, and null-space the operators with $a_{ii} = 0$, all i.*

Proof. E is additive and homogeneous as easily follows from [2]. $E^2 = E$, and the characterization of the range and null-spaces are apparent.

From the chain

$$\begin{aligned}\infty > \| A \| &= \sup_{\|x\|_p \le 1} \| Ax \|_q \ge \sup_j \| A\varphi_j \|_q \\ &= \sup_j \| \textstyle\sum_i a_{ij} \psi_j \|_q \ge \sup_j \| a_{jj}\psi_j \|_q = \sup_j | a_{jj} | \\ &\ge \sup_{\sum|\xi_i|^p \le 1} (\textstyle\sum | a_{ii}\xi_i |^p)^{1/p} \ge \sup_{\|x\|_p \le 1} (\textstyle\sum | a_{ii}\xi_i |^q)^{1/q} = \| EA \| ,\end{aligned}$$

where the last $\ge$ is by Jensen's inequality, we see that E sends bounded operators into bounded operators and, further, $\| E \| = 1$. Also

$$\| EA \| \le \sup_j | a_{jj} | .$$

In fact,

$$\| EA \| = \sup_j | a_{jj} |$$

because

$$\| EA \| \ge \sup_j \| EA\varphi_j \| = \sup_j | a_{jj} | .$$

LEMMA 2. *The mapping γ from the set of diagonal operators onto ℓ^∞ defined by $\gamma(a_{ii}) = (a_{11}, a_{22}, \cdots)$ is an isometry which carries the compact diagonal operators onto c_0.*

Proof. That γ is an isometry from the diagonal operators onto ℓ^∞ follows from the previous observation that $\| EA \| = \sup_j | a_{jj} |$. Hence it suffices to show that the compact diagonal operators are exactly those with the additional condition $\lim_i | a_{ii} | = 0$. This condition is necessary;

otherwise for some $\varepsilon > 0$ there is an infinite index set I such that $|a_{ii}| \geq \varepsilon$ whenever $i \in I$. Then the bounded sequence $(\varphi_i)_{i\in I}$ would be carried into the sequence $(a_{ii}\psi_i)_{i\in I}$, which has no convergent subsequence, showing (a_{ii}) is not compact. The condition is sufficient because, if $\|x\|_p \leq 1$ then

$$\left(\sum_{i=1}^{\infty}|a_{ii}\xi_i|^q\right)^{1/q} \leq (\sup_{i\geq n}|a_{ii}|)\,\|x\|_q \leq \sup_{i\geq n}|a_{ii}|$$

and [2; Th. 2] applies. The last inequality follows from Jensen's inequality and our assumptions $p \leq q$, $\|x\|_p \leq 1$.

LEMMA 3. *Suppose X is a Banach space with a closed subspace $\mathfrak{M}$ onto which there is a bounded projection E. Let $\mathfrak{N}$ be the null-space of E. Let $\mathfrak{P}$ be any closed linear manifold of X such that if $f \in \mathfrak{P}$ then $f = g + h$, with $g \in \mathfrak{P} \cap \mathfrak{M}$ and $h \in \mathfrak{P} \cap \mathfrak{N}$. Then, given any bounded projection F onto $\mathfrak{P}$, EF is a bounded projection onto $\mathfrak{P} \cap \mathfrak{M}$ such that $\|EF\| \leq \|E\|\,\|F\|$.*

The proof is an obvious modification of [3; Lemma 1.2.1].

Let $\mathfrak{P}$ be the set of compact operators, $\mathfrak{M}$ the set of diagonal operators, E the projection of Lemma 1, and $\mathfrak{N}$ its null-space. In order to apply Lemma 3 it remains to show: given any compact operator f, Ef and $f - Ef$ are compact. Ef is compact because, if f is compact,

$$\lim_n\left\|\sum_{i=n}^{\infty}a_{ij}\psi_i\right\| = \lim_n\left(\sum_{i=n}^{\infty}|a_{ij}|^q\right)^{1/q} = 0$$

uniformly in j. This implies $\lim_n|a_{nn}| = 0$, which shows that Ef is compact. Hence $f - Ef$ is compact.

To prove the theorem for $[l^p, l^q]$, $1 \leq p \leq q < \infty$, assume there is a bounded projection F from $[l^p, l^q]$ onto $\mathfrak{P}$. By Lemma 3, the restriction of EF to $\mathfrak{M}$ is a bounded projection from $\mathfrak{M}$ onto $\mathfrak{M} \cap \mathfrak{P}$. By Lemma 2 there must be a corresponding bounded projection from l^∞ onto c_0. This contradicts [4; Cor. 7.5]. For the remaining X and Y pairs of the main theorem, the proof is similar except that the existence of expressions for $\|A\|$ in terms of the matrix coefficients (e.g., see [5]) makes some of the work simpler.

Next we extend the theorem to $[U, V]$. Let $\tilde{E}$ be the function on $[U, V]$ defined by $\tilde{E}f = P_2fP_1$ for all f in $[U, V]$. $\tilde{E}$ is linear and homogeneous and bounded. $\tilde{E}^2f = P_2(P_2fP_1)P_1 = P_2fP_1 = \tilde{E}f$ so $\tilde{E}$ is a projection. The range of $\tilde{E}$ is the set of operators g such that $P_2gP_1 = g$ and is isomorphic with $[X, Y]$. The null-space of $\tilde{E}$ is the set of operators h such that $P_2hP_1 = 0$. If Q_i is the projection $I - P_i$, the decomposition $f = g + h$ is given by

$$f = (P_2 + Q_2)f(P_1 + Q_1) = \underbrace{P_2 f P_1}_{g} + \underbrace{(P_2 f Q_1 + Q_2 f P_1 + Q_2 f Q_1)}_{h} .$$

If f is compact, so are g and h. We apply Lemma 3 with $X = [U, V]$, $\mathfrak{M}$ the range of $\tilde{E}$, $\tilde{E}$ acting as the projection E of that lemma, and $\mathfrak{P}$ the set of compact operators from U to V. The conclusion is that if there were a bounded projection F from X to $\mathfrak{P}$, the restriction of $\tilde{E}F$ to $\mathfrak{M}$ would be a bounded projection from $\mathfrak{M}$ onto $\mathfrak{P} \cap \mathfrak{M}$, contradicting our result for $[X, Y]$.

REMARK. The problem of finding a bounded projection onto the compact operators is trivial when all the bounded operators are compact. This happens, for example, for $[\ell^p, \ell^q]$, $\infty > p > q \geq 1$, [2, p. 700], or $p = \infty$, $q = 1$, and for $[c_0, \ell^q]$, $[c, \ell^q]$, $\infty > q \geq 1$. Whether there exists a pair of normed spaces with a bounded proper projection from the bounded operators onto the compact operators seems to be unknown.

REFERENCES

1. S. Banach, *Theorie des opérations linéarès* Warsaw, 1932.

2. L. W. Cohen and N. Dunford, *Transformations in sequence spaces*, Duke Math. J., **3** (1937), 689–701.

3. F. J. Murray, *On complementary manifolds and projections in spaces L_p and ℓ_p*, Trans. Amer. Math. Soc., **41** (1937), 138–152.

4. R. S. Phillips, *On linear transformations*, Trans. Amer. Math. Soc., **48** (1940), 516–541.

5. A. E. Taylor, *Introduction to functional analysis*, John Wiley and Co., (1958).

UNIVERSITY OF CALIFORNIA, LOS ANGELES, CALIFORNIA

CONCERNING CERTAIN LOCALLY PERIPHERALLY SEPARABLE SPACES

L. B. TREYBIG

In 1954, F. Burton Jones raised the question [2] "Is every connected, locally peripherally separable [3], metric space separable?" In this paper it will be shown that there exists a connected, semi-locally-connected, space Σ satisfying R. L. Moore's axioms 0 and C[1], in which every region has a separable boundary, every pair of points is a subset of some separable continuum[2], and the set of all points at which Σ is not locally separable is separable. It will also be shown that every compactly connected, locally peripherally separable, metric space is completely separable.

PART 1

Let S' denote the set of all points of the Euclidean plane E. A square disk in E will be said to be horizontal if it has two horizonta sides. A point set in E will be called an *H-disk* only if that set is a horizontal square disk. By the *width* of a square disk will be meant the length of one of its sides.

Let K denote a definite H-disk of width d. Let $R_0(K)$ *denote the H-disk of* width $d/4$ whose center is on the vertical line that contains the center of K, and whose upper side lies at a distance of $d/16$ below the upper side of K. Let $R_{00}(K)$ and $R_{01}(K)$ denote the H-disks of width $d/8$ whose upper sides are at a distance of $d/32$ above the lower side of $R_0(K)$ and whose centers are on the vertical lines containing the left and right sides, respectively, of $R_0(K)$.

In general, for each positive integer n let $U_n(K)$ denote a collection of 2^n mutually exclusive congruent H-disks such that

(1) $R_{01}(K)$ and $R_{00}(K)$ are the elements of $U_1(K)$,

(2) if n is a positive integer and y is an element of $U_n(K)$, and x and z are H-disks of width $d/4(2)^{n+1}$ whose centers lie on the same vertical lines as the left and right sides of y, respectively, and whose upper sides lie at a distance of $d/32(2)^n$ above the lower side of y, then x and z are elements of $U_{n+1}(K)$.

If n is a positive integer and $R_{x_1x_2\cdots x_n}(K)$ is an elements of $U_n(K)$, then let the elements x and y of $U_{n+1}(K)$ whose centers lie on the same

Presented to the American Mathematical Society, June 15, 1957; received by the editors September 4, 1958. This paper is part of a dissertation submitted to the Graduate School of the University of Texas in partial fulfillment of the requirements for the Ph. D. degree.

[1] The proof that every space which satisfies axioms 0 and C is metric is due to R. L. Moore.

[2] A continuum is a connected, closed set.

vertical lines as the left and right sides of $R_{x_1x_2\cdots x_n}(K)$, respectively, be denoted by $R_{x_1x_2\cdots x_n0}(K)$ and $R_{x_1x_2\cdots x_n1}(K)$, respectively. Let $C(K)$ be a collection to which x belongs if and only if x is $R_0(K)$ or in one of the collections $U_1(K)$, $U_2(K)$, $\cdots$.

Let $L(K)$ denote the H-disk of width $d/8$ whose center is on the same vertical line as the center of K, and whose lower side is at a distance of $3d/16$ above the lower side of K. Let $P_l(K)$ and $P_r(K)$ denote the left and right-hand end points, respectively, of the lower side of $L(K)$. Let $M(K)$ denote the point set such that a point P belongs to it if and only if P is a point of the interval $P_l(K)P_r(K)$ such that there is no nonnegative integer p and positive integer q such that $PP_l(K)/P_l(K)P_r(K) = p/2^q$. Let $I(K)$ denote the collection to which x belongs if and only if x is a vertical interval containing a point of $M(K)$, and with both end points on the boundary of $L(K)$. Let an interval i of $I(K)$ be denoted by $i_x(I(K))$ if and only if it is true that if P is the lowest point of i, then $P_l(K)P/P_l(K)P_r(K) = x$.

Let R denote some definite H-disk. Let $R_0(R)$ be denoted by Q_0; let $R_{00}(R)$ and $R_{01}(R)$ be denoted by Q_{00} and Q_{01}, respectively. Let $R_{000}(R)$, $R_{001}(R)$, $R_{010}(R)$, and $R_{011}(R)$ be denoted by Q_{000}, Q_{001}, Q_{010}, and Q_{011}, respectively, and so forth. Let $C(R)$ be denoted by C_1 and let $I(R)$ be denoted by I_0.

Let C_2 denote the collection to which x belongs if and only if x is an element of $C(y)$, for some element y of C_1 distinct from Q_0. Let $R_0(Q_{00})$ be denoted by $Q_{00,0}$; let $R_{01}(Q_{00})$ be denote by $Q_{00,01}$. In general, let $R_x(Q_y)$ be denoted by $Q_{y,x}$. Also, if Q_x is in C_1 and $x \neq 0$, let $I(Q_x)$ be denoted by I_x.

In general, let C_{n+1} denote the collection to which x belongs if and only if x is an element of $C(y)$, for some elements y of C_n, which, in case x_n is 0, is distinct from $Q_{x_1,x_2,\cdots,x_n}$. Let the element $R_{x_{n+1}}[R_{x_n}[R_{x_{n-1}}[\cdots[R_{x_1}(R)]\cdots]]]$ of C_{n+1} be denoted by $Q_{x_1,x_2,\cdots,x_{n+1}}$. Also if w is the element $Q_{x_1,x_2,\cdots,x_n}$ of C_n and $x_n \neq 0$, then let $I(w)$ be denoted by $I_{x_1,x_2,\cdots,x_n}$. For each n let I_n be the collection to which x belongs if and only if there is an element $Q_{x_1,x_2,\cdots,x_n}$ of C_n such that $x_n \neq 0$ and x is in $I(Q_{x_1,x_2,\cdots,x_n})$.

Let W denote the point set to which a point P belongs if and only if P belongs to C_n^*[3] for each positive integer n. For each positive integer n let B_n denote the collection of all the boundaries of the elements of C_n. The boundary of $Q_{x_1,x_2,\cdots,x_n}$ will be denoted by $J_{x_1,x_2\cdots,x_n}$.

Let S denote $[I_0^* + I_1^* + \cdots] + [B_1^* + B_2^* + \cdots] + W$.

Let C' be a collection to which w belongs if and only if w is R or $I(w)$ is a subset of S and there is a positive integer n such that w is in C_n.

---

[3] C_n^* Means the sum of all the point sets of the collection C_n.

For each positive integer n let H_n denote a collection to which x belongs if and only if x is the common part of S and the interior of some square of $[B_n + B_{n+1} + \cdots]$. For each element $Q_{x_1,x_2,\cdots,x_n}$ of C_n, let the set of all points of S in the interior of $J_{x_1,x_2,\cdots,x_n}$ be denoted by $r_{x_1,x_2,\cdots,x_n}$.

For each positive integer n let K_n denote a collection to which x belongs if and only if, either (1) x is a segment of an arc lying on some square J of $(B_1 + B_2 + \cdots)$, having length less than $1/4^n$ times the perimeter of J, and intersecting no square of the collection $(B_1 + B_2 + \cdots)$ except J, or (2) x is the sum of two straight line segments p and q intersecting at their midpoints and lying on different squares J_p and J_q of $(B_1 + B_2 + \cdots)$, such that p and q each have length less than $1/4^n$ times the perimeters of J_p and J_q, respectively, and such that neither p nor q intersects three squares of $(B_1 + B_2 + \cdots)$.

Suppose x is a positive number such that $i_x[I_{j_1,j_2,\cdots,j_n}]$ is an interval of $I_{j_1,j_2,\cdots,j_n}$. For each positive integer n there exists a unique pair (k_n, x_n) such that k_n is a non-negative integer, x_n is a positive number less than one, and $x = (k_n + x_n)/2^n$. By $i_n[i_x(I_{j_1,j_2,\cdots,j_n})]$ will be meant the vertical interval $i_{x_n}(I(y))$, where y is the H-disk of $U_n[Q_{j_1,j_2,\cdots,j_n}]$ with only k_n disks of $U_n(Q_{j_1,j_2,\cdots,j_n})$ to the left of it.

Suppose, for some y in C', P is the highest point of the interval $i_x(I(y))$. By $R_n(P)$ will be meant the sum of all the sects z such that either

(1) for some positive integer d greater than or eqal to n, z is the subset of $i_d[i_x(I(y))]$ with length $1/2^n$ times the length of $i_d[i_x(I(y))]$ that contains the lowest point of $i_d[i_x(I(y))]$, or

(2) z is the subset of $i_x[I(y)]$ with length $1/2^n$ times the length of $i_x(I(y))$ that contains the highest point of $i_x(I(y))$.

For each positive integer n let L_n denote a collection such that x belongs to it if and only if there exists a positive integer d greater than or equal to n, an element y of C', and an interval of the collection $I(y)$ such that if P denotes the highest point of that interval, then $x = R_d(P)$.

For each positive integer n let N_n denote a collection to which x belongs if and only if either

(1) for some element y of C' there exists an interval i of the collection $I(y)$ such that x is a segment of i of length less than $1/2^n$ times the length of i, or

(2) for some element y of C' there exists an element i of $I(y)$ such that x is a sect lying in i, containing the lowest point of i and of length less than $1/2^n$ times the length of i.

For each positive integer n let G_n denote a collection to which x belongs if and only if it lies in $H_n + K_n + L_n + N_n$. S is the set of all points of Σ. A subset r of S is a *region* in Σ if and only if r belongs to G_1.[4]

R. L. Moore's axioms 0 and C are as follows:

Axiom 0. Every region is a point set.

Axiom C. There exists a sequence $G_1, G_2, \cdots$ such that

(1) for each positive integer n, G_n is a collection such that each element of G_n is of region and G_n covers S,

(2) for each n, G_{n+1} is a subcollection of G_n,

(3) if A is a point, B is a point and R is a region containing A, then there exists a positive integer n such that if x is a region of G_n containing A and y is a region of G_n intersecting x, then

(a) y is a subset of R and

(b) if B is not A, y does not contain B,

(4) if $M_1, M_2, \cdots$ is a sequence of closed point sets such that for each n there exists a region g_n of G_n such that M_n is a subset of $\bar{g}_n$ and for each n M_n contains M_{n+1}, then there is a point common to all the point sets of this sequence.

It is obvious that in the space Σ each region has a countable, and therefore separable, boundary, and that the sequence $G_1, G_2, \cdots$ defined for the space Σ satisfies conditions (1) and (2) of axiom C. It will be shown that it also satisfies conditions (3) and (4) of this axiom.

Suppose that P is a point of W, that $r = r_{x_1, x_2, \cdots, x_n}$ is a region of H_n containing P, and that Q is a point of r distinct from P. If q is a region containing a point of W, then q must belong to H_1. Since each element of C_{n+1} which contains P has a side of length less than or equal 1/4 times the length of a side of $Q_{x_1, x_2, \cdots, x_n}$, and each element of C_{n+2} which contains P has a side of length less than or equal $1/4^2$ times the length of side of $Q_{x_1, x_2, \cdots, x_n}$, and so forth; it is obvious that there is a $d > n$ such that if q is a region of H_d which contains P, then $\bar{q}$ does not intersect Q and is a subset of r. Suppose that x and y are two intersecting regions of G_{n+1} such that x contains P. x belongs to H_{n+1} and is therefore a subset of r. Every region of G_{n+1} which intersects x is a subset of r, so clearly, y is a subset of r.

Now suppose that P is a point of $J_{x_1, x_2, \cdots, x_n}$ of B_n and r is a region containing P, and Q is a point of r distinct from P. There exists a circle J in E with center at P such that every point of S in the interior of J belongs to r, but Q is not in the interior of J. There exists a positive integer d such that $1/4^d$ times the perimeter of any square of $(B_1 + B_2 + \cdots)$ to which P belongs is less than the radius of J, and such that no region of H_d contains P. If R^1 is a region of G_{d+1} containing P, then $\bar{R}^1$ does not contain Q and is a subset of r. If $n > d + 2$

[4] The collection G_1 of regions is a basis for the space Σ.

and x and y are two intersecting regions of G_n such that x contains P, then $x + y$ is a subset of r.

Now suppose that P is a point of $i_x(I(y))$, for y in C', and that r is a region containing P and that Q is a point of r distinct from P.

Case 1. Suppose P is not the highest point of $i_x(I(y))$. There exists a segment t containing P, or a sect in case P is the lowest point of $i_x(I(y))$, such that t is a subset of r and does not contain Q nor the highest point of $i_x(I(y))$. There exists a positive number ε such that every point of $i_x(I(y))$ which is at a distance from P of less than ε lies in t. There exists a positive integer d such that

(1) no region of L_d intersects t and no region of H_d intersects $i_x(I(y))$, and

(2) $1/2^d$ times the length of $i_x(I(y))$ is less than ε. Therefore, if k is a region of G_{d+1} containing P, then $\bar{k}$ is a subset of r and does not contain Q. Also, if x and y are two intersecting regions of G_{d+2} such that x contains P, then $x + y$ is a subset of r.

Case 2. Suppose P is the highest point of $i_x(I(y))$. Whether Q belongs to $i_x(I(y))$ or there is a positive integer p such that Q belongs to $i_p[i_x(I(y))]$ or r is in H_1 and Q does not belong to $i_x(I(y)) + i_1[i_x(I))] + i_2[i_x(I(y))] + \cdots$, there is a positive integer d such that

(1) $R_d(P)$ does not contain Q and is a subset of r, and

(2) no region of H_d contains P. If k is a region of G_{d+1} containing P, then $\bar{k}$ is a subset of r and does not contain Q. Also, if x and y are two intersecting regions of G_{d+3} such that x contains P, then $x + y$ is a subset of r.

Therefore $G_1, G_2, \cdots$ satisfies the third part of axiom C.

Suppose that $M_1, M_2, \cdots$ is a sequence of closed point sets such that

(1) for each n M_n contains M_{n+1}, and

(2) for each n there is a region g_n of G_n such that M_n is a subset of $\bar{g}_n$.

In case, for each n, g_n is in H_n, then by definition of W, there is a point common to $M_1, M_2, \cdots$ because some point of W can be easily shown to be a limit point or point of M_n for each n.

In case there is a positive integer j such that g_j belongs to K_j, then for $n > j, g_n$ belongs to K_n. But $M_j, M_{j+1}, \cdots$ is a sequence of closed and compact point sets such that for $n \geqq j$ M_n contains M_{n+1}. So there is a point common to $M_j, M_{j+1}, \cdots$ and thus common to $M_1, M_2, \cdots$.

In case there is a positive integer j such that g_j belongs to N_j, then for $n > j, g_n$ belongs to N_n. So, for the same reason as in the previous case, there is a point common to $M_1, M_2, \cdots$.

The only case not considered is the one where there is a positive integer j_1 such that, for $n \geqq j_1$, g_n belongs to L_n. In this case g_{j_1} must be $R_{x_1}(P)$ for some point P and positive integer x_1. There is a positive integer $j_2 > j_1$ such that $g_{j_2} = R_{x_2}(P)$, where $x_2 > x_1$. There is a positive integer $j_3 > j_2$ such that $g_{j_3} = R_{x_3}(P)$, for $x_3 > x_2$, and so forth. P is common to the sets $R_{x_1}(P), R_{x_2}(P), \cdots$. But if P does not belong to each of the sets $M_{j_1}, M_{j_2}, \cdots$ then there is a positive integer d such that $\bar{R}_{x_d}(P)$ contains no point of M_{x_j} for any j. But $R_{x_d}(P)$ contains $M_{j_{d+1}}$. So P is common to the sets $M_{j_1}, M_{j_2}, \cdots$ and thus common to $M_1, M_2, \cdots$.

Thus, Σ satisfies the fourth part of axiom C.

In order to show that Σ is connected, an indirect argument will be used. Suppose that S is the sum of two mutually separated sets H and K. Since $W + (B_1^* + B_2^* + \cdots)$ is connected, let H' be the one of the sets H and K that contains this set and let K' be the other. There exists an element y of C' such that for some $x\, i_x[I(y)]$ is a subset of K'. But there exists a positive integer d_1 such that for $n \geqq d_1$, $i_n[i_x(I(y))]$, belongs to K'. There exists a positive integer d_2 such that for $n \geqq d_2\, i_n[i_{d_1}(i_x(I(y)))]$ belongs to K'. So, obviously, there is a positive integer sequence, $d_1, d_2, \cdots$ such that if j is a positive integer and $n \geqq d_j$, then $i_n(i_{d_{j-1}}(i_{d_{j-2}}(\cdots i_{d_1}(i_x(I(y))) \cdots)))$ belongs to K'. But from this fact it is easily seen that some point of W is a limit point of K'. So Σ is connected.

It has been shown that in any space satisfying axioms 0 and C (1) if M is a separable point set, M is completely separable, and (2) if M is separable, any subset of M is separable.

In order to show that any two points of S lie in a separable continuum, suppose first that P and Q are two points of S. Obviously, $(B_1^* + B_2^* + \cdots)$ is separable and connected, and therefore $W + (B_1^* + B_2^* + \cdots)$ is a separable continuum. In case P and Q both lie in $W + (B_1^* + B_2^* + \cdots)$, this continuum has the desired properties. In case P does not belong to this set, P belongs to $i_x[I(y)]$ for some y in C'. Let M_P be the set to which point R belongs if and only if, either

(1) there is a finite positive integer sequence $x_1, x_2, \cdots, x_n$ such that R belongs to $i_{x_1}[i_{x_2}[\cdots i_{x_n}[i_x(I(y))] \cdots]]$, or

(2) there is a positive integer q such that R belongs to $i_q[i_x(I(y))]$, or

(3) R belongs to $i_x[I(y)]$. $M_P + (B_1^* + B_2^* + \cdots) + W$ is a separable continuum. If Q does not belongs to this set, let M_Q be a set related to Q like M_P was related to P. The continuum $M_P + M_Q + (B_1^* + B_2^* + \cdots) + W$ is separable.

The statement that Σ is locally separable at the point P means that there is a region R containing P such that R is separable. Alexandroff [1] has shown that if β is a connected, locally completely separable, space satisfying axioms 0 and C, then β is completely separable. It is interesting to note that Σ is locally separable, and therefore locally completely separable, at each point except those of a separable set, and yet, Σ is not separable.

Σ is obviously locally separable at all points not belonging to W. Since every region that contains a point of W contains uncountably many mutually exclusive domains, Σ is not locally separable at any point of W. Furthermore $(B_1^* + B_2^* + \cdots)$ is separable, and so $\overline{(B_1^* + B_2^* + \cdots)}$ is separable, and thus, since W is a subset of the latter, W is separable.

Σ is said to be *semi-locally-connected* [5] at point P if and only if it is true that if R is a region containing P, R contains a region R' containing P such that $S - R$ does not intersect infinitely many components of $S - R'$. Σ is said to be semi-locally-connected if and only if Σ is semi-locally-connected at each point.

The space Σ is obviously semi-locally-connected because S minus any region has only a finite number of components.

PART 2

Suppose that Σ is a space satisfying the conditions specified on the first page of this paper.

For each positive integer j let G_j denote the collection of all open sets which have diameter less that j^{-1}.

Let P denote some definite point, and suppose n is a positive integer such that no countable subcollection of G_n covers S. Let R_n be some region of G_n which contains P, let $H_1 = \{R_n\}$, and let K_1 be the boundary of R_n.

For each point Q of S let $\Delta(Q)$ be the least integer $j > n$ such that some region $R(Q)$ of G_n contains every region of G_j that intersects a region of G_j that contains Q.

It has been shown that in a space satisfying these axioms if L is a separable point set and G is a collection of open sets covering L, then some countable subcollection of G covers L. Therefore, there is a countable point set T_1 dense in K_1 such that the collection H_2 of all $R(Q)$'s, for Q's in T_1, covers K_1. Let K_2 be the sum of the boundaries of all the sets in $H_1 + H_2$. There is a countable point set T_2 dense in K_2 such that the collection H_3 of all $R(Q)$'s for Q's in T_2, covers K_2. Let K_3 be the sum of the boundaries of the sets in $H_1 + H_2 + H_3$, and so forth.

There is a point B not in the closure of $H = (H_1 + H_2 + \cdots)^*$. Let M be a compact continuum containing P and B.

Case 1. Suppose some point A of $M - M \cdot H$ is a limit point of $K = K_1 + K_2 + \cdots$.

Let R_1' be a region of G_n containing A, let Q_1 be a point of $T = T_1 + T_2 + \cdots$ in R_1', and let x_1 be the largest integer i such that R_1' belongs to G_i. Let R_2' be a region of C_{x1+1} containing A such that $\bar{R}_2'$ lies in $R_1' - Q_1$. Let Q_2 be a point of T in R_2' and let x_2 be the largest integer i such that R_2' is in G_i. Obtain R_3', Q_3, and x_3 similarly, and so forth. $n \leqq x_1 < x_2 < x_3 < \cdots$. For each i, $\varDelta(Q_i) > x_i$. Otherwise, for some i, $R(Q_i)$ would contain R_i', and thus A. However, there is a positive integer $t > n$ such that if x, y, and z are regions of G_t such that $x \cdot y$ and $y \cdot z$ exist and x contains A, then R_n contains $x + y + z$. For some $s > t$, $\varDelta(Q_s) > t$. But R_n contains every region of G_t that intersects a region of G_t that contains G_s. So $\varDelta(Q_s) \leqq t$, which is a contradiction.

Case 2. Suppose no point of $M - M \cdot H$ is a limit point of K. For each point Q of $M - M \cdot H$ let g_Q be a region containing Q such that g_Q contains no point of $K + P$. Some finite subcollection C of the g_Q's covers this set of limit points. Let $D = H - H \cdot \bar{C}^*$. Let C_1 be the component of $M - M \cdot \bar{D}$ which contains B. Some point z of $M \cdot \bar{D}$ is a limit point of C_1. But z lies in a region r of H, and therefore C_1 would intersect the boundary of r, and thus contain a limit point of K. This yields a contradiction.

Since, for each n, some countable subcollection of G_n covers S, Σ is completely separable.

References

1. Paul Alexandroff, *Uber die Metrization der im Kleinen Kompakten topologischen Raume,* Mathematische Annalen, **92** (1924) 294–301.

2. F. B. Jones, *On a property related to separability in metric spaces,* Journal of the Elisha Mitchell Scientific Society, **70** (1954), 30–33.

3. ———, *A theorem concerning locally peripherally separable spaces,* Bull. Amer. Math. Soc. **41** (1935), 437–439.

4. R. L. Moore, *Foundations of point set theory,* Amer. Math. Soc. Colloq. Publ., vol 13 1932.

5. G. T. Whyburn, *Semi-locally-connected sets,* Amer. J. Math. **61** (1939), 733–749.

The University of Texas

ON THE COMMUTATIVITY OF A CORRESPONDENCE AND A PERMUTATION

MILO W. WEAVER

Foreword. A permutation is a one-to-one mapping of a finite set *onto* itself. The necessary and sufficient conditions for two permutations S_1 and S_2 to satisfy

$$s_1 s_2 \cong s_2 s_1 \tag{0.1}$$

are known[1]: S_1 and S_2 satisfy (0.1) if and only if S_2 is a product PQ of a permutation P which is a product of powers of cycles of S_1 and a permutation Q which permutes cycles of S_1 with equal numbers of symbols. For example if $S_1 \cong (1\,2\,3\,4)(5\,6\,7\,8)$, $P \cong (1\,3)(2\,4)$, and $Q \cong (1\,5)(2\,6)(3\,7)(4\,8)$, then PQ commutes with S_1. A correspondence is a mapping of a finite set *into* itself. Hence a permutation is a special case of a correspondence. It is our major object in this paper to find the necessary and sufficient conditions for a permutation to commute with a correspondence. These conditions are stated in Theorem 3.15 below.

As the literature[2] has very little on "correspondences," all the fundamental definitions needed in this paper and pertaining to correspondences are given.

It is assumed that the reader knows a little about groups of permutations.

1. Fundamental definitions.[3]

A *correspondence* relates each symbol of a finite set $\mathfrak{N}$ to exactly one symbol of $\mathfrak{N}$. A permutation is a correspondence such that each image symbol is the image of exactly one symbol of $\mathfrak{N}$. The statement, *m is the image of n under the corre-*

Received April 27, 1959. The work on this paper was done under National Science Foundation Grant 8238. The writer wishes to express his appreciation to his 1958 University of Texas class, and particularly to Robert R. Bunten, for suggestions concerning terminology and explanations. He also wishes to thank the referee for a valuable suggestion relating to the definition at the beginning of Section 3.

[1] Burnside, *Theory of Groups of Finite Order*, Cambridge University Press, 1897, pp. 215, 216.

[2] Two papers on correspondences are: R. R. Stoll, "Representations of Finite Simple Semigroups," *Duke Math J.*, vol. 11, no. 2 (1944), 251–265; Milo Weaver, "On the Imbedding of a Finite Commutative Semigroup of Idempotents in a Uniquely Factorable Semigroup, "*Proc. Nat. Acad. Sci.*, vol. 42, no. 10 (1956), 772–775.

[3] Most of the definitions in this section and Theorem 1.5 were given: H. S. Vandiver and M. W. Weaver, " A Development of Associative Algebra and an Algebraic Theory of Numbers, III," *Math. Mag.*, vol. 29 (1956), 135–149.

spondence D is abbreviated $nD = m$.

The notation for a correspondence D:

$$(1.1) \qquad \begin{pmatrix} a_1 a_2 & & a_r \\ & \cdots & \\ b_1 b_2 & & b_r \end{pmatrix}$$

is interpreted: "the a's are distinct symbols of $\mathfrak{N}$ and $a_i D = b_i$, $i = 1, 2, \cdots, r$." If $n \in \mathfrak{N}$ and $nD = n$ and $xD = n$ has no solution x, $x \neq n$, $x \in \mathfrak{N}$, n may be omitted from both lines of (1.1). The single-lined notation for a *cycle* C:

$$(1.2) \qquad (d_1 d_2 \cdots d_s)$$

means that the d's are distinct symbols of $\mathfrak{N}$, $d_i C = d_{i+1}$, $i = 1, 2, \cdots, s-1$, but $d_s C = d_1$; and that $nC = n$ if $n \in \mathfrak{N}$ and n is not one of the d's. If $s = 1$, (1.2) becomes (d_1) and means that this cycle is the *identity* permutation, E, defined by $nE = n$, for each n of $\mathfrak{N}$. The example $\binom{1\,2\,3}{2\,3\,3}$ suggests that some correspondence cannot be described either by (1.2) or by a "product" of cycles. We describe the particular correspondence D' by the notation

$$(1.3) \qquad (d_1 d_2 \cdots d_s\}$$

and interpret this exactly as we did (1.2), except here $s > 1$ and $d_s D' = d_s$. A correspondence of the type (1.3) is called a 1-1-*excycle*, or just a 1-excycle.

The correspondences D_1 and D_2 are said to be *equivalent* if $nD_1 = nD_2$, for each $n \in \mathfrak{N}$. We describe this by $D_1 \cong D_2$.

The *product* $D_3 \cong D_1 D_2$ is defined by $nD_3 = (nD_1)D_2 = nD_1 D_2$ for each $n \in \mathfrak{N}$. We illustrate: if $P \cong \binom{1\,2\,3\,4\,5\,6\,7\,8\,9}{2\,3\,3\,2\,6\,3\,6\,9\,8}$ and $S \cong \binom{1\,4\,5\,7\,2\,6\,8\,9}{5\,7\,4\,1\,6\,2\,9\,8}$ then

$$(1.4) \qquad \begin{aligned} PS \cong SP &\cong (3)\,(1\;2\;3\}\,(4\;2\}\,(5\;6\;3\}\,(7\;6\}\,(8\;9)\cdot(1\;5\;4\;7)\,(8\;9)\,(2\;6) \\ &\cong (3)\,(1\;6\;3\}\,(4\;6\}\,(5\;2\;3\}\,(7\;2\}\;. \end{aligned}$$

Positive integral exponents will be interpreted exactly as in permutation theory. If it is convenient, $m \in \mathfrak{N}$, and A is a correspondence, mA^0 may be used to denote m. Only non-negative exponents will be used for correspondences which are not permutations.

In (1.3) above, the set of d's are elements of a set called $\mathfrak{J}(D')$; d_1 is the only element of a set called $\mathscr{E}(D')$; and d_s is the only element of a set called $\mathfrak{K}(D')$. These sets get their notations, respectively, from the words: *involved*, *end*, and *core*, spelled *k-o-r-e*. We now define these sets, formally.

If D is a correspondence, the set $\mathfrak{J}(D)$ is defined by $i \in \mathfrak{J}(D)$ if and only if $i \in N$ and either $iD \neq i$ or $xD = i$ has a solution x, $x \in \mathfrak{N}$,

$x \neq i$. If $i \in \mathfrak{J}(D)$, we notice that $iD^r \in \mathfrak{J}(D)$ also, for each positive integer r.

The set $\mathscr{E}(D)$ is defined by $j \in \mathscr{E}(D)$ if and only if $xD = j$, $j \in \mathfrak{N}$, has no solution x, $x \in \mathfrak{N}$. Clearly, $jD \neq j$ and $\mathscr{E}(D) \subseteq \mathfrak{J}(D)$.

The set $\mathfrak{K}(D)$ is defined by $k \in \mathfrak{K}(D)$ if and only if $k \in \mathfrak{J}(D)$ and $kD^s = k$ for some non-negative integer s. We note that D acts either as a cycle or as a product of cycles on $\mathfrak{K}(D)$. If $k \in \mathfrak{K}(D)$, $kD^r \in \mathfrak{K}(D)$ also, for each positive integer r. The d's of (1.3) exemplify the fact that it is not necessarily true that $\mathfrak{K}(D) \cup \mathscr{E}(D) = \mathfrak{J}(D)$.

Let D be a correspondence and $k \in \mathfrak{K}(D)$. If each symbol of $\mathfrak{K}(D)$ is one of the symbols $k, kD, kD^2, \cdots$, then D is called an *excycle*. Apparently, if $i \in \mathfrak{J}(D)$, there exists a non-negative integer r such that $iD^r \in \mathfrak{K}(D)$. If D is an excycle and $\mathscr{E}(D)$ and $\mathfrak{K}(D)$ contain exactly r and s symbols, respectively, then D is called an *r-s-excycle*. This explains the term, 1-1-*excycle*. A 0-s-excycle is a cycle with s symbols. The product PS of (1.4) is a 4-1-excycle.

THEOREM 1.5 (known). *Each correspondence is either an excycle or a product of excycles with disjoint* $\mathfrak{J}$*-sets.*

The proof is not given here as it is very similar to that for the well-known theorem: *Each permutation, not a cycle, is a product of cycles with disjoint* $\mathfrak{J}$*-sets.* The excycles (*cycles*) of Theorem 1.5 are called *excycles* (*cycles*) *of the given correspondence.* The excycles of P of (1.4) are $(3)(1\,2\,3\}\,(4\,2\}\,(5\,6\,3\}\,(7\,6\}$ and $(8\,9)$.

If $j \in \mathscr{E}(D)$, clearly, for some u and v, the operation of D on a subset of $\mathfrak{J}(D)$ is described by $D_j \cong (jD^v\, jD^{v+1} \cdots jD^u)\,(j\; jD \cdots jD^v\}$. We call D_j a 1-$(u-v+1)$-*subexcycle of* D determined by j and the first factor of D_j a *subcycle* of D. D_j may also be called simply a 1-subexcycle.

2. Some properties of a correspondence and a permutation which commute.

We next make three simple remarks about commutativity of correspondences. The usual proofs of the corresponding remarks about permutations are valid here.

The identity E commutes with each correspondence.

If L is a correspondence, then $L^aL^b \cong L^bL^a$.

If L and M are correspondences and $\mathfrak{J}(L) \cap \mathfrak{J}(M) = 0$, then $LM \cong ML$.

The relation (1.4) illustrates Theorem 2.1 and Theorem 2.4 below.

THEOREM 2.1. *If S is a permutation* on $\mathfrak{N}$ *and P is a correspondence, not a substitution on* $\mathfrak{N}$ *such that $SP \cong PS$, then S maps* $\mathfrak{J}(P)$ *onto itself* and $\mathscr{E}(P)$ *onto itself.*

Suppose that the hypothesis of the theorem is satisfied and that $n \in \mathfrak{J}(P)$, but that $nS \notin \mathfrak{J}(P)$. Then if $nP = m$, we have

$$nS = nSP = nPS = mS . \tag{2.2}$$

Whence $m = n$. Since $n \in \mathfrak{J}(P)$ and $nP = n$, there exists an a, $a \in \mathfrak{J}(P)$ such that $aP = n \neq a$. And since $nS \notin \mathfrak{J}(P)$, it follows from the equation

$$aSP = aPS = nS \tag{2.3}$$

that $aS = nS$ and $a = n$, a contradiction to $a \neq n$. Hence $nS \in \mathfrak{J}(P)$, and since S is a permutation S maps $\mathfrak{J}(P)$ onto itself. Also if we assume $n \in \mathscr{E}(P)$ and $nP = m$ in (2.2), the conclusion $nP = n$ contradicts the hypothesis, $n \in \mathscr{E}(P)$. Whence S maps $\mathscr{E}(P)$ onto itself.

The following is also a theorem, but we shall not prove it as it is not needed in this paper.

THEOREM 2.4. *If P is a correspondence with $j \in \mathscr{E}(P)$ and P_j is a 1-$(u - v + 1)$-subexcycle of P, determined by j, and if S is a permutation such that $SP \cong PS$ and $jS^bP^m = jP^n$, for $b > 0$, $m \leq u$, $n \leq u$, and either $m < v$ or $n < v$, then $m = n$.*

3. **Products of cycles which permute 1-(u − v + 1)-excycles.** We shall first generalize the idea of a permution permuting cyclically a set of cycles of equal numbers of symbols. Let u, v, and t be any integers such that $u \geq v > 0$ and $t \geq 1$, and F_0, $F_1, \cdots, F_t$ be 1-$(u - v + 1)$-excycles whose $\mathscr{E}$-symobols are, respectively, the distinct symbols, $j_0, j_1, \cdots, j_t$ such that if c is an integer, $0 < c < u$, and d is the least nonnegative residue of the positive integer e, $e \leq t$, modulo $t_c + 1$, with $t_c + 1$, defined below, then

$$j_e F_e^c = j_d F_d^c . \tag{3.1}$$

Let $C_0, C_1, \cdots, C_u$ be cycles of a permutation S such that

$$C_i \cong (j_0 F_0^i \; j_1 F_1^i \cdots j_{t_i} F_{t_i}^i), \tag{3.2}$$

with $t_0 = t$ and the order $t_w + 1$ of C_w dividing that $t_z + 1$ of C_z whenever $0 \leq z \leq w \leq u$. Then S is said to *permute cyclically* the 1-excycles $F_0, F_1, \cdots F_t$.

We give examples here. The permutation (1 4) (2 5) (3 6) permutes cyclically each of the pairs: (2 3) (1 2}, (5 6) (4 5}; (1 2 3}, (4 5 6}; (1 2 3 7}, (4 5 6 7}. Also (1 4 6 7) (2 5) permutes cyclically the set (1 2 3}, (4 5 3}, (6 2 3}, (7 5 3}; and (5 6) (1 3) (2 4) permutes cyclically the set (1 2 3 4) (5 1}, (3 4 1 2) (6 3}. The reader should study each of these examples and refer to them, frequently, while studying the rest of this paper.

We shall use the above terminology for the F's and C's, hereafter.

LEMMA 3.3. *If $C_x \cong E$ and $0 \leq x \leq y \leq u$, then $C_y \cong E$, also.*

This is true, since $t_y + 1$ divides $t_x + 1$.

Let r be the largest integer x such that $x \leq u$ and $C_x \not\cong E$. We impose the added restriction on r, that it be the smallest integer x such that C_i, $i = 0, 1, \cdots, x$ gives all the distinct C_i's.

THEOREM 3.4. *Let P be a correspondence and R be a permutation such that*:

(i) $\mathfrak{J}(P) \supseteq \mathfrak{J}(R)$.

(ii) *R is a product of the distinct cycles of a set of permutations, each of which permutes cyclically 1-subexcycles of P.*

(iii) *If F is a 1-subexcycle of P, then $\mathfrak{J}(F) \cap \mathfrak{J}(R) = 0$, unless F is one of a set permuted cyclically by R.*
Then

$$RP \cong PR . \tag{3.5}$$

If $n \notin \mathfrak{J}(R)$, then neither is nP, by (iii); and

$$nPR = nP = nRP . \tag{3.6}$$

If $n \in \mathfrak{J}(P)$. Let $n \in \mathfrak{J}(\prod_{i=1}^{r} C_i)$, where $\prod_{i=1}^{r} C_i$ permutes cyclically the 1-subexcycles F_l, $l = 0, 1, \cdots, t$; further let $n \in \mathfrak{J}(C_q)$ and $n = j_p F_p^q$, for $0 \leq q \leq r$, and $j_p \in \mathscr{E}(P)$. Then from (3.2), for $0 \leq p \leq t_q$,

$$\begin{aligned} nPR &= nP\Big(\prod_{i=1}^{r} C_i\Big) = j_p F_p^{q+1} C_{q+1} = j_{p+1} F_{p+1}^{q+1} \\ &= j_{p+1} F_{q+1}^{q} F_{p+1} = j_p F_p^q \Big(\prod_{i=1}^{r} C_i\Big) P = n\Big(\prod_{i=1}^{r} C_i\Big) P = nRP ; \end{aligned} \tag{3.7}$$

while if $p = t_q$, both the leftmost and rightmost members of (3.7) yield $j_0 E_0^{q+1}$. Hence, by (3.6) and (3.7), we have (3.5).

THEOREM 3.8. *Let P be a correspondence and S be a permutation such that $SP \cong PS$, with $j \in \mathscr{E}(P)$ and $jP^s \in \mathfrak{J}(S)$ for some non-negative s; further let $t + 1$ be the least positive integer such that $jS^{t+1} = j$ and F_l, $l = 0, 1, \cdots, t$ be the 1-subexcycle of P whose $\mathscr{E}$-symbol is jS^l. Then S permutes the set $F_0, F_1, \cdots, F_t$ cyclically.*

Let g be the largest value, if there is one, of x such that $jP^x \in \mathfrak{J}(S)$, with $0 \leq x \leq u$, $u + 1$ the order of the subexcycle P_j, and $u - v + 1$ the order of its subcycle. Let C_i, $i = 0, 1, \cdots, g$ be the cycle of S, possibly the identity, such that

$$C_i \cong (jP^i \; jP^iS \; jP^iS^2 \cdots jP^iS^{t_i}) \tag{3.9}$$

for some non-negative integer t_i. Certainly $t = t_0$. The order of C_i is

$t_i + 1$. By Theorem 2.1, $jP^iS^l \in \mathfrak{J}(P)$ and $jS^l \in \mathscr{E}(P)$, for $i \leq g$, $0 \leq l \leq t_i$. Since S is a permutation, we have cancellation by S^l and both equations in each of the pairs of equations hold simultaneously:

$$(3.10) \qquad \begin{aligned} jP^{u_0+1} &= j^{v_0}\,, & jS^lP^{u_0+1} &= jS^lP^{v_0} \\ jP^{u_l+1} &= jP^{v_l}\,, & jS^lP^{u_l+1} &= jS^lP^{v_l}\,. \end{aligned}$$

Hence, $u_l = u_0$ and $v_l = v_0$. We notice that for $0 \leq z \leq w \leq g$, and $h + z = w$, we have

$$(3.11) \qquad jP^w = jP^zS^{t_z+1}P^h = jP^{z+h}S^{t_z+1} = jP^wS^{t_z+1}\,.$$

Therefore, since $t_w + 1$ is the order of C_w, it follows from group theory that $t_w + 1$ divides $t_z + 1$. Also if $e \equiv d (\mathrm{mod}\ t_c + 1)$, we have $e = m(t_c + 1) + d$, m a non-negative integer and

$$(3.12) \qquad jS^eF^c_e = jS^{m(t_c+1)+d}P^c = jP^cS^d = jS^dP^c = jS^dF^c_d\,,$$

which gives (3.1), since here $j_e = jS^e$ and $j_d = jS^d$. Hence S permutes the F's cyclically.

Let S be a permutation and P be a correspondence, which is not a permutation. Clearly, P is expressible in the form

$$(3.13) \qquad P \cong T_1T_2\,,$$

where either $T_1 \cong E$ or T_1 is a product of cycles of P, and T_2 is product of those excycles of P which are not cycles. And S is expressible in the form

$$(3.14) \qquad S \cong S_1S_2\,,$$

where S_1 is either a product of those cycles C of S such that $I(C) \cap I(T_2) = 0$ or $S_1 \cong E$, depending on whether or not such C's exist, and S_2 is either a product of those cycles D of S such that $\mathfrak{J}(D) \cap I(T_2) \neq 0$ or $S_2 \cong E$, depending on whether or not such D's exist.

THEOREM 3.15. *If S is a permutation and P is a correspondence, not a permutation, and S_1, S_2, T_1, and T_2 satisfy* (3.13) *and* (3.14), *then $SP \cong PS$ if and only if*:

(*i*) $S_1T_1 \cong T_1S_1$;

(*ii*) *Whenever $j \in \mathscr{E}(T_2)$ such that for some non-negative integer s, $jP^s \in \mathfrak{J}(S)$, then S_2 permutes cyclically the set of* 1-*excycles of T_2 whose $\mathscr{E}$-symbols are the distinct symbols obtained by applying all powers of S to j.*

Suppose that $PS \cong SP$. By Theorem 3.8, if $j \in \mathscr{E}(P)$ and $jP^s \in \mathfrak{J}(S)$, a product π defined as in (3.2) of cycles of S permutes cyclically a set of 1-subexcycles of P, and therefore of T_2, having powers of S applied to j as their $\mathscr{E}$-symbols. By (3.2) $I(\pi)$ is contained in the union of the $\mathfrak{J}$-sets of the subexcycles which it permutes. Clearly, S_2 is a product of the distinct cycles of all such π's, or $S_2 \cong E$, depending on whether or not such π's exist, and S_2 satisfies (ii) of Theorem 3.15. From (3.13) and (3.14)

$$\mathfrak{J}(T_1) \cap \mathfrak{J}(T_2) = \mathfrak{J}(S_1) \cap \mathfrak{J}(T_2) = 0 . \tag{3.16}$$

Since $\mathfrak{J}(S_2) \subseteq \mathfrak{J}(T_2)$, we have

$$\mathfrak{J}(T_1) \cap \mathfrak{J}(S_2) = 0 . \tag{3.17}$$

Hence $\mathfrak{J}(T_1) \cup \mathfrak{J}(S_1) \cap \mathfrak{J}(T_2) \cup \mathfrak{J}(S_2) = 0$, and S_1T_1 and T_1S_1 operate on $\mathfrak{J}(S_1) \cup \mathfrak{J}(T_1)$ exactly as ST and TS do; and for $n \notin \mathfrak{J}(T_1) \cup \mathfrak{J}(S_1)$, $nS_1T_1 = nT_1S_1 = n$. Whence $S_1T_1 \cong T_1S_1$, and (i) is satisfied. Now assume that (i) and (ii) of Theorem 3.15 are satisfied by S and P. From Theorem 3.4, we have $S_2T_2 \cong T_2S_2$. By (i), $S_1T_1 \cong T_1S_1$. From (3.16) and (3.17), $S_1T_2 \cong T_2S_1$, $T_1T_2 \cong T_2T_1$, and $S_2T_1 \cong T_1S_2$. Hence

$$SP \cong S_1S_2T_1T_2 \cong S_1T_1S_2T_2 \cong T_1S_1T_2S_2 \cong T_1T_2S_1S_2 \cong PS . \tag{3.18}$$

This completes the proof of Theorem 3.15 which was the major objective of this paper.

The necessary and sufficient conditions for (i) to hold were stated in the foreword. In each of the examples below (3.2), if S is taken to be the permutation and P to be the correspondence whose 1-subexcycles are permuted by S, then S and P obey (i) and (ii) of Theorem 3.15. A more complicated example of such a P and S is: $P \cong (4)\,(1\;2\;3\}$ $(2\;3\}\,(8)\,(5\;7\;8\}\,(6\;7\}$, $S \cong (1\;5)\,(2\;6)\,(3\;7)\,(4\;8)$. On the other hand if $S \cong (1\;4\;6)\,(2\;5)$ and $P \cong (1\;2\;3\}\,(4\;5\;3\}\,(6\;5\}$, then $SP \not\cong PS$, since the order of $(2\;5)$ fails to divide that of $(1\;4\;6)$ and S_2 fails to permute cyclically the 1-1-subexcycles $(1\;2\;3\}$, $(4\;5\;3\}$, and $(6\;5\;3\}$ of P.

ON THE ZEROS OF SOLUTIONS OF SOME LINEAR COMPLEX DIFFERENTIAL EQUATIONS

DAVID V. V. WEND

Introduction. In this paper Green's function methods are used to investigate the distribution on the real axis of zeros of solutions of the complex differential equations

$$(p(x)y')' + f(x)y = 0 \tag{1}$$

and

$$y''' + f(x)y = 0 . \tag{2}$$

In both cases the coefficient $f(x)$ is assumed to be complex-valued and continuous on a half-line $I: x_0 \leq x < \infty$, while $p(x)$ in equation (1) is assumed to belong to a special class of complex-valued functions to be defined in Section I.

Equation (1) or equation (2) is said to be *nonoscillatory* on a set E if no nontrivial solution has an infinite number of zeros in E. In what follows a solution shall mean a nontrivial solution. Suppose in equation (1) x is a complex variable and $p(x)$ and $f(x)$ are analytic in a simply-connected region R. Consider the well known Green's function $g(x, s)$ for the differential system

$$(p(x)y')' = 0, \quad y(a) = y(b) = 0 , \tag{3}$$

where a and b are distinct points of R[1]. If a and b are zeros of a solution of equation (1), then

$$1 \leq \int_a^b |g(x, s)|\,|f(x)|\,|dx| ,$$

where the integral is taken along a path C in R and s is an interior point of C. Starting with this inequality and imposing various bounds on $|f(x)|$, Z. Nehari [7] and P. R. Beesack [3] have obtained nonoscillation theorems for $y'' + f(x)y = 0$ in various regions of the complex plane where $f(x)$ is analytic. By the same methods the author [2] has extended some of these theorems and obtained similar results for equation (1). The methods used in this paper are essentially those employed in the sources mentioned above. However, by restricting the independent variable to the real axis the condition of analyticity is relaxed and

Received May 21, 1959. This research was sponsored by the Office of Ordnance Research under Contract DA-04-495-ORD-1088. Presented to the American Mathematical Society November, 1958.

[1] Sufficient conditions for the existence of $g(x, s)$ are given in [**2**, p 15].

upper bounds on the number of zeros of a solution on a given interval are obtained not only for equation (1) but also for the third order equation (2).

1. A nonoscillation theorem. In this section we will consider equation (1). It will be assumed that $p(x)$ is continuous and different from zero on I. In order to make use of Green's function we wish to have the system (3) incompatible, i.e., possess no (nontrivial) solution on I.

If $p(x)$ is allowed to be complex-valued on I, then the system (3) may be compatible. For example the system $(e^{-ix}y')' = 0$, $y(2m\pi) = 0$, $y(2n\pi) = 0$, m and n distinct positive integers, has the nontrivial solution $y(x) = e^{ix} - 1$ on $I: 0 \leq x < \infty$. In order to avoid such examples and also to be able to make use of certain estimates of Green's function, only a restricted class of functions $p(x)$ will be considered.

DEFINITION. Let $G(I)$ denote the class of all complex-valued, continuous and non-zero functions $p(x)$ defined on $I: x_0 \leq x < \infty$ which possess the further property that for any three numbers a, b and c such that $x_0 \leq a < b < c < \infty$,

$$(4) \qquad \begin{cases} \text{(a)} & \left|\int_a^b \frac{dx}{p(x)}\right| < \left|\int_a^c \frac{dx}{p(x)}\right| \\ \text{(b)} & \left|\int_b^c \frac{dx}{p(x)}\right| < \left|\int_a^c \frac{dx}{p(x)}\right| \end{cases} .$$

Note. The class $G(I)$ contains the functions $p(x) > 0$ which are continuous on I.

An interesting subclass of $G(I)$ is the collection of complex-valued functions $p(x)$ in $G(I)$ which possess the additional property that if

$$\varphi(x) = \int_{x_0}^x \frac{dt}{p(t)} = u(x) + iv(x) ,$$

then

$$\left(\frac{du}{dx}\right)^2 + \left(\frac{dv}{dx}\right)^2 \neq 0 \text{ on } I$$

and for any x', x'' in I, $\theta' = \arctan\,(dv/dx/du/dx)\,|_{x=x'}$ and

$$\theta'' = \arctan\left(\frac{dv/dx}{du/dx}\right)\bigg|_{x=x''}$$

can be chosen so that $|\,\theta' - \theta''\,| < \pi/2$. In effect, the image curve of I under $\varphi(x)$ cannot change direction by more than $\pi/2$.

Suppose $p(x) \in G(I)$. Then the differential system (3) is incompatible. Therefore the Green's function for this system exists, and it has the explicit form

$$(5) \qquad g(x, s) = \begin{cases} \displaystyle\int_a^x \frac{dt}{p(t)} \int_s^b \frac{dt}{p(t)} \Big/ \int_a^b \frac{dt}{p(t)}, & a \leq x \leq s \\ \displaystyle\int_a^s \frac{dt}{p(t)} \int_x^b \frac{dt}{p(t)} \Big/ \int_a^b \frac{dt}{p(t)}, & s \leq x \leq b \end{cases},$$

$a < s < b$.

Since $p(x) \in G(I)$, the inequalities (4) are satisfied and these inequalities together with the above expressions for $g(x, s)$ show that

$$(6) \qquad |g(x, s)| < \begin{cases} \text{(a)} & \displaystyle\int_a^b \frac{dt}{|p(t)|} \\ \text{(b)} & \displaystyle\int_a^x \frac{dt}{|p(t)|} \\ \text{(c)} & \displaystyle\int_x^b \frac{dt}{|p(t)|} \end{cases},$$

$x \neq a, b$ and $s \neq a, b$.

If $y(x)$ is a nontrivial solution of equation (1) on the interval $a \leq x \leq b$ such that $y(a) = y(b) = 0$, then the inequalities (4) imply $f(x)$ is not identically zero on $a \leq x \leq b$ and

$$y(x) = \int_a^b g(x, s) y(s) f(s) ds .$$

If x is chosen so that $|y(x)|$ is a maximum on the interval $a \leq x \leq b$, then (following Z. Nehari [7]):

$$(7) \qquad 1 < \int_a^b |g(x, s)| \, |f(s)| \, ds .$$

As a consequence of inequalities (6) and (7),

$$(8) \qquad 1 < \begin{cases} \text{(a)} & \displaystyle\int_a^b |f(x)| \, dx \int_a^b \frac{dx}{|p(x)|} \\ \text{(b)} & \displaystyle\int_a^b |f(x)| \left(\int_a^x \frac{dt}{|p(t)|} \right) dx \\ \text{(c)} & \displaystyle\int_a^b |f(x)| \left(\int_x^b \frac{dt}{|p(t)|} \right) dx \end{cases}.$$

THEOREM 1. *Suppose* $p(x) \in G(I)$, *and* $a_1 < a_2 \cdots < a_n$ *are* n *consecutive zeros of a solution of* $(p(x)y')' + f(x)y = 0$, $a_1 \geq x_0$. *Then* a_n

must satisfy the inequalities

$$(9)\qquad \begin{cases} \text{(a)}\quad n-1 < \int_{x_0}^{a_n} |f(x)|\,dx \int_{x_0}^{a_n} \frac{dx}{|p(x)|} \\ \text{(b)}\quad n-1 < \int_{x_0}^{a_n} |f(x)| \left(\int_{x_0}^{x} \frac{dt}{|p(t)|}\right) dx\,. \\ \text{(c)}\quad n-1 < \int_{x_0}^{a_n} |f(x)| \left(\int_{x}^{a_n} \frac{dt}{|p(t)|}\right) dx \end{cases}$$

Proof. Since $p(x) \in G(I)$, the inequalities (8) are satisfied for a and b zeros of a solution of equation (1). From inequality (8a)

$$1 < \int_{a_j}^{a_{j+1}} |f(x)| \int_{a_j}^{a_{j+1}} \frac{dx}{|p(x)|}\,, \qquad j = 1, 2, \cdots, n-1\,.$$

Adding these $n-1$ inequalities,

$$n-1 < \sum_{j=1}^{n-1} \int_{a_j}^{a_{j+1}} |f(x)|\,dx \int_{a_j}^{a_{j+1}} \frac{dx}{|p(x)|} \le \int_{x_0}^{a_n} |f(x)|\,dx \int_{x_0}^{a_n} \frac{dx}{|p(x)|}\,,$$

giving the inequality (9a). The inequalities (9b) and (9c) follow in similar fashion from the inequalities (8b) and (8c), respectively.

The following theorem is an immediate corollary of Theorem 1.

THEOREM 2. *Nonoscillation theorem. Suppose* $p(x) \in G(I)$ *and*

$$L = \int_{x_0}^{\infty} |f(x)| \left(\int_{x_0}^{x} \frac{dt}{|p(t)|}\right) dx\,,$$

$$M = \int_{x_0}^{\infty} |f(x)| \left(\int_{x}^{\infty} \frac{dt}{|p(t)|}\right) dx\,,$$

where L and M many assume the value $+\infty$. Then $(p(x)y')' + f(x)y = 0$ is nonoscillatory on I if either L or M is finite, and if either L or M is less than 1, then equation (1) is *disconjugate* on I, i.e., no solution has more than one zero on I.

In the case $f(x)$ and $p(x)$ are real, the tests in Theorem 2 compare with known criteria, for example those of W. Leighton [5, Corollary 4.2], E. Hille [4, p. 238], R. A. Moore [6, Theorems 3, 4 and 7 Corollary 1] and R. L. Potter [8, Theorem 4.2].

2. An example. The substitution $y = v/\sqrt{p}$ transforms equation (1) into the normal form

$$(10)\qquad v'' + F(x)v = 0$$

where

$$F(x) = \frac{f}{p} + \frac{1}{4}\left(\frac{p'}{p}\right)^2 - \frac{1}{2}\left(\frac{p''}{p}\right),$$

and equation (1) is nonoscillatory if and only if equation (10) is nonoscillatory. With $p(x) \equiv 1$, the constant M in Theorem 2 is infinite, while the nonoscillation condition

$$L = \int_{x_0}^{\infty} |f(x)| \left(\int_{x_0}^{x} \frac{dt}{|p(t)|}\right) dx < \infty . \tag{11}$$

is equivalent to

$$\int_{x_0}^{\infty} x\,|F(x)|\,dx < \infty . \tag{12}$$

In the following differential equation the integral in (12) is infinite, hence fails to show nonoscillation, while in (11) $L < 2$, showing that no solution of the equation can have more than two zeros on I. Let

$$\left(\frac{x^2}{2 - \sin\log x}\, y'\right)' + \frac{1}{(x+i)^2}\, y = 0,\ x_0 = 1 . \tag{13}$$

Since $p(x) = x^2/(2 - \sin\log x) > 0$ on I, $p(x) \in G(I)$, and it is easily estimated that

$$L = \int_1^{\infty} \frac{1}{|x+i|^2} \left(\int_1^x \frac{2 - \sin\log t}{t^2}\, dt\right) dx < \frac{3}{2} .$$

For equation (13)

$$F(x) = \frac{2 - \sin\log x}{x^2(x+i)^2} + \frac{1}{4x^2}\left(\frac{-3\cos^2\log x}{(2 - \sin\log x)^2} + \frac{2\sin\log x - 2\cos\log x}{2 - \sin\log x}\right),$$

and easy estimations give

$$\int_1^{\infty} x\,|F(x)|\,dx \geq \int_1^{\infty} \frac{1}{4x} \left| \frac{2\sin\log x - 2\cos\log x}{2 - \sin\log x} - \frac{3\cos^2\log x}{(2 - \sin\log x)^2} \right| dx$$
$$- \int_1^{\infty} \frac{2 - \sin\log x}{x(x^2+1)}\, dx = I_1 - I_2 ,$$

where $0 < I_2 < 3/4$. Letting $t = \log x$ in I_1,

$$I_1 \geq \frac{1}{36} \int_0^{\infty} |1 + (\cos t - \sin t)^2 + \cos^2 t - 4(\sin t - \cos t)|\, dt$$
$$= \frac{1}{36} \int_0^{\infty} |k(t)|\, dt .$$

From the graph of $k(t)$, $k(t) > 1$ for $0 < t < \pi/4$ and $5\pi/4 < t < 2\pi$, while $k(t) < -1$ for $3\pi/4 < t < \pi$, so $|k(t)| > 1$ for intervals of length $5\pi/4$ out of each interval of length 2π on $0 \leq t < \infty$. Therefore $I_1 = \infty$, so $\int_1^\infty x\,|F(x)|\,dx = \infty$.

3. Distribution of zeros. Suppose the upper limits of the integrals on the right in the inequalities (9) are considered as continuous variables and $f(x)$ is not identically zero on any subinterval of I. Then in each case the integral is a strictly monotone increasing function of the upper limit and there exists at most one value of the upper limit for which equality will hold. If x_1 is such a value, then no solution of equation (1) can have more than n zeros on the interval $x_0 \leq x \leq x_1$. Since $a_n \geq x_1$, the value x_1 also gives a lower bound on the magnitude of the nth consecutive zero on I of any solution of equation (1).

Adapting the notation used in [**6**], let $N(x_1, x_2)$ be the maximum number of zeros any solution of equation (1) may have on the interval $x_1 \leq x \leq x_2$. Since in the complex case there is often no zero separation theorem, the number $N(x_1, x_2)$ merely puts an upper bound on the number of zeros a particular solution may have. See [**1**, Theorem 1.2].

As an application of the above discussion we give the following theorem:

THEOREM 3. *Suppose* $a_1 < a_2 < \cdots < a_n$, $1 \leq x_0 \leq a_1$, *are* n *consecutive zeros of a solution of*

$$(x^\sigma y')' + f(x)y = 0$$

and $H = \int_{x_0}^\infty |f(x)|\,dx < \infty$. *If* $0 \leq \sigma < 1$, *then*

$$[1 + (n-1)(1-\sigma)/H]^{1/(1-\sigma)} = x_1 < a_n \tag{14}$$

and $N(x_0, x_1) < n$. *If* $\sigma = 1$, *then*

$$\exp\left[\frac{n-1}{H}\right] = x_2 < a_n \tag{15}$$

and $N(x_0, x_2) < n$.

Proof. Inequality (14) follows from inequality (9a). Inequality (15) may be obtained from inequality (14) by letting $\sigma \to 1$ or directly from inequality (9b).

Other lower bounds on the magnitude of the zeros of solutions of equation (1) can be obtained by considering the maximum value of $|g(x, s)|$ on $a \leq x \leq b$, $a < s < b$. We assume $p(x) > 0$ and continuous on I. From the expressions for $g(x, s)$ given in (5) it can be shown that the maximum value of $|g(x, s)|$ occurs when $x = s$ and s satisfies the equation

$$\int_a^s \frac{dx}{p(x)} = \int_s^b \frac{dx}{p(x)} . \tag{16}$$

(Compare [3, p. 231].)

As an illustration of this result we give the following theorem:

THEOREM 4. *Suppose* $a_1 < a_2 < \cdots < a_n$, $0 \leq x_0 \leq a_1$, *are n consecutive zeros of a solution of equation* (1) *and* $H = \int_{x_0}^{\infty} |f(x)|\, dx < \infty$. *If* $p(x) \equiv 1$ *on* I, *then* $N(x_0, 4(n-1)/H) < n$. *If* $p(x) \equiv x$ *on* I, *then* $N(x_0, x_0 \exp[4(n-1)/H]) < n$. *If* $p(x) \equiv x^2$ *on* I, *then* $N(x_0, \infty) < (H/4x_0) + 1$, $x_0 > 0$, *hence the equation is nonoscillatory on* I.

Proof. If $p(x) \equiv 1$ on I, then from equation (16) $s = (a+b)/2$ and the maximum value of $|g(x, s)| = (b-a)/4$. From inequality (7),

$$1 < \max |g(x, s)| \int_{a_j}^{a_{j+1}} |f(x)|\, dx , \qquad j = 1, 2, \cdots, n-1$$

so that

$$n - 1 < \frac{a_n}{4} \int_{x_0}^{\infty} |f(x)|\, dx = \frac{Ha_n}{4} ,$$

and $a_n > 4(n-1)/H$. The results for $p(x) \equiv x$ and $p(x) \equiv x^2$ can be obtained in a similar fashion.

4. The equation $y''' + f(x)y = 0$. The differential system $y''' = 0$, $y(a) = y(b) = y(c) = 0$, $a < b < c$, is incompatible, so that the Green's function for this system exists and has the explicit form

$$g(x, s) = \begin{cases} \dfrac{1}{2}\dfrac{(c-s)^2}{(c-a)(c-b)}(x-a)(x-b) = g_{11},\ b < s < c,\ a \leq x \leq s \\ g_{11} - \dfrac{1}{2}(x-s)^2 = g_{12},\ b < s < c,\ s \leq x \leq c \\ g_{11} - \dfrac{1}{2}\dfrac{(b-s)^2}{(c-b)(b-a)}(x-a)(x-c) = g_{21},\ a < s < b,\ a \leq x \leq s \\ g_{21} - \dfrac{1}{2}(x-s)^2 = g_{22},\ a < s < b,\ s \leq x \leq c \\ \dfrac{1}{2}\dfrac{c-b}{c-a}(x-a)(x-b) = g_{31},\ s = b,\ a \leq x \leq s \\ g_{31} - \dfrac{1}{2}(x-s)^2 = g_{32},\ s = b,\ s \leq x \leq c , \end{cases} \tag{17}$$

An upper bound for $|g(x, s)|$ on $a \leq x \leq c, a < s < c$ can be obtained when $a \geq 0$ by considering each of the expressions g_{ij} above. It is easily found that

$$|g_{11}| < \frac{c^2}{2}, |g_{12}| < c^2, |g_{31}| < \frac{c^2}{2}, |g_{32}| < c^2 .$$

The expression for $g_{22}(x, s)$ can be written as

$$g_{22} = \frac{1}{2} \frac{(s-a)^2}{(c-a)(b-a)} (x-b)(c-x) ,$$

whence $|g_{22}| < c^2/2$, and $|g_{21}| \leq |g_{22}| + (1/2)(x-s)^2 < c^2$. Thus in each case $|g_{ij}| < c^2$, so

$$|g(x, s)| < c^2 \text{ for } a \leq x \leq c, a < s < c . \tag{18}$$

Assume $f(x)$ is continuous on I. If $y(x)$ is a nontrivial solution of equation (2) on the interval $a \leq x \leq c$ for which $y(a) = y(b) = y(c) = 0$, $0 \leq x_0 \leq a < b < c$, then

$$y(x) = \int_a^c g(x, s)y(s)f(s)ds ,$$

and as in § 1,

$$1 < \int_a^c |g(x, s)| |f(s)| ds .$$

Using inequality (18),

$$1 < c^2 \int_a^c |f(x)| dx . \tag{19}$$

THEOREM 5. *Suppose $f(x)$ is continuous on I: $x_0 \leq x < \infty, x_0 \geq 0$, and $\int_{x_0}^{\infty} |f(x)| dx = N$. If $a_1 < a_2 < \cdots < a_n$ are n consecutive zero of a solution of $y'' + f(x)y = 0, a_1 \geq x_0$, then*

$$a_n > \sqrt{[(n-1) - (1 + (-1)^n)/2]/2N}, n \geq 3 . \tag{20}$$

Proof. From inequality (19),

$$1 < a_{j+2}^2 \int_{a_j}^{a_{j+2}} |f(x)| dx, \quad j = 1, 2, \cdots, n-2 . \tag{21}$$

Let $n = 2m$. Then adding the inequalities in (21) for $j = 2, 4, \cdots, 2m-2$,

$$m - 1 < a_{2m}^2 \int_{a_2}^{a_{2m}} |f(x)| dx < a_{2m}^2 N .$$

Therefore $a_{2m} = a_n > \sqrt{(n-2)/2N}$. If $n = 2m + 1$, then adding inequalities in (21) for $j = 1, 3, \cdots, 2m-1$, $m < a_{2m+1}^2 N$, so $a_{2m+1} = a_n > \sqrt{(n-1)/2N}$. Combining these two cases the inequality (20) results.

Note 2. Adding the $n-2$ inequalities in (21),

$$n - 2 < 2a_n^2 \int_{x_0}^{a_n} |f(x)|\, dx \, .$$

In the case when $N = \int_{x_0}^{\infty} |f(x)|\, dx = \infty$, this last inequality still yields lower bounds for the zeros a_n. For example, if $f(x) = \sqrt{x} + i$ and $x_0 = 0$, then

$$n - 2 < \frac{4}{3} a_n^2 [(a_n + 1)^{3/2} - 1], \; n \geq 3 \, ,$$

and the positive root of

$$x^7 + 3x^6 + 3x^5 - \frac{3}{2}(n-2)x^2 - \frac{9}{16}(n-2)^2 = 0$$

is a lower bound for a_n.

5. **Higher order equations.** The methods employed in deriving inequalities (14) ($\sigma = 0$) and (20) can be applied to the kth order differential equation

(22) $$y^{(k)} + f(x)y = 0 \, ,$$

where $f(x)$ is continuous and complex-valued on I. For suppose $a_1 < a_2 < \cdots < a_n$ are n consecutive zeros of a solution of equation (22), $a_1 \geq x_0 \geq 0$, $n = kq + r \geq k$. Then

$$1 < \int_{a_j}^{a_{j+k-1}} |g(x, s)|\, |f(x)|\, dx, \quad j = 1, 2, \cdots, n - k + 1 \, ,$$

where $g(x, s)$ is the Green's function for the system.

$$y^{(k)} = 0, \; y(a_j) = y(a_{j+1}) = \cdots = y(a_{j+k-1}) = 0 \, .$$

Suppose a bound can be found for $|g(x, s)|$ on the interval $a_j \leq x \leq a_{j+k-1}$ which is a monotone function, say $B(a_{j+k-1})$, of a_{j+k-1}. Then

$$q < B(a_n) \int_{x_0}^{a_n} |f(x)|\, dx \leq B(a_n) N \, ,$$

where $N = \int_{x_0}^{\infty} |f(x)|\, dx \leq \infty$. In particular if we conjecture[2] $B(a_n) < a_n^{k-1}$, as is the case for $k = 2, 3$, then

$$a_n > \sqrt[k-1]{\frac{n-(k-1)}{kN}} \, .$$

[2] This conjecture has been verified for $n < 6$.

REFERENCES

1. J. H. Barrett, *Second order complex differential equations with a real independent variable,* Pacific J. Math. **8** (1958), 187–200.

2. J. H. Barrett, W. J. Coles, D. V. V. Wend and R. W. Hunt, *Complex differential equations I,* O. O. R. Technical Report, University of Utah (1958), 1–39.

3. P. R. Beesack, *Nonoscillation and disconjugacy in the complex domain,* Trans. Amer. Math. Soc. **81** (1956), 211–242.

4. E. Hille, *Non-oscillation theorems,* Trans. Amer. Math. Soc. **64** (1948), 234–252.

5. W. Leighton, *On self-adjoint differential equations of second order,* J. Lond. Math. Soc. **27** (1952), 37–47.

6. R. A. Moore, *The behavior of solutions of a linear differential equation of the second order,* Pacific J. Math. **5** (1955), 125–145.

7. Z. Nehari, *On the zeros of solutions of second-order linear differential equations,* Amer. J. Math. **76** (1954), 689–697.

8. R. L. Potter, *On self-adjoint differential equations of second order,* Pacific J. Math. **3** (1953), 467–491.

UNIVERSITY OF UTAH

POLARITY AND DUALITY

FRED B. WRIGHT

One of the most frequently encountered situations in mathematics is the existence of a Galois correspondence between two partially ordered systems. An abstract formulation of this concept has been given by Garrett Birkhoff [1, pp. 54-57] and Ore [5], in the following terms.

DEFINITION. Let A and B be two partially ordered sets, and let $\sharp : A \to B$ and $+ : A \to B$ be two mappings such that:

(i) if $p_1 \leq p_2$ in A, then $p_2^\sharp \leq p_1^\sharp$ in B;

(ii) if $q_1 \leq q_2$ in B, then $q_2^+ \leq q_1^+$ in A; and

(iii) for any $p \in A$ and any $q \in B$, $p \leq q^{\sharp +}$ and $q \leq q^{+\sharp}$.

Then the mappings $\sharp$ and $+$ are said to define a Galois correspondence between A and B.

The number of ways in which a Galois correspondence can arise is quite large, and most of them are very well known instances of what is usually called "duality theory". Perhaps the commonest source is the existence of a relation between the elements of two sets. Birkhoff has described this procedure as follows. Let S and T be two sets, and let ρ be a relation from S to T. That is, ρ is a subset of the cartesian product $S \times T$; we write $s\rho t$ to denote $(s, t) \in \rho$, as is customary. For any subset $S_1 \subset S$, define $S_1^\sharp$ to be the set of all those elements $t \in T$ such that $s_1 \rho t$ for all $s_1 \in S_1$. Similarly, for any subset $T_1 \subset T$, denote by T_1^+ the set of all those $s \in S$ such that $s \rho t_1$ for all $t_1 \in T_1$. Then the mapping $\sharp : A \to B$ and $+ : B \to A$ define a Galois correspondence between the Boolean algebra A of all subsets of S and the Boolean algebra B of all subsets of T.

This example has some special features which are not available for general partially ordered systems. If ϕ denotes the empty set of S, then $\phi^\sharp = T$, and if S_1 and S_2 are any two subsets of S, then $(S_1 \cup S_2)^\sharp = S_1^\sharp \cup S_2^\sharp$. A similar result holds for the other mapping $+$. This is due to the fact that Boolean algebras are special cases of lattices which satisfy the conditions of the following result.

LEMMA. *Let A and B be lattices, each having a greatest element* 1 *and a least element* 0, *and let $\sharp : A \to B$ and $+ : B \to A$ define a Galois correspondence between A and B. Then* $0^\sharp = 1$, $0^+ = 1$, *and* $(p_1 \vee p_2)^\sharp = p_1^\sharp \wedge p_2^\sharp$, $(q_1 \vee q_2)^+ = q_1^+ \wedge q_2^+$, *for any* $p_1, p_2 \in A$ *and* $q_1, q_2 \in B$.

This result is well known and, in any event, easily proved. (The

Received May 8, 1959. The author is an Alfred P. Sloan Research Fellow.

symbols 0, 1 denote ambiguously the least and greatest elements of both A and B.)

This suggests that these two conditons might perhaps be taken as more primitive embodiments of general duality concepts. In so doing, of course, one loses the full generality of partially ordered systems. The purpose of this note is to consider mappings of Boolean algebras which have these two properties. It will be shown that, in this case, the method of Birkhoff described above is not only sufficient for constructing a Galois corresponcence but is also necessary.

To be precise, we introduce the following terminolgy.

DEFINITION. Let A and B be two Boolean algebras. By a polarity of A into B, we shall mean a mapping $\#$ of A into B satisfying the two requirements: (i) $0^\# = 1$, and (ii) for any $p, q \in A$, $(p \vee q)^\# = p^\# \wedge q^\#$.

Some recent developments in the duality theory of Boolean algebras may be used to characterize completely such mappings. It may be well to summarize these developments.

If A is any Boolean algebra, its dual space X is a Boolean space—that is, a compact, totally disconnected, Hausdorff space. The algebra A is isomorphic with the Boolean algebra $\Phi(X)$ of all continuous functions from the space X to the (discrete) two-element Boolean algebra Φ. The algebra A will, in fact, be identified with the algebra $\Phi(X)$, so that each element $p \in A$ is a continuous function from X to Φ, and each such function is an element of A. This relationship between Boolean algebras A and Boolean spaces X is the basis of the duality theory of M. H. Stone [6, 7].

Let A and B be two Boolean algebras, with dual spaces X and Y respectively, so that $A = \Phi(X)$ and $B = \Phi(Y)$. By a hemimorphism α of A into B is meant a mapping $\alpha : AB$ such that (i) $\alpha 0 = 0$, and (ii) $\alpha(p \vee q) = \alpha p \vee \alpha q$, for any $p, q \in A$. Every hemimorphism α of A into B defines a relation, denoted by α^*, of Y into X, as follows: $y\alpha^*x$ if and only if $p(x) \leq \alpha p(y)$ for every $p \in A$. The relation α^* so defined has two special topological properties. If E is any subset of X, let $\alpha^{*-1}E$ denote the set of all those $y \in Y$ such that $y\alpha^*x$ for some $x \in E$. Then α^* has the property that $\alpha^{*-1}P$ is a clopen set in Y whenever P is a clopen set in X. (A clopen set in a topological space is one which is both closed and open.) Another way of expressing this property is to say that $\alpha^{*-1}\{x \in X : p(x) = 1\} = \{y \in Y : \alpha p(y) = 1\}$, for each $p \in A$. Moreover, if $y \in Y$ is any point, then the set of all those $x \in X$ such that $y\alpha^*x$ is compact.

Conversely, let ρ be any relation from the space Y to the space X. For any element $p \in A$, we define a function ρ^*p from Y to Φ by setting $\rho^*p(y) = \text{lub}\,\{p(x) : y\rho x\}$. It is easily seen that $\rho^*0(y) = 0$ for every $y \in Y$ and that $\rho^*(p \vee q)(y) = \rho^*p(y) \vee \rho^*q(y)$ for any $y \in Y$ and any two elements $p, q \in A$. If ρ has the property that $\rho^{-1}P$ is a clopen set in Y whenever P is a clopen set in X, then $\rho^{*}p$ will be a continuous function for each element $p \in A$, and hence ρ^{*} will be a hemimorphism of A into B.

If α is a hemimorphism of A into B, and if α^{*} is the relation from Y to X described above, then α^{**} is also a hemimorphism of A into B. One easily shows that $\alpha^{**} = \alpha$. (Any mapping α has a dual relation α^{*}, defined as before; in order that $\alpha = \alpha^{**}$, it is necessary and sufficient that α be a hemimorphism.) On the other hand, suppose that ρ is a relation from Y to X such that $\rho^{-1}P$ is clopen in Y whenever P is clopen in X; then ρ^{*} is a hemimorphism of A into B, and hence ρ^{**} is a relation from Y to X. A necessary and sufficient condition that $\rho = \rho^{**}$ is that for each $y \in Y$, the set $\{x \in X : y\rho x\}$ is compact. Such a relation is called a Boolean relation. The correspondence between hemimorphisms and Boolean relations just described is one-to-one. This extension of Stone's duality theory is due to Halmos [2]. See also Jonsson and Tarski [4] and Wright [8].

Cognizance should be taken of the fact that topological considerations may be ignored when the algebras A and B are the algebras of *all* subsets of two sets, say of S and T, respectively. If $A = \Phi(X)$ and $B = \Phi(Y)$, then the Boolean spaces X and Y are the Stone-Čech compactifications of the discrete spaces S and T respectively. Then a Boolean relation from Y to X defines a relation from T to S, and any relation from T to S may be extended to a Boolean relation from Y to X.

The duality between hemimorphisms and Boolean relations is sufficient to describe completely the structure of polarities, because the theory of polarities is coextensive with the theory of hemimorphisms. (If p is an element of a Boolean algebra, we denote the complement of p by the symbol p'.)

THEOREM 1. *If $\sharp$ is a polarity of a Boolean algebra A into a Boolean algebra B, and if, for each $p \in A$, we set $\alpha p = (p^{\sharp})'$, then α is a hemimorphism of A into B. Conversely, if α is a hemimorphism of A into B, and if, for each $p \in A$, we set $p^{\sharp} = (\alpha p)'$, then $\sharp$ is a polarity of A into B.*

Proof. This is quite trivial: let α and $\sharp$ be two mappings of A into B such that $(\alpha p)' = p^{\sharp}$ for each $p \in A$. Then $\alpha 0 = 0$ if and only if $0^{\sharp} = 1$, and $\alpha(p \vee q) = \alpha p \vee \alpha q$ if and only if $(p \vee q)^{\sharp} = p^{\sharp} \wedge q^{\sharp}$.

This means that every special property of a polarity can be translated into a corresponding special property of a hemimorphism, and consequently into a special property of a Boolean relation. It is, however, sometimes more revealing to use the complementary relation. If ρ is a relation from Y to X, the complementary relation ρ' from Y to X is the complement of ρ in the cartesian product $Y \times X$; that is, the settheoretic complement of ρ considered as a subset of $Y \times X$. Since it will be convenient to use such complementary relations, we shall introduce the following name for them.

DEFINITION. A relation ρ from one Boolean space Y to another Boolean space X will be called a polarity relation if it is the complementary relation σ' of a Boolean relation σ of Y into X. If $\#$ is a polarity of one Boolean algebra A into another Boolean algebra B, and if α is the hemimorphism of A into B defined by $\alpha p = (p^{\#})'$, then α and $\#$ will be said to be associated. If α^* is the dual Boolean relation for the hemimorphism α, the polarity relation $\alpha^{*\prime}$ will be called the conjugate relation of the polarity $\#$ associated with α.

Suppose, in the notation of this definition, that $\#$ is a polarity from A to B. For any clopen set P in X, there is an element $p \in A$ such that $p = \{x \in X : p(x) = 1\}$. We may, temporarily, denote by $p^{\#}$ the set $\{y \in Y : p^{\#}(y) = 1\}$. The comprement $(p^{\#})'$ of $p^{\#}$ in Y is given by the formulas

$$(p^{\#})' = \{y \in Y : (p^{\#})'(y) = 1\} = \{y \in Y : \alpha p(y) = 1\}$$
$$= \alpha^{*-1}\{x \in X : p(x) = 1\} = \alpha^{*-1}p \,.$$

Thus $y \in (P^{\#})'$ if and only if there is an element $x \in P$ such that $y\alpha^* x$, and hence $y \in P^{\#}$ if and only if $y\alpha^{*\prime}x$ for all $x \in P$. If ρ denotes the polarity relation $\rho = \alpha^{*\prime}$, then $y \in P^{\#}$ if and only if $y\rho x$ for all $x \in P$. In other words, every polarity has the form given by Birkhoff, if consideration is given to the topological structure of the dual spaces. Then Theorem 1 may be restated in the following (somewhat telegraphic) form.

THEOREM 2. *There is a one-to-one correspondence between polarities of Boolean algebras and polarity relations of Boolean spaces.*

Special properties of hemimorphisms have been investigated in terms of their dual Boolean relations [8]. It is thus quite easy to obtain the corresponding facts about polarities.

DEFINITION. A polarity $\#$ of A into B is called a DeMorgan polarity if $(p \wedge q)^{\#} = p^{\#} \vee q^{\#}$, for each $p, q \in A$, and if $1^{\#} = 0$.

THEOREM 3. *Let $\#$ be a polarity of A into B. The following are then equivalent*:

(i) *$\#$ is a DeMorgan polarity*;

(ii) *the associated hemimorphism α is a homomorphism*:

(iii) *the Boolean relation α^* is a function*;

(iv) *the polarity relation $\rho = \alpha^{*\prime}$, has the property that for any $y \in Y$, if x_1 and x_2 are distinct elements of X, then either $y\rho x_1$ or $y\rho x_2$; and*

(v) $(p^{\#})' = (p')^{\#}$, *for any* $p \in A$.

Proof. A hemimorphism α is a homomorphism if and only if $\alpha(p \wedge q) = \alpha p \wedge \alpha q$ for all $p, q \in A$, and $\alpha 1 = 1$. The theorem follows from this and from the fact that α is a homomorphism if and only if α^* is a function [2].

Let A, B, C be Boolean algebras, with dual spaces X, Y, Z respectively. If $\alpha : A \rightarrow B$ and $\beta : B \rightarrow C$ are hemimorphisms, then the product $\beta\alpha : A \rightarrow C$ is also a hemimorphism. Correspondingly if α is a Boolean from Z to Y and if ρ is a Boolean relation from Y to X, then the prodnct $\rho\sigma$ is a Boolean relation from Z to X [3]. Recall then $z(\rho\sigma)x$ if and only if there is an element $y \in Y$ such that $z\sigma y$ and $y\rho x$. If $\alpha^* = \rho$ and $\beta^* = \sigma$, then $(\beta\alpha)^* = \rho\sigma = \alpha^*\beta^*$ [2].

The iteration of two polarities is not, is general, a polarity. However, it may have other important properties; in particular, we are interested in the properties of a Galois correspondence. Note that if $\# : A \rightarrow B$ is a polarity, then $p_1 \leq p_2$ in A implies $p_2^{\#} \leq p_1^{\#}$ in B. This means that it is only the third condition in the definition of a Galois correspondence which needs investigation.

Recall that if ρ is a relation from Y to X, the inverse relation ρ^{-1} is a relation from X to Y, defined by declaring $x\rho^{-1}y$ if and only if $y\rho x$.

THEOREM 4. *Let* $\#$ *be a polarity from a Boolean algebra* A, *with dual space* X, *to a Boolean algebra* B, *with dual space* Y. *Let* α *be the associated hemimorphism, let* α^* *be the dual Boolean relation of* α, *and let* ρ *be the conjugate polarity relation of* $\#$. *Let* $+$ *be a polarity from* B *into* A, *let* β *be the associated hemimorphism, let* β^* *be the dual Boolean relation of* β, *and let* σ *be the conjugate polarity relation of* $\#$. *Then the following are equivalent*:

(i) $p \leq p^{\#+}$ *for each* $p \in A$;

(ii) $\beta(\alpha p)' \leq p'$ *for each* $p \in A$;

(iii) $x\beta^* y$ *implies* $y\alpha^* x$ *for each* $x \in X, y \in Y$;

(iv) $\beta^* \subset \alpha^{*-1}$:

(v) $y\rho x$ *implies* $x\sigma y$ *for each* $x \in X, y \in Y$; *and*

(vi) $\rho \subset \sigma^{-1}$.

Proof. The only problem is to show that (ii) and (iii) are equivalent. This will follow from the slightly more general result: for any two elements x_1 and $x_2 \in X$, we have $\beta(\alpha p)'(x_1) \leq p'(x_2)$ for all $p \in A$ if and only if, for any $y \in Y$, $x_1\beta^* y$ implies $y\alpha^* x_2$. For we have $\beta(\alpha p)'(x_1) = \text{lub}\{(\alpha p)'(y) : x_1\beta^* y\}$, so that $\beta(\alpha p)'(x_1) \leq p'(x_2)$ if and only if $x_1\beta^* y$ implies $p(x_2) \leq \alpha p(y)$. This last inequality holds for each $p \in A$ if and only if $y\alpha^* x_2$.

This result has a number of immediate corollaries which clarify the nature of Galois correspondences between Boolean algebras.

THEOREM 5. *In the notation of Theorem 3, the polarities* $\#$ *and* $+$ *define a Galois correspondence of A and B if and only if* $\beta^* = \alpha^{*-1}$.

This means that a given polarity $\#$ can have at most one other polarity $+$ which may be paired with it to yield a Galois correspondence. Theorem 2 showed that the method given by Birkhoff is the only way to obtain a polarity; this result shows that the same method is the only way to obtain a Galois correspondence. These facts also given an answer to an important question in connection with Boolean relations themselves: when is the inverse of a Boolean relation again a Boolean relation?

THEOREM 6. *Let* θ *be a Boolean relation from a Boolean space Y to another Boolean space X. A necessary and sufficient condition that the inverse relation* θ^{-1} *be a Boolean relation is that the dual hemimorphism* θ^* *of* $\phi(X)$ *into* $\phi(Y)$ *be the associated hemimorphism of a polarity of* $\phi(X)$ *to* $\phi(Y)$ *which is part of a Galois correspondence.*

In the special case of most importance, when $X = Y$, this condition becomes very simple.

THEOREM 7. *Let* $\#$ *be a polarity of a Boolean algebra A into itself, let* α *be its associated hemimorphism, let* α^* *be the dual relation of* α, *and let* ρ *be the conjugate polarity relation of* $\#$. *Then the following are equivalent*:

(i) $p \leq p^{\#\#}$ *for each* $p \in A$;

(ii) $\alpha(\alpha p)' \leq p'$ *for each* $p \in A$;

(iii) α^* *is symmetric*; *and*

(iv) ρ *is symmetric*.

A polarity has some to the properties of the complementation mapping $p \rightarrow p'$. We may ask what other properties of complementation it can have, and in particular, we may seek a characterization of complementation. Since we are given a Boolean algebra at the outset, there is already available one characterization of complementation: In any distributive lattice with 0 and 1, if every element p has an element p' such that $p \vee p' = 1$ and $p \wedge p' = 0$, then the element p' is unique. Furthermore, the mapping $p \rightarrow p'$ is a DeMorgan polarity satisfying $p = p''$. When we ask for a characterization of complementation, we ask for additional assumptions about a polarity $\#$ which imply that $p^{\#} = p'$ for each element p.

Let $\#$ be a polarity of a Boolean algebra A into itself. Theorem 3 gives precise conditions that $\#$ satisfy DeMorgan's laws, and Theorem 7 gives equally precise conditions that $p \leq p^{\#\#}$ for each $p \in A$. From the above list of properties of complementation, this leaves three attributes to be investigated; (1) $p \vee p^{\#} = 1$ for each $p \in A$; (2) $p \wedge p^{\#} = 0$ for each $p \in A$; and (3) $p^{\#\#} \leq p$ for each $p \in A$.

In the next two theorems, let α be the associated hemimorphism of the polarity $\#$ of A into itself, let α^* be the dual Boolean relation of α, and let ρ be the conjugate polarity ralation of $\#$.

THEOREM 8. *The following are equivalent*:

(i) $p \vee p^{\#} = 1$ *for each* $p \in A$;

(ii) $p^{\#} = p' \vee b$ *for each* $p \in A$, *where* $b \in A$ *is some fixed element*;

(iii) *if* $x_1 \alpha^* x_2$, *then* $x_1 = x_2$; *and*

(iv) *if* $x_1 \neq x_2$, *then* $x_1 \rho x_2$.

In particular, the following are equivalent:

(I) $p \vee p^{\#} = 1$ *for each* $p \in A$, *and* $1^{\#} = 0$;

(II) $p^{\#} = p'$ *for each* $p \in A$;

(III) $x_1 \alpha^* x_2$ *if and only if* $x_1 = x_2$; *and*

(IV) $x_1 \rho x_2$ *if and only if* $x_1 \neq x_2$.

Proof. It is easily seen that $p \vee p^{\#} = 1$ if and only if $\alpha p \leq p$. It is known [8] that a hemimorphism α satisfies this condition for each $p \in A$ if and only if $\alpha p = p \wedge a$, for some fixed $a \in A$. The theorem follows from this fact.

THEOREM 9. *The following are equivalent*:

(i) $p \wedge p^{\#} = 0$ *for each* $p \in A$;

(ii) $p \leq \alpha p$ *for each* $p \in A$;

(iii) α^* *is reflexive*; *and*

(iv) ρ *is irreflexive*.

Proof. The equivalence of (ii) with (iii) is proved in [8]; the equivalence of the others is then trivial. (Note that an irreflexive relation ρ is one such that either $x\rho' x$ or else $x \rho x$ implies $x \rho y$ for all y.)

The problem of condition (3) can be treated in more generality. Return for a moment to the definition of a Galois correspondence. If we retain (i) and (ii) of this definition, but alter (iii) to read (iii′) $p^{\#+} \leq p$ and $q^{+\#} \leq q$, then the lemma stated at the beginning must also be altered. In fact, the conclusion becomes $1^{\#} = 0$, $1^{+} = 0$, $(p_1 \wedge p_2)^{\#} = p_1^{\#} \vee p_2^{\#}$, $(q_1 \wedge q_2)^{+} = q_1^{+} \vee q_2^{+}$. These properties might also be considered in the manner in which we have treated of polarities. There is obviously no need to do this.

However, if we have two *polarities* $\#$ and $+$ having the property given by (iii′) above, then the altered lemma shows that we have two DeMorgan polarities. If α and β are the associated hemimorphisms, then both α and β are homomorphisms. Furthermore, $\beta\alpha p = p^{\#'+'} = p^{\#+} \leq p$, for each $p \in A$. Then [8], $\beta\alpha p = p \wedge a$, and since $\beta\alpha 1 = 1$, we have $\beta\alpha p = p$ for each $p \in A$. This gives the following result.

THEOREM 10. *Let* $\sharp$ *be a polarity of* A *into* B *and let* $+$ *be a polarity of* B *into* A. *In the notation of Theorem* 4, *the following are equivalent*:

(i) $p^{\sharp +} \leq p$ *and* $q^{+\sharp} \leq q$ *for each* $p \in A$ *and* $q \in B$;

(ii) $p^{\sharp +} = p$ *and* $q^{+\sharp} = q$ *for each* $p \in A$ *and* $q \in B$;

(iii) α *and* β *are reciprocal isomorphisms of* A *onto* B *and of* B *onto* A *respectively*; *and*

(iv) α^* *and* β^* *are reciprocal homeomorphisms of* Y *onto* X *and of* X *onto* Y *respectively.*

REFERENCES

1. G. Birkhoff, *Lattice theory*, Amer. Math. Soc. Colloq. Publ. **XXV**, New York, 1948.
2. P. R. Halmos, *Algebraic logic I*, Compositio Mathematica **12** (1955), 217–249.
3. ———, *Algebraic logic II*, Fundamenta Mathematica **43** (1956), 255–325.
4. B. Jónsson and A. Tarski, *Boolean algebras with operators*, Amer. J. Math. **73** (1951), 891–939.
5. O. Ore, *Galois connexions*, Trans. Amer. Math. Soc. **55** (1944), 493–513.
6. M. H. Stone, *A theory of representations of Boolean algebras*, Trans. Amer. Math. **40** (1936), 37–111.
7. ———, *Applications of the theory of Boolean rings to general topology*, Trans. Amer. Math. Soc. **41** (1937), 375–481.
8. F. B. Wright, *Some remarks on Boolean duality*, Portugaliae Mathematica **16** (1957), 109–117.

KING'S COLLEGE, CAMBRIDGE

Pacific Journal of Mathematics

Vol. 10, No. 2 October, 1960

Maynard G. Arsove, *The Paley-Wiener theorem in metric linear spaces* 365

Robert (Yisrael) John Aumann, *Acceptable points in games of perfect information* 381

A. V. Balakrishnan, *Fractional powers of closed operators and the semigroups generated by them* 419

Dallas O. Banks, *Bounds for the eigenvalues of some vibrating systems* 439

Billy Joe Boyer, *On the summability of derived Fourier series* 475

Robert Breusch, *An elementary proof of the prime number theorem with remainder term* 487

Edward David Callender, Jr., *Hölder continuity of n-dimensional quasi-conformal mappings* 499

L. Carlitz, *Note on Alder's polynomials* 517

P. H. Doyle, III, *Unions of cell pairs in* E^3 521

James Eells, Jr., *A class of smooth bundles over a manifold* 525

Shaul Foguel, *Computations of the multiplicity function* 539

James G. Glimm and Richard Vincent Kadison, *Unitary operators in* C^**-algebras* 547

Hugh Gordon, *Measure defined by abstract* L_p *spaces* 557

Robert Clarke James, *Separable conjugate spaces* 563

William Elliott Jenner, *On non-associative algebras associated with bilinear forms* 573

Harold H. Johnson, *Terminating prolongation procedures* 577

John W. Milnor and Edwin Spanier, *Two remarks on fiber homotopy type* 585

Donald Alan Norton, *A note on associativity* 591

Ronald John Nunke, *On the extensions of a torsion module* 597

Joseph J. Rotman, *Mixed modules over valuations rings* 607

A. Sade, *Théorie des systèmes demosiens de groupoï des* 625

Wolfgang M. Schmidt, *On normal numbers* 661

Berthold Schweizer, Abe Sklar and Edward Oakley Thorp, *The metrization of statistical metric spaces* 673

John P. Shanahan, *On uniqueness questions for hyperbolic differential equations* 677

A. H. Stone, *Sequences of coverings* 689

Edward Oakley Thorp, *Projections onto the subspace of compact operators* 693

L. Bruce Treybig, *Concerning certain locally peripherally separable spaces* 697

Milo Wesley Weaver, *On the commutativity of a correspondence and a permutation* 705

David Van Vranken Wend, *On the zeros of solutions of some linear complex differential equations* 713

Fred Boyer Wright, Jr., *Polarity and duality* 723