Pacific Journal of Mathematics

IN THIS ISSUE—

Vol. 4, No. 4 **August, 1954**

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 of the editors. Manuscripts intended for the outgoing editors should be sent to their successors. All other communications to the editors should be addressed to the managing editor, E.G. Straus, at the University of California Los Angeles 24, California.

50 reprints 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 (Volumes 1, 2, 3) are available at $2.50 per copy. 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 the publishers, University of California Press, Berkeley 4, California.

Printed at Ann Arbor, Michigan. Entered as second class matter at the Post Office, Berkeley, California.

*To be succeeded in 1955, by H.L. Royden, Stanford University, Stanford, California.

**To be succeeded in 1955, by E.G. Straus, University of California, Los Angeles 24, Calif.

UNIVERSITY OF CALIFORNIA PRESS · BERKELEY AND LOS ANGELES

ORTHONORMAL CYCLIC GROUPS

PAUL CIVIN

In an earlier paper [1] a characterization was given of the Walsh functions in terms of their group structure and orthogonality. The object of the present note is to present a similar result concerning the complex exponentials.

THEOREM. *Let* $\{A_n(x)\}$ $(n = 0, \pm 1, \cdots; 0 \leq x \leq 1)$ *be a set of complex-valued measurable functions which is a multiplicative cyclic group. A necessary and sufficient condition that* $\{A_n(x)\}$ *be an orthonormal system over* $0 \leq x \leq 1$ *is that the generator of the group admit a representation* $\exp(2\pi i\, c(x))$ *almost everywhere, with* $c(x)$ *equimeasurable with* x.

As the sufficiency is immediate, we present only the proof of the necessity. Let the notation be chosen so that the generator of the group is $A_1(x)$, and

$$A_n(x) = (A_1(x))^n \qquad (n = 0, \pm 1, \cdots).$$

The normality implies $|A_1(x)| = 1$ almost everywhere. Hence there is a measurable $a(x)$, $0 \leq a(x) < 1$, such that

$$A_1(x) = \exp(2\pi i\, a(x))$$

almost everywhere. Let $b(x)$ be a function [2, p. 207] monotonically increasing and equimeasurable with $a(x)$. Also let

$$c(x) = m\{u : 0 \leq u \leq 1,\ b(u) \leq x\} \qquad (-\infty < x < \infty).$$

The orthonormal condition becomes

$$\delta_{0,n} = \int_0^1 \exp(2\pi ni\ b(x))dx = \int_{-\infty}^{\infty} \exp(2\pi niy)\, dc(y),$$

where the latter integral is a Lebesgue-Stieltjes integral. Thus for any $\epsilon > 0$,

Received February 6, 1953.
Pacific J. Math. 4 (1954), 481-482

$$\delta_{0,n} = \int_{b(0)-\epsilon}^{b(1)} \exp(2\pi niy)\, dc(v)$$

$$= \int_{b(0)}^{b(1)} \exp(2\pi niy)\, dc(y) + \exp(2\pi ni\, b(0))\ \ m\{x : b(x) = b(0)\},$$

and the latter integral is interpretable as a Riemann-Stieltjes integral.

Integration by parts yields

$$(1) \qquad \delta_{0,n} = \exp(2\pi ni\, b(1)) - 2\pi ni \int_{b(0)}^{b(1)} c(y) \exp(2\pi niy)\, dy.$$

If $f(y) = y$, $0 < y \leq 1$, and $f(y+1) = f(y)$, a direct calculation shows that

$$(2) \qquad \delta_{0,n} = \exp(2\pi ni\, b(1)) - 2\pi ni \int_0^1 f(y - b(1)) \exp(2\pi niy)\, dy.$$

Formulas (1) and (2), and the completeness of the complex exponentials, imply the existence of a constant k such that for almost all y, $0 < y \leq 1$,

$$f(y - b(1)) + k = \begin{cases} 0, & 0 < y \leq b(0) \\ c(y), & b(0) < y \leq b(1) \\ 0, & b(1) < y \leq 1. \end{cases}$$

Since the supremum of $c(y)$ is one, and $f(y)$ has no interval of constancy, one infers that $k = 0$, $b(0) = 0$, and $b(1) = 1$. Thus $c(y) = y$, $0 < y \leq 1$, which is equivalent to the proposition that was asserted.

References

1. P. Civin, *Multiplicative closure and the Walsh functions*, Pacific J. Math. 2 (1952), 291-296.

2. N. J. Fine, *On groups of orthonormal functions*, Pacific J. Math. (to appear).

3. ———, *On groups of orthonormal functions*, II, Pacific J. Math. (to appear).

4. D. Jackson, *Proof of a theorem of Haskins*, Transa. Amer. Math. Soc. 17 (1916), 178-180.

5. A. Zygmund, *Trigonometrical series*, Warsaw-Lvov, 1935.

THE UNIVERSITY OF OREGON

THE RATE OF INCREASE OF REAL CONTINUOUS SOLUTIONS OF ALGEBRAIC DIFFERENTIAL-DIFFERENCE EQUATIONS OF THE FIRST ORDER

K. L. COOKE

1. Introduction. It is the purpose of this paper to prove several theorems describing the rate of increase, as $t \longrightarrow +\infty$, of real solutions of algebraic differential-difference equations of the form

$$(1) \qquad P(t, u(t), u'(t), u(t+1), u'(t+1)) = 0.$$

In this equation, and throughout this paper, $P(t, u, v, \cdots)$ denotes a polynomial in the variables $t, u, v, \cdots$, with real coefficients, and a prime denotes differentiation with respect to t. In order to explain the significance and limitations of these theorems, it is first necessary to summarize the work, by other investigators, which suggested the present discussion.

In 1899, E. Borel, [1], published a memoir in which he studied the magnitude of solutions of algebraic differential equations. His result, as later improved by E. Lindelöf, [4], is quoted here for reference:

Let $u(t)$ be a real function which is defined and which has a continuous first derivative for all t larger than t_0, and which satisfies the first order algebraic differential equation

$$(2) \qquad P(t, u(t), u'(t)) = 0$$

for $t > t_0$. Then there is a positive number k, which depends only on P, such that

$$|u(t)| < \exp(t^k/k)$$

for $t \geq t_0$.

It is noteworthy that it is impossible to prove a result of the above type for higher order equations. For a discussion of this point, refer to Vijayaraghavan, [7].

Received January 10, 1953.
Pacific J. Math. 4 (1954), 483-501

Extensions of the Borel-Lindelöf method to difference equations have already been effected by Lancaster, [3], and Shah, [5] and [6]. Shah demonstrated that no theorem comparable to that of Borel and Lindelöf can be obtained for the class of algebraic difference equations of the form

$$P(t, u(t), u(t+1)) = 0. \tag{3}$$

For, let $g(t)$ be an arbitrary increasing function which becomes indefinitely large as $t \longrightarrow +\infty$. Shah proved that it is possible to construct an equation of the type (3) with a real solution $u(t)$ which exists and is continuous for $t \geq t_0$ and which exceeds $g(t)$ at each point of a sequence $\{t_n\}$ such that $t_n \longrightarrow +\infty$ as $n \longrightarrow \infty$. The situation with respect to higher order equations is similar. Shah did, however, obtain the following weaker results concerning the possible rate of growth of solutions of (3):

There exists a positive number A, which depends only on the polynomial P, with the following property: if $u(t)$ is, for all $t \geq t_0$, a real continuous solution of (3), *then there is no number T such that*[1]

$$|u(t)| > e_2(At) \qquad \textit{for all } t \geq T.$$

That is, for each such solution $u(t)$ there is a sequence $t_1, t_2, \cdots$ ($t_n \longrightarrow +\infty$ as $n \longrightarrow \infty$) such that

$$|u(t)| \leq e_2(At) \tag{4}$$

for $t = t_1, t_2, \cdots$. If $u(t)$ is a real, continuous, monotonic solution of (3) *for $t \geq t_0$, then there exists a number $\tau \geq t_0$ such that* (4) *holds for all $t \geq \tau$.*

We shall now turn to a discussion of the class of differential-difference equations of the form (1). We first make the following definition.

DEFINITION. A real function $u(t)$ will be said to be a *proper solution* of a differential-difference equation (1) if there exists a number t_0 such that $u(t)$ exists and is a solution of (1) for all $t \geq t_0$, and such that $u(t)$ has a continuous first derivative for $t \geq t_0$.

In view of Shah's results on difference equations, it is not to be expected that a theorem analogous to the Borel-Lindelöf theorem should hold for first

[1] We here employ the notation $e_2(x) = \exp(e^x)$ which was adopted by G. H. Hardy, [2].

order differential-difference equations. However, it might be expected that a result like that of Shah could be obtained for equations of the class (1). This is not the case, as is shown by the following theorem.

THEOREM 1. *Let $g(t)$ be an arbitrary increasing function which becomes indefinitely large as $t \longrightarrow +\infty$. It is possible to construct an algebraic differential-difference equation of the form*

$$P(t, u(t), u'(t), u(t+1), u'(t+1)) = 0 \tag{1}$$

which has a proper solution $u(t)$ which exceeds $g(t)$ for all t. This statement remains valid if equation (1) is replaced by the equation

$$P(t, u(t), u'(t), u'(t+1)) = 0. \tag{5}$$

Proof. We shall prove this theorem at once by constructing a suitable example. Define a function $u(t)$ as follows. Let $u(t) = g(n+2)+1$ in the interval $[n, n+1]$, for $n = 0, 2, 4, \cdots$. In the intervals $[n, n+1]$, where $n = 1, 3, 5, \cdots$, let $u(t)$ be any continuous, non-decreasing function which has a continuous first derivative, and for which

$$u(n) = g(n+1)+1, \; u(n+1) = g(n+3)+1, \; u'(n) = u'(n+1) = 0.$$

It is clear that the function so defined satisfies the equation

$$u'(t)u'(t+1) = 0 \tag{6}$$

for all $t > 0$. Furthermore, $u(t)$ is non-decreasing for all t and $u(t) > g(t)$ for $t \geq 0$. Since equation (6) is in the class of equations of the form (1), and in the class of equations of the form (5), the proof of Theorem 1 is complete.

This theorem is in sharp contrast to those for algebraic differential or difference equations. It shows that no bound at all can be placed on the rate of growth of solutions of differential-difference equations of the form (1). The same difficulty intrudes even if we speak only of monotone solutions.

It is, however, possible to obtain useful bounds on the rate of growth of solutions of less general classes of differential-difference equations. We observe first of all that, according to Theorem 1, no results like those of Borel or Shah can be obtained for the class of equations of the form (5). We shall, however, prove analogous results for equations of the following types:

$$P(t, u(t), u'(t+1)) = 0 \tag{7}$$

$$(8) \qquad P(t, u(t), u'(t), u(t+1)) = 0$$

$$(9) \qquad P(t, u'(t), u(t+1)) = 0.$$

Even for such equations it is not possible to establish a theorem like the Borel-Lindelöf theorem. This may be seen from the following simple counterexample. Let $g(t)$ be an arbitrary real, continuous, increasing function which becomes indefinitely large as $t \longrightarrow +\infty$. Let m be any non-negative integer. Let $u(t) = t^m$ for t in the intervals $[2n, 2n+1]$, $n = 0, 1, 2, \cdots$. For t in the intervals $(2n+1, 2n+2)$, $n = 0, 1, 2, \cdots$, let $u(t)$ be defined in any convenient fashion for which $u'(t)$ is continuous and $u(2n + 3/2) > g(2n + 3/2)$. This function $u(t)$ exceeds $g(t)$ for arbitrarily large values of t, and satisfies each of the following equations for all $t > 0$:

$$(10) \qquad \{u'(t+1) - m(t+1)^{m-1}\}\{u(t) - t^m\} = 0$$

$$(11) \qquad [u(t+1) - (t+1)^m][u'(t) - mt^{m-1}] = 0.$$

Note that (10) is an equation in the class (7) and equation (11) is in the class (8) and (9). Furthermore, all the above remarks are correct for $m = 0$, in which case (10) and (11) are equations with constant coefficients. The following theorem has therefore been proved.

THEOREM 2. *Let $g(t)$ be an arbitrary increasing function which becomes indefinitely large as $t \longrightarrow +\infty$. It is possible to construct a first order algebraic differential-difference equation of the form*

$$(7) \qquad P(t, u(t), u'(t+1)) = 0$$

with a proper solution $u(t)$ which exceeds $g(t)$ at each point of a sequence $\{t_n\}$ for which $t_n \longrightarrow +\infty$ as $n \longrightarrow \infty$. The statement remains true if (7) is replaced by equation (8) or equation (9), or by one of the equations

$$(12) \qquad P(u(t), u'(t+1)) = 0$$

$$(13) \qquad P(u'(t), u(t+1)) = 0.$$

Although we cannot establish theorems of the Borel-Lindelöf type for the classes of equations mentioned above, we have proved several results analogous to those of Shah. These results are stated in Theorems 3, 5, and 6 of § 3 below. Moreover, in Theorem 4, stated below, we have proved a theorem of the Borel-Lindelöf type for a certain subclass of equations of the type (7). No theorems are given in this paper for equations with higher order derivatives or differences, since results like those mentioned above can be obtained only for rather special classes of such equations.

2. Lemmas. In this section, we shall prove several lemmas which will be required in proving the theorems of § 3.

LEMMA 1. *Suppose that* $u(t)$ *is, for all* $t \geq t_0$, *a positive function with a continuous first derivative. Let* A *and* B *be two positive numbers for which* $B < e^A$. *Let* C *be an arbitrary non-negative number. Suppose that there is a sequence* $\{\tau_n\}$ *for which* $\tau_n \longrightarrow +\infty$ *as* $n \longrightarrow \infty$ *and for which* $u(\tau_n) \geq e_2(A\tau_n)$. *Then there exists a sequence* $\{t_n\}$ *for which* $t_n \longrightarrow +\infty$ *as* $n \longrightarrow \infty$ *and for which*

$$u'(t_n + 1) > t_n^C u(t_n)^B.$$

Proof. Assume that $u(t)$ is a positive function with a continuous first derivative, and that

$$u'(t+1) \leq t^C u(t)^B \tag{14}$$

for all $t \geq T$. We shall prove that as a consequence there is a number T_2 such that

$$u(t) < e_2(At) \tag{15}$$

for $t \geq T_2$. This will prove Lemma 1. We divide the proof of (15) into two cases.

Case 1. We assume that $B > 1$. We may, of course, suppose that T is as large as is convenient; choose T so large that

$$B^{j-1} \log T > j - 1 \qquad (j = 1, 2, 3, \cdots). \tag{16}$$

This is certainly true for j sufficiently large if $\log T > 0$, and by choosing T large enough we can ensure that it is true for all j. Then for $j = 1, 2, 3, \cdots$,

$$(2T)^{B^{j-1}} \geq 2T^{B^{j-1}} > T + e^{j-1} \geq T + j. \tag{17}$$

Having chosen T, define

$$M' = \max_{T \leq t \leq T+1} u(t), \qquad M = \max(M', 1).$$

We shall now prove by induction that

$$u(t) \leq M^{B^n} \prod_{j=0}^{n} (T+j)^{(C+1)B^{n-j}} \tag{18}$$

for $T + n \leq t \leq T + n + 1$ $(n = 0, 1, 2, \cdots)$. This is evident for $n = 0$. Suppose that (18) has been proved for $n = k - 1$ $(k \geq 1)$. Then by (18) and (14)

$$u'(t+1) \leq (T+k)^C M^{B^k} \prod_{j=0}^{k-1} (T+j)^{(C+1)B^{k-j}}$$

for $T + k - 1 \leq t \leq T + k$. Upon observing that the right hand side of inequality (18) is an increasing function of n, and employing (14) again, we get

$$u(T+k) \leq u(T+1) + \int_T^{T+k-1} u'(t+1)\,dt$$

$$\leq M + (T+k)^C (k-1) M^{B^k} \prod_{j=0}^{k-1} (T+j)^{(C+1)B^{k-j}}$$

$$\leq (T+k)^C (T+k-1) M^{B^k} \prod_{j=0}^{k-1} (T+j)^{(C+1)B^{k-j}}.$$

On integrating the first inequality under (18) from $T + k - 1$ to t, where $t \leq T + k$, and combining with the inequality just derived, we obtain

$$u(t+1) \leq t(T+k)^C M^{B^k} \prod_{j=0}^{k-1} (T+j)^{(C+1)B^{k-j}} \quad (T+k-1 \leq t \leq T+k).$$

Replacing t by $T + k$ in the right member of the above inequality, we see that (18) is valid for $n = k$. This completes the inductive proof of (18).

We now employ (17). (18) takes the form

$$u(t) \leq \left[M(2T)^{(n+1)(C+1)}\right]^{B^n} = e_2\left[n \log B + \log\{(1+n)(C+1)\log(2T) + \log M\}\right]$$

for $T + n \leq t \leq T + n + 1$. Let $R = \max(2T, M)$. Then

$$u(t) \leq e_2\left[n \log B + \log(nC + n + C + 2) + \log\log R\right]$$

for $T + n \leq t \leq T + n + 1$. Since $\log B < A$ by hypothesis, (15) follows.

Case 2. We now assume that $B \le 1$. Using the same method as in Case 1, we can easily prove by induction that

$$(19) \qquad u(t) \le M \prod_{j=0}^{n} (T+j)^{C+1}$$

for $T + n \le t \le T + n + 1$ $(n = 0, 1, 2, \cdots)$. Hence

$$u(t) \le M(T+n)^{(n+1)(C+1)} \le Mt^{(C+1)(t-T+1)}$$

for $T + n \le t \le T + n + 1$. (15) follows at once. This completes the proof of Lemma 1.

LEMMA 2. *Suppose that $u(t)$ is, for all $t \ge t_0$, a positive non-decreasing function with a continuous first derivative, and that $u(t) \ge e_2(At)$ for $t \ge t_0$. Let B and C be any non-negative numbers for which $B + C < e^A$, and let D be any non-negative number. Then, given any positive number ϵ, there exists a sequence $t_1, t_2, \cdots$ ($t_n \longrightarrow +\infty$ as $n \longrightarrow \infty$) such that*

$$(20) \qquad u(t_n + 1) \ge u(t_n)^B u'(t_n)^C$$

$$(n = 1, 2, \cdots).$$

$$(21) \qquad t_n^D u(t_n) \le u'(t_n) \le u(t_n)^{1+\epsilon}$$

Proof. We divide the proof into two cases.

Case 1. Suppose that $u'(t) \ge t^D u(t)$ for all sufficiently large t, say for $t \ge t_0$. It will be sufficient to prove the lemma for values of ϵ so small that $(B + C)(1 + \epsilon) < e^A$. Let ϵ be any such number, and let $\alpha = (B + C)(1 + \epsilon)$.

Borel, [1], proved that if a function $u(t)$ is, for all $t \ge t_0$, a positive, non-decreasing function with a continuous first derivative, then, given any positive number ϵ, $u'(t) \ge u(t)^{1+\epsilon}$ at most on a set of intervals the sum of whose lengths is a finite number (which depends on ϵ). This result will hereafter be referred to as Borel's Lemma.

If $u(t)$ satisfies the hypotheses of Lemma 2, then by Borel's Lemma there is a number $T \ge t_0$ such that $u'(t) \le u(t)^{1+\epsilon}$ for all $t \ge T$, except for t belonging to a set E of intervals of total length less than $1/2$. We can now choose a number $\tau > T$ such that no point of the sequence $\tau, \tau + 1, \tau + 2, \cdots$, belongs to E. It follows that (21) holds for each point $t_n = \tau + n$. We shall now show that (20) holds at the points of an infinite subsequence of the sequence $\{\tau+n\}$. If this is not true, there is an integer N such that

$$u(\tau+n+1) < u(\tau+n)^B u'(\tau+n)^C \qquad \text{for all } n \geq N.$$

This implies that

$$u(\tau+n+1) < u(\tau+n)^{\alpha} \qquad \text{for } n \geq N.$$

It follows that

$$u(\tau+N+m) < e_2[m \log \alpha + \log\log u(\tau+N)]$$

for $m = 1, 2, 3, \ldots$. Since $\log \alpha < A$, this contradicts the hypothesis that $u(t) \geq e_2(At)$ for $t \geq t_0$. It follows that there is an infinite subsequence of the sequence $\{\tau+n\}$ at which (20) is valid. This completes the proof in Case 1.

Case 2. The alternative to the supposition of Case 1 is that $u'(t) < t^D u(t)$ for arbitrarily large values of t. We define α as in Case 1, and again suppose ϵ so small that $\log \alpha < A$. From the fact that $u(t) \geq e_2(At)$ it follows that $u'(t) > t^D u(t)$ for arbitrarily large t. By continuity of $u(t)$ and $u'(t)$, $u'(t) < t^D u(t)$ in open intervals, and $u'(t) \geq t^D u(t)$ in closed intervals. Let the open intervals be

$$(a_1, b_1), (a_2, b_2), \ldots, (a_n, b_n), \ldots$$

$(a_1 \geq t_0)$ and let the closed intervals be

$$[b_1, a_2], \ [b_2, a_3], \ldots, [b_n, a_{n+1}], \ldots.$$

Note that $a_n \longrightarrow +\infty$ and $b_n \longrightarrow +\infty$, and that

$$u'(a_n) = a_n^D u(a_n) \ \text{ and } \ u'(b_n) = b_n^D u(b_n). \tag{22}$$

By Borel's Lemma, $u'(t) \leq u(t)^{1+\epsilon}$ for all t except for t in a set E of open intervals of finite total length. Let E_n be the subset of E contained in $[b_n, a_{n+1}]$ and let L_n be the sum of the lengths of the intervals of E_n. Then $\lim L_n = 0$ as $n \longrightarrow \infty$. We shall prove that there are arbitrarily large values of n for which there is at least one point t_n in the interval $[b_n, a_{n+1}]$ such that

$$u(t_n+1) \geq u(t_n)^B u'(t_n)^C$$

and such that t_n is not in E_n. The proof will be by contradiction. Assume the contrary. Then there is a positive integer N such that, for every $n \geq N$,

$$(23) \qquad u(t+1) < u(t)^B u'(t)^C$$

for all t which are in $[b_n, a_{n+1}]$ but not in E_n.

First we suppose that $0 < \alpha \leq 1$. Since $u(t) \geq e_2(At)$, we may select an integer $p \geq N$ such that

$$u(b_n)^{\epsilon} > b_n^D \qquad \text{for all } n \geq p.$$

Equations (22) therefore imply that

$$u(b_n)^{1+\epsilon} > u'(b_n) \qquad \text{for all } n \geq p.$$

Hence b_n is not in E_n if $n \geq p$. Consequently (23) implies that

$$u(b_p + 1) < u(b_p)^B u'(b_p)^C < u(b_p)^{\alpha} \leq u(b_p).$$

But $u(t)$ is non-decreasing. Thus we have reached a contradiction, and (23) cannot be true if $0 < \alpha \leq 1$.

Suppose, then, that $\alpha > 1$. Just as before, we may select an integer $p \geq N$ such that b_n is not in E_n for $n \geq p$. We also choose p so large that $L_n < 1$ for $n \geq p$ and so large that

$$(24) \qquad \max_{\zeta \geq b_p} \frac{(D+1)\zeta^D}{\alpha^{\zeta} \log \alpha} < \alpha^{-b_p} \log u(b_p).$$

This is possible because the right-hand member becomes indefinitely large as $p \longrightarrow +\infty$, since $u(t) \geq e_2(At)$ and $A > \log \alpha$, and because the maximum in the left member is finite. Define $c_p = b_p$. We shall now employ an inductive method to establish the existence of a sequence $c_p, c_{p+1}, c_{p+2}, \cdots$, for which

$$(25) \qquad \log u(c_{p+i}) \leq \alpha^{c_{p+i} - c_p + \sum \delta_j} \log u(c_p)$$

$(i = 0, 1, 2, \cdots)$, where the summation is over all $j \geq p$ for which $b_j \leq c_{p+i-1}$, and where the δ_j are defined below. In the first place, it is clear that (25) is true for $i = 0$. Suppose that we have established the existence of points c_{p+1}, $c_{p+2}, \cdots, c_{p+k-1}$ $(k \geq 1)$ for which (25) holds. There are now two possibilities:

(a) One possibility is that the point c_{p+k-1} lies in an interval $[b_q, a_{q+1}]$ for some value of q. If this is so, c_{p+k-1} may lie in E_q, or it may not. Let $\epsilon_{q,1}$ be the smallest non-negative number such that $c_{p+k-1} - \epsilon_{q,1}$ is in $[b_q, a_{q+1}]$ but not in E_q. Such a number exists, since b_q is not in E_q. Then, by (23),

$$u(c_{p+k-1} + 1 - \epsilon_{q,1}) < u(c_{p+k-1} - \epsilon_{q,1})^B u'(c_{p+k-1} - \epsilon_{q,1})^C.$$

By Borel's Lemma and the fact that $u(t)$ is non-decreasing, this gives rise to

$$u(c_{p+k-1} + 1 - \epsilon_{q,1}) < u(c_{p+k-1})^\alpha.$$

Since $\epsilon_{q,1} < L_q < 1$, the points c_{p+k-1} and $c_{p+k-1} + 1 - \epsilon_{q,1}$ cannot lie in the same interval of E_q. If $c_{p+k-1} + 1 - \epsilon_{q,1} > a_{q+1}$, we define $c_{p+k} = c_{p+k-1} + 1 - \epsilon_{q,1}$. If not, we proceed as follows. Let $\epsilon_{q,2}$ be the smallest non-negative number such that $c_{p+k-1} + 1 - \epsilon_{q,1} - \epsilon_{q,2}$ is in $[b_q, a_{q+1}]$ but not in E_q. Using (23) again, we find that

$$u(c_{p+k-1} + 2 - \epsilon_{q,1} - \epsilon_{q,2}) < u(c_{p+k-1} + 1 - \epsilon_{q,1} - \epsilon_{q,2})^\alpha$$

and therefore that

$$\log u(c_{p+k-1} + 2 - \epsilon_{q,1} - \epsilon_{q,2}) < \alpha^2 \log u(c_{p+k-1}).$$

We continue in this manner, obtaining a sequence of points

$$c_{p+k-1},\ c_{p+k-1} + 1 - \epsilon_{q,1},\ c_{p+k-1} + 2 - \epsilon_{q,1} - \epsilon_{q,2},\ \cdots,$$

no two of which can lie in the same interval of E_q. Let

$$c_{p+k-1} + r - \epsilon_{q,1} - \cdots - \epsilon_{q,r}$$

be the first point in this sequence which is larger than a_{q+1}; such a point must exist since

$$\delta_q = \epsilon_{q,1} + \epsilon_{q,2} + \cdots + \epsilon_{q,r} \leq L_q < 1.$$

Define

$$c_{p+k} = c_{p+k-1} + r - \epsilon_{q,1} - \cdots - \epsilon_{q,r}.$$

Then

$$\log u(c_{p+k}) < \alpha^r \log u(c_{p+k-1}) = \alpha^{c_{p+k}-c_{p+k-1}+\delta_q} \log u(c_{p+k-1}).$$

Combining this result with (25) for $i = k - 1$, we find that (25) holds for $i = k$, with the choice of c_{p+k} made above.

(b) The alternative to (a) is that c_{p+k-1} lies in an interval (a_q, b_q) for some value of q. (In this case $k \geq 2$, since $c_p = b_p$.) Now $u'(t) < t^D u(t)$ for all t in this interval. Hence, by integration,

$$(26) \qquad u(b_q) < u(c_{p+k-1}) \exp(b_q^{D+1} - c_{p+k-1}^{D+1}).$$

In this case, we define $c_{p+k} = b_q$. We shall now show that, with this choice of c_{p+k}, (25) is satisfied for $i = k$. From the extended theorem of the mean and the inequality (24) we can deduce

$$c_{p+k}^{D+1} - c_{p+k-1}^{D+1} < (\alpha^{c_{p+k}} - \alpha^{c_{p+k-1}})\alpha^{-c_p} \log u(c_p).$$

This inequality will still be true if, in the right member, we place the additional factor

$$\alpha^{\sum \delta_j},$$

where the summation is over all $j \geq p$ for which $b_j \leq c_{p+k-2}$. Using this result, (26), and (25) for $i = k - 1$, we obtain

$$\log u(c_{p+k}) < c_{p+k}^{D+1} - c_{p+k-1}^{D+1} + \log u(c_{p+k-1}) < \alpha^{c_{p+k}-c_p+\sum \delta_j} \log u(c_p).$$

The inequality (25) for $i = k$ is an immediate consequence.

This completes the proof that there is a sequence of points c_{p+i} for which (25) is valid. It is clear that $c_{p+i} \longrightarrow +\infty$ as $i \longrightarrow \infty$. Since the sum of the δ_j $(j = 1, 2, \cdots)$ is no greater than the sum L of the lengths of all the intervals of E,

$$\log u(t) \leq \alpha^{t-c_p+L} \log u(c_p) = \exp[(t - c_p + L)\log \alpha + \log \log u(c_p)]$$

for $t = c_{p+i}$ $(i = 0, 1, 2, \cdots)$. Since $\log \alpha < A$, it is a consequence of the above inequality that there is a positive integer I such that $\log u(t) < \exp(At)$ for $t = c_{p+i}$ $(i \geq I)$. This contradicts the hypothesis of Lemma 2. Therefore (23) cannot be true if $\alpha > 1$. Hence, no matter what the value of α, there are arbitrarily large values of n for which there is at least one point t_n in the interval

$[b_n, a_{n+1}]$ such that

$$u(t_n+1) \geq u(t_n)^B u'(t_n)^C,$$

and such that t_n is not in E_n. Since t_n is not in E_n, $u'(t_n) \leq u(t_n)^{1+\epsilon}$ for each such t_n. Since t_n lies in $[b_n, a_{n+1}]$, we have $u'(t_n) \geq t_n^D u(t_n)$. This completes the proof of the lemma.

LEMMA 3. *Suppose that $u(t)$ is, for all $t \geq t_0$, a positive function with a continuous first derivative, and that $u(t) \geq e_2(At)$ for all $t \geq t_0$. Let C be any non-negative number less than e^A. Then there is a sequence $t_1, t_2, \cdots$ ($t_n \longrightarrow +\infty$ as $n \longrightarrow \infty$) such that*

$$u(t_n+1) \geq u'(t_n)^C, \quad u'(t_n) \geq e^{t_n}.$$

Proof. First we suppose that there is a number $T \geq t_0$ such that $u'(t) \geq e^t$ for $t \geq T$. Then $u(t)$ is non-decreasing for $t \geq T$, and the result follows at once from Lemma 2.

On the other hand, suppose that $u'(t) < e^t$ for arbitrarily large values of t. Since $u(t) \geq e_2(At)$, $u'(t) > e^t$ for arbitrarily large values of t. Therefore there is a sequence of numbers $t_1, t_2, \cdots$ ($t_n \longrightarrow +\infty$ as $n \longrightarrow \infty$) such that $u'(t_n) = \exp(t_n)$. There exists a positive integer N such that

$$u(t_n+1) \geq e_2\{A(t_n+1)\} > (e^{t_n})^C = u'(t_n)^C$$

for $n \geq N$. This completes the proof of Lemma 3.

3. Theorems. We can now state and prove the theorems alluded to in the last paragraph of the introduction. The first of these is the following.

THEOREM 3. *Consider any equation of the form*

$$P(t, u(t), u'(t+1)) = 0. \tag{7}$$

There exists a positive number A, which depends only on the polynomial P, with the following property: to each proper solution $u(t)$ of (7) there corresponds a sequence $t_1, t_2, \cdots$ ($t_n \longrightarrow +\infty$ as $n \longrightarrow \infty$) such that

$$|u(t)| < e_2(At) \tag{4}$$

for $t = t_n$ $(n = 1, 2, 3, \cdots)$. That is, if $u(t)$ is a proper solution of (7) then there *is no number* $T > 0$ *for which* $|u(t)| \geq e_2(At)$ *for all* $t \geq T$.

Proof. Equation (7) may be written in the form

$$\sum_{i=0}^{I} \sum_{j=0}^{J} \sum_{k=0}^{K} T_{ijk} = 0 \tag{27}$$

where

$$T_{ijk} = a_{ijk} t^i u(t)^j u'(t+1)^k.$$

The a_{ijk} are real numbers independent of t. Among the terms T_{ijk} there is one term T_{pqr} selected in the following way:

(1) Choose $r = K$.

(2) Choose q to be the greatest of the values of j among all the terms T_{ijr}.

(3) Choose p to be the greatest of the values of i among all the terms T_{iqr}.

The term T_{pqr} so defined will be called the *principal term.*

Except for constant factors, the ratios T_{ijk}/T_{pqr} are of the following three possible types (excluding the ratio T_{pqr}/T_{pqr}):

$$\left\{\frac{t^{r_0} u(t)^{r_1}}{u'(t+1)}\right\}^{r-k} \tag{a}$$

where r_0 and r_1 are rational numbers and $r > k$.

$$\left\{\frac{t^{r_2}}{u(t)}\right\}^{q-j} \tag{b}$$

where r_2 is a rational number and $q > j$.

$$t^{i-p} \tag{c}$$

where $p > i$. Let R be the least non-negative number which is greater than or equal to the maximum value of r_1 for all ratios of type (a). Let A be any positive number such that $e^A > R$.

Now suppose that $u(t)$ is a proper solution of (7) and that $u(t) \geq e_2(At)$ for $t \geq T$. Choose B so that $R < B < e^A$. It follows from Lemma 1 that there exists a sequence $\{t_n\}$ for which $t_n \longrightarrow +\infty$ as $n \longrightarrow \infty$ and for which $u'(t_n+1) > u(t_n)^B$. For each value $t = t_n$, the function $u(t)$ satisfies not only equation

(7), but also the equation

$$(28)\qquad \sum_{i=0}^{I}\sum_{j=0}^{J}\sum_{k=0}^{K}\frac{T_{ijk}}{T_{pqr}}=0.$$

Since $u(t) \geq e_2(At)$, all ratios of types (b) or (c) approach zero as $t_n \longrightarrow +\infty$. Each ratio of type (a) is bounded by

$$\left\{\frac{t^{r_0}u(t)^R}{u(t)^B}\right\}^{r-k}$$

when $t = t_n$, for appropriate values of r_0 and k. Since $B > R$ and $r > k$, each such ratio approaches zero on the sequence $\{t_n\}$. It now follows that we may find a positive integer N such that the sum of all ratios T_{ijk}/T_{pqr} is less than one in absolute value when $t = t_N$, whereas $T_{pqr}/T_{pqr} = 1$. Thus (28) cannot be satisfied at the point t_N. This contradiction shows that a proper solution $u(t)$ of (7) cannot satisfy $u(t) \geq e_2(At)$ for all $t \geq T$.

Moreover, a proper solution $u(t)$ of (7) cannot satisfy $u(t) \leq -e_2(At)$ for all $t \geq T$. For if it could, the function $U(t) = -u(t)$ would satisfy $U(t) \geq e_2(At)$ for $t \geq T$ and would be a proper solution of an equation of the type (7). We have just shown that this is impossible. Since a proper solution is continuous, this completes the proof of Theorem 3.

The following theorem gives a much stronger result than does Theorem 3, but for a smaller class of equations.

THEOREM 4. *Let $u(t)$ be a non-decreasing or non-increasing proper solution of an equation of the form*

$$(29)\qquad \sum_{i=0}^{I} a_{iLK}t^i u(t)^L u'(t+1)^K + \sum_{i,j,k} a_{ijk}t^i u(t)^j u'(t+1)^k = 0,$$

wherein the a_{ijk} are constants and the latter summation is a triple summation over the ranges $i = 0, 1, \cdots, I$; $j = 0, 1, \cdots, J$; $k = 0, 1, \cdots, K-1$. (L may be greater than J, equal to J, or less than J.) Then there exists a number $A > 0$, which depends only on the form of (29), and there exists a number $T > 0$, which depends on (29) and on $u(t)$, such that

$$(4)\qquad |u(t)| < e_2(At)$$

for all $t \geq T$.

Proof. The method used in the proof of Theorem 3 for selecting the principal term T_{pqr} leads to the choice $p = I$, $q = L$, $r = K$ for the equation (29). Except for constant factors, the ratios T_{ijk}/T_{pqr} are of the following two possible types (excluding the ratio T_{pqr}/T_{pqr}):

(a)
$$\left\{\frac{t^{r_0} u(t)^{r_1}}{u'(t+1)}\right\}^{K-k}$$

where r_0 and r_1 are rational and $K > k$.

(b)
$$t^{i-I}$$

where $I > i$. Define R, A, and B as in the proof of Theorem 3. Let C be any positive number for which $C/2$ is larger than the maximum value of r_0 for all ratios of type (a).

Now suppose that $u(t)$ is a proper, non-decreasing solution of (29) for which $u(t) \geq e_2(At)$ for a sequence $\{\tau_n\}$ of values of t for which $\tau_n \longrightarrow +\infty$ as $n \longrightarrow \infty$. It follows from Lemma 1 that there exists a sequence $\{t_n\}$ for which $t_n \longrightarrow +\infty$ and for which $u'(t_n + 1) > t_n^C u(t_n)^B$. For each value $t = t_n$, the function $u(t)$ satisfies not only equation (29), but also the equation (28) obtained by dividing by the principal term. But for $t = t_n$ all ratios of type (b) approach zero as $n \longrightarrow \infty$. Each ratio of type (a) is bounded by

$$\left\{\frac{t^{C/2} u(t)^R}{t^C u(t)^B}\right\}^{K-k}.$$

Since $B > R$ and $K > k$, and since $u(t_n) \longrightarrow +\infty$ as $t_n \longrightarrow +\infty$, each such ratio approaches zero. We thus obtain the same contradiction as in the proof of Theorem 3. No such solution $u(t)$ can exist. Therefore to each proper non-decreasing solution $u(t)$ there corresponds a $T > 0$ such that $|u(t)| < e_2(At)$ for all $t \geq T$.

If a proper, non-increasing solution $u(t)$ exists for which $u(t) \leq -e_2(At)$ for $t = \tau_n$ $(n = 1, 2, \cdots)$, where $\tau_n \longrightarrow +\infty$ as $n \longrightarrow \infty$, we define $U(t) = -u(t)$, and obtain the same contradiction. Therefore to each proper, non-increasing solution $u(t)$ there corresponds a $T > 0$ such that $|u(t)| < e_2(At)$ for all $t \geq T$. This completes the proof of Theorem 4.

Our next theorem is as follows.

THEOREM 5. *Consider any equation of the form*

$$P(t, u(t), u'(t), u(t+1)) = 0. \tag{8}$$

There exists a positive number A, which depends only on the polynomial P, with the following property: to each proper non-decreasing or non-increasing solution $u(t)$ of (8) there corresponds a sequence $t_1, t_2, \cdots$ ($t_n \longrightarrow +\infty$ as $n \longrightarrow \infty$) such that

$$|u(t)| < e_2(At) \tag{4}$$

for $t = t_n$ ($n = 1, 2, \cdots$). That is, if $u(t)$ is any proper non-decreasing or non-increasing solution of (8), there is no number $T > 0$ for which $|u(t)| \geq e_2(At)$ for all $t \geq T$.

Proof. Equation (8) may be written in the form

$$\sum_{h=0}^{H} \sum_{i=0}^{I} \sum_{j=0}^{J} \sum_{k=0}^{K} T_{hijk} = 0$$

where

$$T_{hijk} = a_{hijk} t^h u(t)^i u'(t)^j u(t+1)^k.$$

The a_{hijk} are real numbers independent of t. We select a principal term T_{pqrs} in the following way. Let S be the set of all terms T_{hijk}. Let S_1 be the subset of S consisting of those terms for which $k = K$. Let M_1 be the maximum value of $i + j$ for all terms in S_1. Let S_2 be the set consisting of those terms of S_1 for which $i + j = M_1$. Let M_2 be the maximum value of j for all terms in S_2. Let S_3 be the set containing those terms of S_2 for which $j = M_2$. Let M_3 be the maximum value of h for all terms in S_3. There is a unique term in S_3 for which $h = M_3$. This term will be called the principal term. We shall use the symbol T_{pqrs} for it.

Except for constant factors, the ratios T_{hijk}/T_{pqrs} are of the following possible types (excluding the ratio T_{pqrs}/T_{pqrs}):

$$\left\{\frac{t^{r_0} u(t)^{r_1} u'(t)^{r_2}}{u(t+1)}\right\}^{s-k} \tag{a}$$

where r_0, r_1, and r_2 are rational numbers and $s > k$.

(b) $$t^{h-p}u(t)^{i-q}u'(t)^{j-r}$$

where $q + r > i + j$. Since i, j, q, and r are integers, terms of type (b) fall into one of the following two sub-classes.

(1) $$\frac{t^{h-p}u(t)^m}{u'(t)^{m+n}}$$

where m is an integer, $h - p$ is an integer, n is a positive integer, and $m + n$ is a non-negative integer.

(2) $$\frac{t^{h-p}u'(t)^m}{u(t)^{m+n}} = \left\{\frac{t^{r_3}u'(t)}{u(t)^{1+r_4}}\right\}^m$$

where m and n are positive integers, $h - p$ is an integer, r_3 is a rational number, and r_4 is a positive rational number.

(c) $$\left\{\frac{t^{r_5}u(t)}{u'(t)}\right\}^{r-j}$$

where r_5 is a rational number and $r > j$.

(d) $$t^{h-p}$$

where $p > h$.

Let R_0 be the maximum value of r_0 for all ratios of type (a). Let R_1' be the maximum value of r_1 for all ratios of type (a), and let $R_1 = \max(0, R_1')$. Let R_2' be the maximum value of r_2 for all ratios of type (a), and let $R_2 = \max(0, R_2')$. Let A be any number such that $e^A > R_1 + R_2$. Select any numbers B and C for which $B > R_1$, $C > R_2$, and $B + C < e^A$. Let R_3 be the maximum value of r_3 and let M be the maximum value of m for all ratios of type (b2). Let R_4 be the minimum value of r_4 for all ratios of type (b2). Let ϵ be any positive number less than $R_4/2$. Let R_5' be the maximum value of r_5 for all ratios of type (c), and let $R_5 = \max(0, R_5')$. Select any number D for which $D > R_5$.

Now assume that there exists a proper, non-decreasing solution $n(t)$ of (8) for which $u(t) \geq e_2(At)$ for all $t \geq t_0$. By Lemma 2 there exists a sequence $\{t_n\}$ such that (20) and (21) are satisfied. For each value $t = t_n$, $u(t)$ satisfies not only equation (8), but also the equation

$$\sum_{h=0}^{H} \sum_{i=0}^{I} \sum_{j=0}^{J} \sum_{k=0}^{K} \frac{T_{hijk}}{T_{pqrs}} = 0.$$

Now since $u(t) \geq e_2(At)$ and $u'(t) \geq t^D u(t)$ when $t = t_n$, all ratios of types (b1), (c), and (d) approach zero as $t_n \longrightarrow +\infty$. Also, each ratio of type (a) is, according to (20), bounded by

$$\left\{ \frac{t^{R_0} u(t)^{R_1} u'(t)^{R_2}}{u(t)^B u'(t)^C} \right\}^{s-k}$$

when $t = t_n$; and each ratio of type (b2) is, according to (21), bounded by

$$\left\{ \frac{t^{R_3} u'(t)}{u(t)^{1+2\epsilon}} \right\}^M < \left\{ \frac{t^{R_3} u'(t)}{u'(t) u(t)^{\epsilon}} \right\}^M$$

when $t = t_n$. Since $B > R_1$ and $C > R_2$, all these ratios tend to zero as $t_n \longrightarrow +\infty$. This conclusion yields a contradiction, just as in the proofs of the earlier theorems. Therefore no such solution $u(t)$ can exist.

The assumption that a proper, non-increasing solution $u(t)$ satisfies $u(t) \leq -e_2(At)$ for all $t \geq t_0$ may be shown to lead to a contradiction by defining $U(t) = -u(t)$.

The conclusion stated in Theorem 5 follows.

Our final theorem is the following.

THEOREM 6. *Consider any equation of the form*

$$P(t, u'(t), u(t+1)) = 0. \tag{9}$$

There exists a positive number A, which depends only on the polynomial P, with the following property: to each proper solution $u(t)$ of (9) there corresponds a sequence $t_1, t_2, \cdots$ ($t_n \longrightarrow +\infty$ as $n \longrightarrow \infty$) such that

$$|u(t)| < e_2(At) \tag{4}$$

for $t = t_n$ $(n = 1, 2, \cdots)$. That is, if $u(t)$ is any proper solution of (9), there is no positive number T for which $|u(t)| \geq e_2(At)$ for all $t \geq T$.

Proof. Equation (9) may be written in the form (27), where

$$T_{ijk} = a_{ijk} t^i u'(t)^j u(t+1)^k.$$

The principal term T_{pqr} is selected as follows:

(1) $r = K$;

(2) q is the greatest of the values of j among all terms T_{ijr};

(3) p is the greatest of the values of i among all terms T_{iqr}.

By using Lemma 3, the proof of Theorem 6 may now be completed in much the same way as before. We omit the details.

REFERENCES

1. E. Borel, *Mémoire sur les séries divergentes*, Annales de l'Ecole Normale Superieure, 16 (1899), 9-136.

2. G. H. Hardy, *Orders of infinity*, Cambridge Tract No. 12, Cambridge University Press, 1910.

3. O. E. Lancaster, *Some results concerning the behavior at infinity of real continuous solutions of algebraic difference equations*, Bull. Amer. Math. Soc., 46 (1940), 169-177.

4. E. Lindelöf, *Sur la croissance des intégrales des équations différentielles algébriques du premier ordre*, Bull. Soc. Math. France, 27 (1899), 205-215.

5. S. M. Shah, *On real continuous solutions of algebraic difference equations*, Bull. Amer. Math. Soc., 53 (1947), 548-558.

6. ———, *On real continuous solutions of algebraic difference equations*, II, Proc. National Inst. Sci. India 16 (1950), 11-17.

7. T. Vijayaraghavan, N. M. Basu, and S. N. Bose, *A simple example for a theorem of Vijayaraghavan*, Jour. London Math. Soc., 12 (1937), 250-252.

WASHINGTON STATE COLLEGE

LINEAR FUNCTIONAL EQUATIONS AND INTERPOLATION SERIES

PHILIP DAVIS

1. Introduction. The question of obtaining complete sets of solutions for a given linear partial differential equation is of the greatest interest from the theoretical as well as from the computational point of view. For constructing such sets, several methods of considerable generality have been proposed. Thus, for instance, Bergman [3] has introduced an integral operator which provides a means for the generation of complete sets when the differential equation is of the second or the fourth order. Extensions may be made to higher orders. By means of Bergman's operator, the space of analytic functions of a single complex variable is mapped upon the space of solutions of the given differential equation, and the process yields a generalization of the operator Re in the case of harmonic functions.

Complete sets of solutions may also be found by a method which is analogous to Runge's method of approximation in the theory of analytic functions. A description of this may be found in [6, p. 282]. This scheme has the practical drawback of requiring a knowledge of a fundamental singularity for the differential equation, a function which is known explicitly for but few differential equations.

In the present paper, we adopt a different point of view and study possible representations of solutions of linear functional equations of a certain class, and the generation of complete sets of such solutions by means of generalized interpolation series. By this is meant a biorthogonal series of the form

$$(1) \qquad f \sim \sum_{n=0}^{\infty} L_n(f)\phi_n; \quad L_m(\phi_n) = \delta_{mn}.$$

Here $\{L_n\}$ is a sequence of linear functionals. When each L_n is a point or a linear differential operator, then the series (1) reduces to a classical interpolation series. Our method is, essentially, to reduce the problem of the solution of the linear functional equation to a problem involving a denumerable infinity

Received May 7, 1953. This work was performed under a National Bureau of Standards contract with American University, and was sponsored by the Wright Air Development Center, ARDC, USAF.

Pacific J. Math. 4 (1954), 503-532

of interpolatory conditions. An interpolatory procedure then yields an operator which may be cast into integral form, and which maps an appropriate space of functions onto a subspace of solutions.

In order to carry out this method with ease, it is convenient to deal with Hilbert spaces H of functions, H being supposed to possess a reproducing kernel [cf. **5**, **1**], to restrict our basic functional equations to those possessing certain boundedness properties with respect to H, and to consider only solutions which lie in H. These assumptions will cause no difficulty in many instances where the existence and regularity of solutions may be known beforehand from independent considerations. Our work, therefore, falls mainly within the region of representation theory.

It is our principal aim to construct interpolation series which converge in preassigned regions to solutions of linear functional equations, and, by way of corollary, to construct complete systems of solutions. This is carried out in §§2-4. In §§5 and 6 we discuss some related topics, while in the final sections we take up the problem of systems of equations. The work is applicable to linear differential equations, both ordinary and partial, in an arbitrary number of variables, or of systems of such equations.

2. Reduction to an interpolatory problem. For the sake of definiteness, but realizing that restrictions other than the ones about to be set forth may prove useful in other circumstances, we shall deal with n complex variables

$$z_j = x_j + iy_j \qquad (j = 1, \cdots, n),$$

and shall designate by B a fixed $2n$-dimensional region in the space $Z = (z_1, \cdots, z_n)$ of the n complex variables. We shall designate by $L^2(B)$ the class of functions f which are single-valued analytic functions of z, are regular in B, and are such that

$$(2) \qquad \|f\|^2 = \int_B |f|^2 \, d\omega < \infty; \quad d\omega = dx_1 \cdots dx_n \, dy_1 \cdots dy_n .$$

It may sometimes prove expedient to introduce a weight function in (2). By L, we shall designate a fixed linear operator defined on $L^2(B)$ and with the property that $L(f)$, $f \in L^2(B)$, is regular analytic in B. Additional conditions on L will be required below. We shall be concerned with representations of solutions of class $L^2(B)$ of the functional equation

$$(3) \qquad L(f) = 0.$$

A principal application will be the case in which L is a partial differential operator of the kth order:

$$(4) \qquad L(f) = \sum_{i_1+i_2+\cdots+i_n=k} a_{i_1,i_2,\cdots,i_n} \frac{\partial^k f}{\partial z_1^{i_1}\partial z_2^{i_2}\cdots\partial z_n^{i_n}} + \cdots,$$

where the $+\cdots$ in (4) indicates the presence of partial derivatives of order $< k$. If now all the coefficients in (4) are regular in B, then so also will $L(f)$ be regular in B. It is to be remarked that the case $n = 1$ which leads in (4) to an ordinary differential equation is not excluded.

Let $\{L_n\}$ $(n = 0, 1 \cdots)$ be a sequence of linear functionals each of which is defined over the set R of functions which are regular in B and which possess the following two additional properties:

(a) The set $\{L_n\}$ is complete[1] for R; that is, if $f \in R$ and $L_n(f) = 0$ $(n = 0, 1, \cdots)$, then $f \equiv 0$.

(b) Each linear functional $\widetilde{L}_n \equiv L_n(L)$ is bounded over $L^2(B)$; that is, for each k, there exists a positive constant M_k such that

$$(5) \qquad |\widetilde{L}_k(f)| < M_k \, ||f||$$

for all $f \in L^2(B)$.

In connection with (b), let us observe that the composite operator

$$\widetilde{L}_n(f) \equiv L_n(L(f))$$

is a linear functional from $L^2(B)$ onto the complex numbers.

Example. Let L be the differential operator (4) with coefficients regular in B. Set

$$(6) \qquad L_k(f) = \left. \frac{\partial^{m_1+m_2+\cdots+m_n} f}{\partial z_1^{m_1}\partial z_2^{m_2}\cdots\partial z_n^{m_n}} \right|_{z_j = z_j^*},$$

where $k = k(m_1, m_2, \cdots, m_n)$ refers to a fixed indexing of the n-tuples of non-negative integers $m_1, m_2, \cdots, m_n$, and where the point $Z^* = (z_1^*, z_2^*, \cdots, z_n^*)$ is interior to B. It is clear that condition (a) holds for the selection (6). Let us

[1] Banach [2, p. 42] calls such sets *total*.

next examine $\widetilde{L}_n$, which is (6) acting on (4). In this case $\widetilde{L}_n(f)$ is a finite linear combination with constant coefficients of mixed partial derivatives of f evaluated at $Z = Z^*$. To show that condition (b) is satisfied, it suffices to show that any linear functional of the form (6), with Z^* interior to B, is bounded over $L^2(B)$. This is a consequence of the fact that the functionals (6) have a representation as Cauchy integrals with the path of integration lying in B, and hence are applicable term by term to any series of analytic functions which converges uniformly in a neighborhood of Z^*. Now let

$$\phi_m(Z) = \phi_m(z_1, z_2, \cdots, z_n)$$

be a complete orthonormal system for $L^2(B)$. Each $f \in L^2(B)$ possesses a Fourier expansion

$$f(Z) = \sum_{n=0}^{\infty} a_n \phi_n(Z); \quad \sum_{n=0}^{\infty} |a_n|^2 = \|f\|^2 < \infty, \tag{7}$$

convergent uniformly in every closed bounded subregion of B. Hence

$$L_k(f) = \sum_{n=0}^{\infty} a_n L_k(\phi_n(Z)), \tag{8}$$

the series (8) converging for all selections a_n with $\sum_{n=0}^{\infty} |a_n|^2 < \infty$. By a lemma of Landau, this implies that

$$\sum_{n=0}^{\infty} |L_k(\phi_n(Z))|^2 < \infty, \tag{9}$$

and the Schwarz inequality applied to (8) yields

$$|L_k(f)|^2 \leq \|f\|^2 \sum_{n=0}^{\infty} |L_k(\phi_n)|^2. \tag{10}$$

This establishes (b).

For analytic functions of a single complex variable, complete sets of functionals $\{L_n\}$ of a wide variety are known. We can, for instance, replace (6) (in the 1-dimensional case) by

$$L_k(f) = f(Z_k); \quad \lim_{k \to \infty} Z_k = Z^*; \; Z_k, Z^* \text{ interior to } B. \tag{11}$$

If we are dealing with differential equations with nonconstant coefficients, such a selection may reduce the complexity of the subsequent formal work. The points Z_k need not have an accumulation point interior to B as is suggested by (11), but, as in the Blaschke theory for the unit circle, may only have a weak accumulation of points on the boundary. In the theory of analytic functions of several complex variables, questions of the completeness of linear functionals are largely uninvestigated. However, certain sets in addition to (6) are known. Thus, for example, we may select the set (11) with the added restriction that the points Z_k do not lie on an analytic hypersurface [**10**, p.39].

The functionals (6) and (11) are the usual "point" functionals met in interpolatory function theory. However, complete sets of integral functionals usually associated with orthogonal expansions may also be employed here. Within an L^2 theory, for complex analytic functions the distinction between these two types is weak, and persists only in certain discussions [8].

Under the foregoing hypotheses, we have the following result.

THEOREM 1. *The linear functional equation $L(u) = 0$ possesses a nontrivial solution of class $L^2(B)$ if and only if the set of functionals $\{\widetilde{L}_n\}$ is incomplete for $L^2(B)$.*

Proof. Suppose first that $\{\widetilde{L}_k\}$ is incomplete. Then there exists an $f \in L^2(B)$ which does not vanish identically and such that $\widetilde{L}_k(f) = 0$ $(k = 0, 1, \cdots)$. That is, $L_k(L(f)) = 0$ for $k = 0, 1, \cdots$. By hypothesis, $L(f)$ is regular in B. Since therefore $\{L_k\}$ is complete for the class of regular functions in B, we must have $L(f) \equiv 0$. Conversely, let these begin a nontrivial $f \in L^2(B)$ such that $L(f) \equiv 0$ in B. Then

$$\widetilde{L}_k(f) = L_k(L(f)) = 0,$$

so that the incompleteness of $\{\widetilde{L}_k\}$ follows.

In this way, the consideration of the functional equation $L(f) = 0$ may be reduced to a consideration of the denumerable infinity of interpolation conditions of the type[2]

$$\widetilde{L}_k(f) = 0 \qquad (k = 0, 1, \cdots).$$

It is frequently of importance to be able to solve this equation subject to the auxiliary conditions

[2] Interpolation problems of this type, where the functionals involved are point functionals have been considered in Bergman, Mémorial Sciences Math., vol. 106, pp. 46-48.

$$A_n(f) = 0 \qquad (n = 0, 1, 2, \cdots), \tag{12}$$

where A_n designates a linear functional which we shall again assume is bounded over $L^2(B)$. Let now $\{\hat{L}_n\}$ be an augmented set of linear functionals which includes both the sets $\{\tilde{L}_n\}$ and $\{A_n\}$, but only these; that is, each $\hat{L}_n$ is either an $\tilde{L}_n$ or an A_n, while every $\tilde{L}_n$ and every A_n is some $\hat{L}_n$. We may now state the following result.

THEOREM 1′. *The linear functional equation* (3), *under the auxiliary conditions* (12), *possesses a nontrivial solution of class* $L^2(B)$ *if and only if the set of linear functionals* $\{\hat{L}_n\}$ *is incomplete for* $L^2(B)$.

Thus it appears that, from our present point of view, *the role played by the auxiliary conditions* (12) *is indistinguishable from that of the functional equation itself.* In the notation used later, the circumflex ^ will indicate the presence of auxiliary conditions; that is, we deal with the equation (3) and derive from it a set of functionals $\{\tilde{L}_n\}$, but when auxiliary conditions are present, the set $\{\tilde{L}_n\}$ will be augmented to yield $\{\hat{L}_n\}$. It should also be observed that, in eigenvalue problems, the operator L may involve a parameter λ. In such cases, the functionals $\tilde{L}_n$ and $\hat{L}_n$ will also involve this parameter.

3. Representation of solutions. We reproduce here, for convenience of reference, the following theorem on double orthogonality which was established in a previous paper [**12**].

THEOREM 2. *Let* $\{\tilde{L}_k\}$ *be a set of linear functionals each of which is defined and bounded over* $L^2(B)$. *The set* $\{\tilde{L}_k\}$ *will be assumed independent. There then exists a set of functions* $\{\phi_k^*(Z)\}$ $(k = 0, 1, \cdots)$ *and a set of linear functionals* $\{\tilde{L}_k^*\}$ $(k = 0, 1, \cdots)$ *which possess the following properties*:

(a) *Each* ϕ_k^* *is of class* $L^2(B)$ *and the set is orthonormal over* B:

$$\int_B \phi_i^* \overline{\phi_k^*}\, d\omega = \delta_{ik}. \tag{13}$$

(b) *Each* $\tilde{L}_k^*$ *is a finite linear combination of the functionals* $\tilde{L}_k$:

$$\tilde{L}_k^* = \sum_{p=0}^{k} a_{kp} \tilde{L}_p \qquad (k = 0, 1, \cdots) \tag{14}$$

for an appropriate set of constants a_{kp}.

(c) *The sets* $\{\widetilde{L}_k^*\}$ *and* $\{\phi_k^*\}$ *are biorthonormal:*

$$\widetilde{L}_i^*(\phi_k^*) = \delta_{ik}. \tag{15}$$

(d) *For all* $f \in L^2(B)$, *we have*

$$\widetilde{L}_k^*(f) = \int_B f\,\overline{\phi_k^*}\,d\omega. \tag{16}$$

(e) *The functions* ϕ_k^* *may be obtained by taking the set*[2]

$$\phi_n(Z) = \widetilde{L}_{n,\overline{w}} K_B(Z, \overline{W}) \qquad (n = 0, 1, \cdots) \tag{17}$$

and orthonormalizing them by the Gram-Schmidt process.

(f) *The set* $\{\phi_k^*\}$ *is complete for* $L^2(B)$ *if and only if the set* $\{\widetilde{L}_k\}$ *is complete for* $L^2(B)$.

In (17),

$$K_B(Z; \overline{W}) = K_B(z_1, z_2, \cdots, z_n;\ \overline{w}_1, \overline{w}_2, \cdots, \overline{w}_n)$$

designates the Bergman kernel function for the domain B, and in our notation an asterisk * used with the symbols for either functions or functionals indicates that the corresponding set of functions or functionals is orthonormal. Starting from a given B and a given ordered set $\{\widetilde{L}_k\}$, the sets $\{\widetilde{L}_k^*\}$ and $\{\phi_k^*\}$ are determined uniquely, and we shall speak of them as being the biorthogonal sets associated with $\{\widetilde{L}_k\}$ and B.

The inner products

$$(\phi_i, \phi_j) = \int_B \phi_i\,\overline{\phi_j}\,d\omega, \tag{18}$$

which occur in the orthonormalizing process, may be easily evaluated in terms of $K_B(Z; \overline{W})$. We have, from (16), (17), and the orthonormal expansion for K_B,

$$(\phi_i, \phi_j) = L_{j,z}\,[L_{i,\overline{w}} K_B(Z; \overline{W})]. \tag{19}$$

If we introduce the determinants

$$D_n = |(\phi_i, \phi_j)| \qquad (i, j = 0, 1, \cdots, n), \tag{20}$$

[3] The notation $L_{n,\overline{w}}$ means that L_n is to be applied to K_B as a function of $\overline{w}$.

then we have

$$(21) \qquad \phi_n^*(Z) = (D_{n-1} D_n)^{-\frac{1}{2}} \cdot \begin{vmatrix} (\phi_0, \phi_0), & \cdots, & (\phi_0, \phi_n) \\ \cdot & & \\ \cdot & & \\ \cdot & & \\ (\phi_{n-1}, \phi_0), & \cdots, & (\phi_{n-1}, \phi_n) \\ \phi_0(Z), & \cdots, & \phi_n(Z) \end{vmatrix}$$

while

$$(22) \qquad a_{ni} = (-1)^i (D_{n-1} D_n)^{-\frac{1}{2}}$$

$$\cdot \begin{vmatrix} (\phi_0, \phi_0), & \cdots, (\phi_0, \phi_{i-1}), & (\phi_0, \phi_{i+1}), & \cdots, & (\phi_0, \phi_n) \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ (\phi_{n-1}, \phi_0), & \cdots, (\phi_{n-1}, \phi_{i-1}), & (\phi_{n-1}, \phi_{i+1}), & \cdots, & (\phi_{n-1}, \phi_n) \end{vmatrix}.$$

In view of the orthonormality of the functions ϕ_k^*, we may form the kernel function

$$(23) \qquad K_I(Z; \overline{W}) = \sum_{k=0}^{\infty} \phi_k^*(Z) \overline{\phi_k^*(W)},$$

the series (23) converging uniformly and absolutely for Z and W confined to any closed bounded subset of $B \times B$ and defining there an analytic function of $z_1, z_2, \cdots, z_n$ and an anti-analytic function of $w_1, w_2, \cdots, w_n$. If the system $\{\phi_k^*\}$ is complete for $L^2(B)$, then K_I must coincide with K_B. If auxiliary conditions are present, we replace (23) by

$$(23') \qquad K_I(Z; \overline{W}) = \sum_{k=0}^{\infty} \hat{\phi}_k^*(Z) \overline{[\hat{\phi}_k^*(W)]},$$

where $\{\hat{\phi}_k^*\}$ and $\{\hat{L}_k^*\}$ are the biorthonormal sets associated with $\{\hat{L}_k^*\}$ and B.

Combining this observation with Theorem 1, we have the following result:

THEOREM 2. *The functional equation* (3) (*augmented, possibly, by* *auxiliary conditions* (12)) *possesses nontrivial solutions of class* $L^2(B)$ *if and only if* $K_I \not\equiv K_B$, *or if and only if* $K_I(Z;\overline{Z}) < K_B(Z;\overline{Z})$, $Z \in B$.

If we admit the possibility of a nontrivial solution, the kernel K_I may be thought of as an "incomplete" kernel for B relative to the functional equation (3). It is wholly accessible to computation via (19)-(23) once K_B has been established. Moreover, K_I may be used to project the space $L^2(B)$ onto the linear subspace S of solutions:

THEOREM 3. *The function* $g(Z)$ *is a solution of* (3) *of class* $L^2(B)$ *if and only if there exists an* $f \in L^2(B)$ *for which*

$$T(f) \equiv g(Z) = f(Z) - \int_B K_I(Z, \overline{W})\, f(W)\, d\omega_W, \tag{24}$$

or alternately, for which

$$T(f) \equiv g(Z) = f(Z) - \sum_{k=0}^{\infty} \widetilde{L}_k^*(f)\, \phi_k^*(Z). \tag{24'}$$

Appropriate changes must be made if auxiliary conditions (12) *are present.*

Proof. We observe that in view of (16) and (23), (24) and (24′) are equivalent. For a given $f \in L^2(B)$, construct a g by means of (24′). Since the quantities $\widetilde{L}_n^*(f)$ are Fourier coefficients of f, the sum in (24′) is of class $L^2(B)$. Thus also $g \in L^2(B)$. As remarked previously, g will be a solution of (3) if $\widetilde{L}_k(g) = 0$ $(k = 0, 1, \cdots)$. By (14), this is equivalent to $\widetilde{L}_k^*(g) = 0$ $(k = 0, 1, \cdots)$. In view of the boundedness of $\widetilde{L}_k^*$ over $L^2(B)$, we have

$$\widetilde{L}_k^*(g) = \widetilde{L}_k^*(f) - \sum_{n=0}^{\infty} \widetilde{L}_n^*(f)\, \widetilde{L}_k^*(\phi_n^*) = \widetilde{L}_k^*(f) - \widetilde{L}_k^*(f) = 0. \tag{25}$$

The last equality follows from (15). Thus g is a solution. Conversely, if g is a solution of class $L^2(B)$ we shall have

$$\widetilde{L}_k^*(g) = L_k^*(L(g)) = 0 \qquad (k = 0, 1, \cdots),$$

so that (24′) holds with $f \equiv g$.

Equation (24)-(24′) yields a projection of $L^2(B)$ onto the subspace S of solutions. The partial sums of (24′),

$$\sum_{k=0}^{N} \widetilde{L}_k^*(f)\, \phi_k^*(Z),$$

have the usual minimum property of Fourier series; that is, for each N they solve the minimization of the integral

$$(26) \qquad I_N = \int_B \left| f - \sum_{k=0}^{N} a_k \phi_k(Z) \right|^2 d\omega.$$

On the other hand, the series in (24′) has two characters: *it is simultaneously a Fourier series and an interpolation series as well.* This means that the partial sum

$$(27) \qquad S_N = \sum_{k=0}^{N} \widetilde{L}_k^*(f)\, \phi_k^*(Z)$$

is that linear combination of $\phi_0, \phi_1, \cdots, \phi_N$ which interpolates to f in the sense that

$$(28) \qquad \widetilde{L}_k^*(S_N(Z)) = \widetilde{L}_k^*(f) \qquad (k = 0, 1, \cdots, N),$$

or, equivalently,

$$(28') \qquad \widetilde{L}_k(S_N(z)) = \widetilde{L}_k(f) \qquad (k = 0, 1, \cdots, N),$$

Once K_B is known, and if $\{L_k\}$ is a sequence of point or differential operators, then no integrals extended over B of the inner product type need actually be computed to obtain either ϕ_n^* or the series expansion in (24′). The explicit orthogonalization formulas of Gram-Schmidt (20)-(22) are equivalent to an interpolation series of Newton type with respect to the sequence $\{\widetilde{L}_n\}$. For a fixed N, formulas of the Lagrange type may be developed, and may prove to be more convenient.

In view of the reproducing property of K_B, we may write (24) in the form

$$(29) \qquad T(f) = g(Z) = \int_B K_B(Z;\overline{W}) f(W)\, d\omega_W - \int_B K_I(Z;\overline{W}) f(W)\, d\omega_W,$$

so that by introducing the kernel

$$K_S(Z;\overline{W}) = K_B(Z;\overline{W}) - K_I(Z;\overline{W}) \tag{30}$$

we have the representation

$$g(Z) = \int_B K_S(Z;\overline{W}) f(W)\, d\omega_W . \tag{31}$$

If $f \in S$, then $g(Z) \equiv f(Z)$ inasmuch as

$$\widetilde{L}_n^*(f) = L_n^*(L(f)) = 0 \qquad (n = 0, 1, \cdots).$$

Thus, $K_S(Z;\overline{W})$ is a reproducing kernel for the subspace S and as such, may be proved unique (that is, nondependent upon the selection $\{L_n\}$) in the usual way. $K_S(Z,\overline{W})$ may also be defined by

$$K_S(Z;\overline{W}) = \sum_k \psi_k(Z)\overline{\psi_k(W)}, \tag{32}$$

where $\{\psi_k\}$ is any orthonormal set which is complete for S. In the case of ordinary differential equations, the sum in (32) will consist of a finite number of terms. In the case of ordinary differential equations of infinite order or of partial differential equations, there will, in general, be an infinity of terms present.

The incomplete kernel K_I may be identified as the kernel of the orthogonal complement $S^{\perp}$ of S, and the utility of the backward decomposition of K_B given by (30) lies in relative accessibility of K_I as opposed to K_S. Let us note also the orthogonality relationship

$$\int_B K_S(Z;\overline{W}) K_I(W;\overline{X})\, d\omega_W = 0, \tag{33}$$

which follows from (30) and from the reproducing properties of K_S and K_B over S and $L^2(B)$, respectively.

For $f \in S$, we have, by (31) and the Schwarz inequality,

$$\begin{aligned} |f|^2 < \|f\|^2 \int_B K_S(Z;\overline{W})\, K_S(\overline{Z};W)\, d\omega_W \\ = \|f\|^2 [K_B(Z;\overline{Z}) - K_I(Z;\overline{Z})]. \end{aligned} \tag{34}$$

The inequality (34) is a wide generalization of the Schwarz Lemma for functions regular in the unit circle. Finally, we may generate complete sets of solutions in the following way:

THEOREM 4. *Let $\theta_n(Z)$ $(n = 0, 1, \cdots)$ be a complete set for $L^2(B)$; then the functions*

$$\psi_n(Z) = T(\theta_n) = \int_B K_S(Z; \overline{W})\, \theta_n(W)\, d\omega_W \tag{35}$$

$$= \theta_n(Z) - \sum_{k=0}^{\infty} \widetilde{L}_k^*(\theta_n) \phi_k^*(Z) \qquad (n = 0, 1, \cdots)$$

form a complete set of solutions.

Proof. We must show that any solution g can be approximated arbitrarily closely by combinations of $\psi_0, \cdots, \psi_n, \cdots$. Let

$$\left\| g - \sum_{k=0}^{N} a_k \theta_k \right\| < \epsilon.$$

Then by (34), (35), we have

$$\left| g - \sum_{k=0}^{N} a_k \psi_k \right| < \epsilon [K_B(Z; \overline{Z}) - K_I(Z; \overline{Z})]^{1/2},$$

which establishes completeness.

4. The nonhomogeneous case. We consider next the nonhomogeneous linear functional equation

$$L(u) = f, \tag{36}$$

which may be supplemented by auxiliary conditions of the form

$$A_k(u) = \alpha_k \qquad (k = 0, 1, \cdots). \tag{37}$$

We assume that f is regular in B, and that all previous hypotheses regarding L and A_n remain in force. We first reduce (36)-(37) to a problem in interpolation in the following way.

THEOREM 5. *The linear functional equation* (36), *subject to the auxiliary* *conditions* (37), *is equivalent to the interpolation problem*

$$\widetilde{L}_k(u) = L_k(f) \tag{38}$$

$$A_k(u) = \alpha_k \qquad (k = 0, 1, \cdots). \tag{38'}$$

If conditions (37) *are absent we may omit* (38′).

Proof. That (38), (38′) follow from (36), (37) is evident. Suppose conversely that (38) holds. We wish to prove that $L(u) \equiv f$ throughout B. We have

$$L_k\, L(u) - L_k(f) = 0 \qquad (k = 0, 1, \cdots).$$

Thus

$$L_k[L(u) - f] = 0 \qquad (k = 0, 1, \cdots).$$

Now $L(u) - f$ is regular in B, and $\{L_k\}$ is complete for R. Hence the conclusion follows.

It will now be convenient to uniformize our notation. We introduce an augmented set $\{\hat{L}_k\}$ of linear functionals as in the previous paragraph, and introduce a set of constants $\{\beta_k\}$ by means of the definition

$$\begin{aligned} \beta_k &= L_k(f) && \text{if } \hat{L}_k = \widetilde{L}_k, \\ \beta_k &= \alpha_k && \text{if } \hat{L}_k = A_k. \end{aligned} \tag{39a}$$

The interpolation problem is now

$$\hat{L}_k(u) = \beta_k \qquad (k = 0, 1, \cdots). \tag{40}$$

We observe again that *there is no distinct rôle played by the auxiliary conditions*. Boundary value and initial value problems of mathematical physics may be fitted into the pattern (40) providing it is known a priori that the required solutions are regular across the boundary so that the functionals $\hat{L}_k$ will have the required boundedness properties. We next introduce the biorthonormal sets $\{\hat{L}_k^*\}$ and $\{\hat{\phi}_k^*\}$ associated with $\{\hat{L}_k\}$ and B. We have

$$\hat{L}_k^* = \sum_{p=0}^{k} a_{kp}\, \hat{L}_p \tag{41}$$

for constants a_{kp} determined as in the previous paragraph. The following result now holds.

THEOREM 6. *The linear functional equation* (36), (37) *possesses a solution of class* $L^2(B)$ *if and only if*

$$\sum_{k=0}^{\infty} \left| \sum_{p=0}^{k} a_{kp} \beta_p \right|^2 < \infty. \tag{42}$$

The solution is unique (*within* $L^2(B)$) *if and only if* $\{\hat{L}_k\}$ *is complete for* $L^2(B)$. *If* (42) *holds, then the series*

$$u(Z) = \sum_{k=0}^{\infty} \left(\sum_{p=0}^{k} a_{kp} \beta_p \right) \hat{\phi}_k^*(Z) \tag{43}$$

converges to a solution $u(Z)$ *uniformly and absolutely in every closed bounded subset of* B.

Proof. Suppose that a solution $u \in L^2(B)$ exists. Then from (40) and (41) we have

$$\hat{L}_k^*(u) = \sum_{p=0}^{k} a_{kp} \beta_p .$$

But since the $\hat{L}_k^*(u)$ are Fourier coefficients of u with respect to $\{\phi_k^*\}$, we must have (42). If (42) holds, then the series (43) converges uniformly and absolutely in every closed bounded subregion of B to a function $V(Z)$ of class $L^2(B)$. Now,

$$\hat{L}_k^*(V) = \sum_{p=0}^{k} a_{kp} \beta_p$$

in view of the boundedness and biorthogonality properties of these functionals. Hence

$$\hat{L}_k(V) = \beta_k \qquad (k = 0, 1, \cdots).$$

so that V satisfies the equations (36) and (37) by Theorem 5.

A particularly important special case is to solve (36) subject to the auxiliary conditions

$$A_k(u) = 0 \qquad (k = 0, 1, \cdots). \tag{44}$$

As in Theorem 5, we again construct the biorthonormal sets $\{\hat{L}_k^*\}$ and $\{\hat{\phi}_k^*\}$, and note that each functional $\hat{L}_k^*$ is a finite linear combination of functionals $\widetilde{L}_k$ and A_k:

$$\hat{L}_k^* = \sum_{p=0}^{k} a_{kp} \hat{L}_p = \sum_p b_{kp} \widetilde{L}_p + \sum_p c_{kp} A_p, \tag{45}$$

where the coefficients b_{kp} and c_{kp} now contain certain dummy zeros. Let us write

$$\sum_p b_{kp} \widetilde{L}_p = \sum_p b_{kp} L_p L = S_k L;\ S_k = \sum_p b_{kp} L_p. \tag{46}$$

We now have the following result.

THEOREM 6′. *The linear functional equation* (36), (44) *possesses a solution of class* $L^2(B)$ *if and only if*

$$\sum_{k=0}^{\infty} |S_k(f)|^2 < \infty. \tag{47}$$

If (47) *holds, the interpolation series*

$$u(Z) = \sum_{k=0}^{\infty} S_k(f)\, \hat{\phi}_k^*(Z) \tag{48}$$

converges to a solution uniformly and absolutely in every closed bounded subset of B.

Proof. If $\beta_k = \alpha_k = 0$ when $\hat{L}_k = A_k$, then, by (39a),

$$\sum_p a_{kp} \beta_p = \sum_p b_{kp} L_p(f) = S_k(f).$$

Under the assumption that the equation (36), (44) possesses a solution for all $f \in L^2(B)$, we may find a second representation for the interpolation series (48). The functionals S_k are bounded over $L^2(B)$, and hence possess a Riesz representative $s_k(Z)$:

$$(49) \qquad S_k(f) = \int_B f\, \overline{s_k}\, d\omega; \quad f \in L^2(B),$$

where

$$(50) \qquad s_k(Z) = S_{k,\overline{w}}\, K_B(Z; \overline{W}).$$

If (36), (44) possess a solution of class $L^2(B)$ for all $f \in L^2(B)$, then (47) must hold for all $f \in L^2(B)$. In particular, from

$$(51) \qquad \overline{s_k(Z_0)} = S_{k,w}\, K_B(\overline{Z}_0; W), \quad Z_0 \in B,$$

we learn that

$$(52) \qquad \sum_{k=0}^{\infty} |s_k(Z_0)|^2 < \infty \qquad \text{for all } Z_0 \in B.$$

We may therefore form the mixed kernel

$$(53) \qquad D(Z; \overline{W}) = \sum_{k=0}^{\infty} \hat{\phi}_k^*(Z)\, \overline{s_k(W)},$$

which will converge uniformly and absolutely in every closed bounded subregion of $B \times B$. Finally, from (48) and (49), we have the representation

$$(54) \qquad U(Z) = \int_B D(Z; \overline{W})\, f(W)\, d\omega_W.$$

The kernel $D(Z; \overline{W})$ plays a role analogous to a Green's function or to the Duhamel kernel in the superposition theorem of the theory or ordinary linear differential equations. The totality of solutions in $L^2(B)$ of (36), (44) may be written in the form

$$(55) \qquad U(Z) = \int_B D(Z; \overline{W})\, f(W)\, d\omega_W + \int_B K_s(Z; \overline{W})\, h(W)\, d\omega_W; \; h \in L^2(B),$$

or, in interpolatory form,

$$U(Z) = \sum_{k=0}^{\infty} S_k(f)\,\hat{\phi}_k^*(Z) + h(Z) - \sum_{k=0}^{\infty} \hat{L}_k^*(h)\,\hat{\phi}_k^*(Z);\; h \in L^2(B). \tag{56}$$

5. Convergence of interpolation series for $f \notin L^2(B)$. In the present paragraph we return to the interpolation series (24′). This has been discussed under the hypothesis that $f \in L^2(B)$. If, however, each functional $\widetilde{L}_k^*$ (or $\widetilde{L}_k$) is applicable to a wider class of functions than $L^2(B)$, a formal series (24′) may be constructed and its properties examined for f in this wider class. This will be the case, for example, when L_k are differential operators. For the sake of definiteness, let us assume that we are dealing with the ordinary linear differential equation

$$L(f) \equiv f^{(n)} + a_1(z)\,f^{(n-1)} + \cdots + a_n(z)f = 0, \tag{57}$$

and that we have selected

$$L_k(f) = f^{(k)}(0) \qquad (k = 0, 1, \cdots). \tag{58}$$

The coefficients $a_j(z)$ in (57) are assumed regular in a region R containing the origin. If f is regular at $z = 0$, then the series (24′) may be formed. If this series then converges uniformly in a neighborhood of $z = 0$, the difference

$$g(z) = f(z) - \sum_{k=0}^{\infty} \widetilde{L}_k^*(f)\,\phi_k^*(z),$$

which is again regular at $z = 0$, will be a solution; for, since the functionals $\widetilde{L}_k^*$ are applicable term by term, we have

$$\widetilde{L}_k^*(g) = \widetilde{L}_k^*(f) - \sum_{p=0}^{\infty} \widetilde{L}_p^*(f)\,\widetilde{L}_k^*(\phi_p^*) = \widetilde{L}_k^*(f) - \widetilde{L}_k^*(f) = 0, \quad (k = 0, 1, \cdots),$$

and this implies that $L(g) = 0$. The interpolation series (24′) has a doubly orthogonal character, but the above proof will apply to any interpolation series

$$g(z) = f(z) - \sum_{k=0}^{\infty} \widetilde{L}_k^*(f)\,\psi_k(z) \tag{59}$$

in which the regular functions ψ_k are merely biorthogonal:

(60) $$\widetilde{L}_j^*(\psi_k) = \delta_{jk}\,.$$

Such sets are more numerous than doubly orthogonal sets. To determine such a set, we need only start from a given set of functions $\{t_n(z)\}$ $(n = 0, 1, \cdots)$ which has properties of independence with respect to $\{\widetilde{L}_j\}$ and determine linear combinations

(61) $$\psi_k(z) = \sum_{p=0}^{k} e_{kp}\, t_p(z) \qquad (k = 0, 1, \cdots)$$

successively by the requirement (60).

We shall now prove that we may find a set $\{\psi_k(z)\}$ biorthogonal to $\{\widetilde{L}_k^*\}$ with the property that if f is regular in any neighborhood of $z = 0$, the interpolation series

(62) $$g(z) = f(z) - \sum_{k=0}^{\infty} \widetilde{L}_k^*(f)\,\psi_k(z)$$

will converge to a solution of (57) in some neighborhood of $z = 0$. The present proof will generalize to both partial differential equations and to ordinary differential equations of infinite order.

We have, from (57) and (58),

(63) $$\widetilde{L}_k(f) = \sum_{p=0}^{n} \frac{d^k}{dz^k}\,[a_p(z) f^{(n-p)}(z)]\Big|_{z=0} = \sum_{j=0}^{n+k} b_{kj}\, f^{(j)}(0),$$

while

(64) $$\widetilde{L}_k^*(f) = \sum_{j=0}^{n+k} b_{kj}^*\, f^{(j)}(0)$$

for appropriate $b_{j,k}^*$. We assume that B contains the origin and is contained in the region of regularity of $a_i(z)$.

LEMMA. *Let $f(z)$ be regular in $|z| \le \rho$. Then there exist positive constants M and t such that*

(65) $$|\widetilde{L}_k^*(f)| < Mt^k \qquad (k = 0, 1, \cdots).$$

Proof. For any $g \in L^2(B)$, we have

$$\widetilde{L}_k^*(g) = \iint_B g\, \overline{\phi}_k^*\, dx\, dy \qquad (k = 0, 1, \cdots).$$

Thus $\widetilde{L}_k^*(g)$ are Fourier coefficients of g, so that, by the Bessel inequality,

$$\sum_{k=0}^{\infty} |\widetilde{L}_k^*(g)|^2 < \iint_B |g|^2\, dx\, dy. \tag{66}$$

In particular, we may select

$$g = z^p/p! \qquad (p = 0, 1, \cdots),$$

so that

$$\widetilde{L}_k^*(z^p/p!) = \sum_{j=0}^{n+k} b_{kj}^*\, [z^p/p!]^{(j)}\Big|_{z=0} = b_{kp}^*. \tag{67}$$

From (66) we obtain

$$\sum_{k=0}^{\infty} |\widetilde{L}_k^*(z^p/p!)|^2 = \sum_{k=0}^{\infty} |b_{kp}^*|^2 \le \iint_B \left|\frac{z^p}{p!}\right|^2 dx\, dy \tag{68}$$

$$\le \frac{\text{Area } (B)}{(p!)^2}\, d^{2p}, \qquad (p = 0, 1, \cdots),$$

where d designates the maximum distance from the boundary of B to the origin. If now f is regular in $|z| \le \rho$, we have, for some constant M^*,

$$|f^{(j)}(0)| < M^*\, j!/\rho^j \qquad (j = 0, 1, \cdots), \tag{69}$$

so that from (64) and (68),

$$|\widetilde{L}_k^*(f)| < M(d/\rho)^k \qquad (k = 0, 1, \cdots), \tag{70}$$

with

$$M = M^*\, (\text{Area } (B))^{1/2}\, (d/\rho)^{n+1}/(d/\rho) - 1.$$

LEMMA. *There exist positive constants m and σ such that*

$$|L(\phi_k^*(z))| < m|z|^k \tag{71}$$

for all $k = 0, 1, \cdots$, and for all $|z| \le \sigma$.

Proof. The orthonormal functions $\phi_k^*(z)$ satisfy the requirements [**12**, p. 16]

$$\widetilde{L}_0(\phi_k^*) = \widetilde{L}_1(\phi_k^*) = \cdots = \widetilde{L}_{k-1}(\phi_k^*) = 0,$$

or with notation $g_k = L\phi_k^*$,

$$g_k(0) = g_k'(0) = \cdots = g_k^{(k-1)}(0) = 0.$$

Let $|z| = \sigma'$ and $|z| = \sigma$, $\sigma' > \sigma$ both be contained in B. Since ϕ_n^* are orthonormal over B, they are uniformly bounded by some M over $|z| \le \sigma'$; hence, by (57) and Cauchy's inequality,

$$|L(\phi_k^*)| \le \sigma' M \sum_{j=0}^{n} B_j\, j!\, s^{j+1} = m; \ |z| \le \sigma, \tag{72}$$

where

$$s = \sigma' - \sigma \quad \text{and} \quad B_j = \max_{|z| \le \sigma'} |a_j(z)|.$$

Thus the functions $L(\phi_k^*)$ are uniformly bounded in $|z| \le \sigma$ by m. The inequality (71) now follows from Schwarz's lemma.

We observe now that the last two lemmas imply that the series

$$\sum_{k=0}^{\infty} \widetilde{L}_k^*(f) L\phi_k^*(z)$$

will converge absolutely and uniformly in $|z| \le r$, $r < 1/t$. Furthermore, we must have

$$L(f) = \sum_{k=0}^{\infty} \widetilde{L}_k^*(f)\, L\phi_k^*(z); \ |z| \le r. \tag{73}$$

To show this, designate the sum of (73), $|z| \le r$, by $g(z)$. By uniform convergence, we may apply L_p term by term. Thus

$$L_p(g) = \sum_{k=0}^{\infty} \widetilde{L}_k^*(f) L_p L\phi_k^*(z), \tag{74}$$

so that

$$L_p^*(g) = \sum_{k=0}^{\infty} \widetilde{L}_k^*(f) \widetilde{L}_p^*(\phi_k^*(z)) = \widetilde{L}_p^*(f) = L_p^* L(f). \tag{75}$$

By the completeness of $\{L_p^*\}$, $g \equiv L(f)$.

Let now B_1 designate a region containing $z = 0$ and contained in $|z| < r$, and let $D(z, \overline{w}) \equiv D_{B_1}(z, \overline{w})$ be the kernel described in (54)-(55). We have, for each f regular in $|z| \leq r$, the identity

$$f(z) = \iint_{B_1} D(z, \overline{w})\, L(f(w))\, d\omega_w + s(z), \tag{76}$$

where $s(z)$ is some solution of $L(s) = 0$, regular in B_1. Applying this inversion operator to (73), we have

$$\begin{aligned} f(z) - s(z) &= \sum_{k=0}^{\infty} \widetilde{L}_k^*(f) \iint_{B_1} D(z, \overline{w})\, L(\phi_k^*(w))\, d\omega_w \\ &= \sum_{k=0}^{\infty} \widetilde{L}_k^*(f)\, \psi_k(z), \end{aligned} \tag{77}$$

where

$$\psi_k(z) = \iint_{B_1} D(z, \overline{w})\, L(\phi_k^*(w))\, d\omega_w \qquad (k = 0, 1, \cdots). \tag{78}$$

The functions $\{\psi_k\}$ are easily seen to be biorthonormal to the interpolation operators $\{\widetilde{L}_k^*\}$. We therefore have the following result.

THEOREM 7. *For each $f(z)$ regular in $|z| \leq \sigma$, the biorthogonal interpolation series*

$$g(z) = f(z) - \sum_{k=0}^{\infty} \widetilde{L}_k^*(f)\psi_k(z) \tag{79}$$

converges to a solution of the equation (57).

6. Relation to questions of stability. In a previous paragraph we have given necessary and sufficient conditions in order that a given functional equation possess solutions of class $L^2(B)$. If the coefficients of this equation involve a parameter λ, then a criterion may be obtained in terms of λ. Here B designates any region which possesses a kernel function K_B. If B is chosen as an unbounded domain, then membership in $L^2(B)$ acts as a stability criterion.

To elucidate this remark, let us consider the two dimensional case, and let S designate the half-strip

$$\operatorname{Re}(z) \geq 0, \quad |\operatorname{Im}(z)| \leq h.$$

Then we have $f \in L^2(S)$ if and only if[4]

$$\|f\|_S^2 = \int_{-h}^{h} \int_0^{\infty} |f(x+iy)|^2 \, dx\, dy < \infty. \tag{80}$$

Thus, to belong to $L^2(S)$ a function must not become large too rapidly as z approaches the horizontal boundaries of the strip, and indeed, must approach zero with a certain maximal rapidity along any horizontal line.

LEMMA. *Let* $f \in L^2(S)$; *then along each line*

$$y = \sigma, \quad -h < \sigma < h,$$

we must have

$$\lim_{x \to +\infty} f(x + i\sigma) = 0. \tag{81}$$

Proof. If (81) were not true, we could find two positive quantities A and δ and a sequence of values $\lambda_0 < \lambda_1 < \cdots$ such that

$$\lambda_n - \lambda_{n-1} \geq \delta > 0 \qquad (n = 1, 2, \cdots), \tag{82}$$

and

[4] Various authors have considered solutions of class $L^2(0,\infty)$; for example, see [**13**].

$$|f(\lambda_n + i\sigma)| \geq A > 0 \qquad (n = 0, 1, \cdots). \tag{83}$$

In virtue of (82) we may find an $r > 0$ such that the circles

$$C_n : |z - (\lambda_n + i\sigma)| \leq r$$

lie in S and do not overlap. Now

$$\infty > \iint_S |f|^2\, dx\, dy > \sum_{n=0}^{\infty} \iint_{C_n} |f|^2\, dx\, dy. \tag{84}$$

Since f is regular in C_n, it possesses a Taylor series expansion

$$f(z) = f(P_n) + f'(P_n)(z - P_n) + \cdots ; \; P_n \equiv \lambda_n + i\sigma, \tag{85}$$

so that

$$\iint_{C_n} |f(z)|^2\, dx\, dy > \pi r^2 \; |f(P_N)|^2. \tag{86}$$

Combining this with (84), we must have

$$\sum_{n=0}^{\infty} |f(P_n)|^2 < \infty. \tag{87}$$

This contradicts (83) and proves the result.

The 'stability' which is spoken of here is that usually associated with the theory of linear, non time-varying electrical networks; in this theory we are confronted with a differential equation

$$y^{(n)} + a_1 y^{(n-1)} + \cdots + a_n y = f(x), \tag{88}$$

to be solved under initial conditions such as

$$y(0) = y'(0) = \cdots \; y^{(n-1)}(0) = 0. \tag{89}$$

If the characteristic roots of (88) are

$$\tau_j = u_j + i v_j \qquad (j = 1, 2, \cdots, n),$$

assumed distinct, then the n independent solutions of the homogeneous equation are

$$y_i(x) = e^{u_j x} e^{i v_j x}.$$

The equation (88) is called stable if $u_j < 0$ $(j = 1, \cdots, n)$. We observe now that $y_j(z) \in L^2(S)$ if and only if $u_j < 0$. For

$$|y_j(z)|^2 = e^{2(u_j x - v_j y)},$$

and this result is now implied by (80). It appears then that the equation (88) *is stable if and only if the fundamental solutions of the homogeneous equation are in* $L^2(S)$. For the case of more general linear networks, we propose membership in $L^2(S)$ as a possible extension of this type of stability. Added flexibility may be achieved by varying h, and by attaching a weighting function to (80). Inasmuch as the mapping function for S is elementary, the kernel function K_S of S may be computed explicitly, and the criterion of § 3 can be formulated in closed form.

7. Systems of functional equations. The methods of the previous paragraphs may be extended to the case of systems of equations. As the proofs and the principal results parallel those given in § 2-4 very closely, we shall not dwell on these aspects, and shall be content merely with showing how the generalization may be set up.

For the sake of simplicity, we consider here only systems of two functional equations in the two unknown functions, $u_i = u_i(z_1, z_2, \cdots, z_n) = u_i(Z)$, $(i = 1, 2)$,

$$L^1(u_1, u_2) = 0, \tag{90}$$

$$L^2(u_1, u_2) = 0.$$

Introducing the solution vector $\mathbf{u} = (u_1, u_2)$ and the vector operator $\mathbf{L} = (L_1, L_2)$, we may write (90) as

$$\mathbf{L}(\mathbf{u}) = 0. \tag{90$'$}$$

We assume that $L^i(u_1, u_2)$ are regular functions of $z_1, \cdots, z_n$ whenever u_i are, and that $\mathbf{L}$ is linear on the vector $\mathbf{u}$. We shall say that (90′) possesses a solution of class $L^2(B)$ if there exists $u_i \in L^2(B)$ for which (90′) holds. In addition to (90), we may consider an augmented system comprising (90) plus certain auxiliary conditions which may be written in the form

$$(91) \qquad \mathbf{A}_n(\mathbf{u}) = A_n(u_1, u_2) = 0 \qquad (n = 0, 1, \cdots).$$

Here $\mathbf{A}_n$ is a linear functional on $\mathbf{u}$. Let again $\{L_n\}$ designate a fixed set of linear functionals defined on the set of functions regular on B and complete for this set. We introduce

$$(92) \qquad \widetilde{\mathbf{L}}_{2n}(\mathbf{u}) = L_n L^1(u_1, u_2) \qquad (n = 0, 1, \cdots),$$

$$\widetilde{\mathbf{L}}_{2n+1}(\mathbf{u}) = L_n L^2(u_1, u_2) \qquad (n = 0, 1, \cdots).$$

$\widetilde{\mathbf{L}}_{2n}$ and $\widetilde{\mathbf{L}}_{2n+1}$ are linear functionals defined over vectors $\mathbf{u}$, and we shall say that a sequence $\widetilde{\mathbf{L}}_n$ of such functionals is complete for a class S of vectors if $\widetilde{\mathbf{L}}_n(\mathbf{u}) = 0$ $(n = 0, 1, \cdots)$ implies $\mathbf{u} = 0$. We have the following parallel to Theorem 1.

THEOREM 8. *The system* (90′) *possesses a nontrivial solution of class* $L^2(B)$ *if and only if the set of functionals* $\{\widetilde{\mathbf{L}}_n\}$ *is incomplete for* $L^2(B)$. *If auxiliary conditions* (91) *are present, the set* $\{\widetilde{\mathbf{L}}_n\}$ *must be augmented by the addition of* $\{\mathbf{A}_n\}$.

It is now convenient to introduce the direct sum of $L^2(B)$ with itself:

$$L_2^2(B) = L^2(B) \oplus L^2(B).$$

This space consists of pairs

$$\mathbf{u} = (u_1, u_2), \; u_i \in L^2(B).$$

Vector addition and scalar multiplication are defined by

$$\mathbf{u} + \mathbf{v} = (u_1, u_2) + (v_1, v_2) = (u_1 + v_1, \; u_2 + v_2)$$

and

$$a\,\mathbf{u} = a(u_1, u_2) = (a\,u_1, a\,u_2).$$

We introduce an inner product in $L_2^2(B)$ by means of[5]

$$(93) \qquad \{\mathbf{u}, \mathbf{v}\} = \{(u_1, u_2), (v_1, v_2)\} = \int_B (u_1 \bar{v}_1 + u_2 \bar{v}_2)\, d\omega,$$

[5] In what follows, the parenthesis is used solely for the element pairs of $L_2^2(B)$. If the inner product in $L^2(B)$ is required, we shall write $(u, v)_{L^2(B)}$.

and a norm by

$$(94) \qquad ||\mathbf{u}||^2 = ||(u_1, u_2)||^2 = \{(u_1, u_2), (u_1, u_2)\}$$

$$\int_B (|u_1|^2 + |u_2|^2) d\omega = ||u_1||^2_{L^2(B)} + ||u_2||^2_{L^2(B)}.$$

Under this norm, it is known that $L_2^2(B)$ becomes a Hilbert Space. By classical results, any bounded linear functional $\mathbf{T}$ over $L_2^2(B)$ possesses a representation of the form

$$(95) \qquad \mathbf{T}(\mathbf{u}) = \{\mathbf{u}, \mathbf{v}\} = \int_B (u_1 \bar{v}_1 + u_2 \bar{v}_2) d\omega$$

for some $\mathbf{v} \in L_2^2(B)$. Hence we have the decomposition

$$(96) \qquad \mathbf{T}(\mathbf{u}) = T_1(u_1) + T_2(u_2),$$

where T_1 and T_2 are bounded linear functionals over $L^2(B)$. The converse evidently holds also. Moreover,

$$(97) \qquad ||\mathbf{T}||^2 = ||\mathbf{V}||^2 = ||v_1||^2 + ||v_2||^2 = ||T_1||^2 + ||T_2||^2.$$

In what follows, we shall assume that $\widetilde{\mathbf{L}}_n$ as well as $\mathbf{A}_n$ are linear and bounded over $L_2^2(B)$. The examples given in § 2 are easily extended to the present case.

If $\{\mathbf{u}^{(n)}\}$ is a complete orthonormal system for $L^2(B)$, it may be shown by an extension of the usual proofs that each of the series $\sum_{n=0}^{\infty} u_i^{(n)}(z)\overline{[u_j^{(n)}(w)]}$ converges uniformly and absolutely in every closed bounded subdomain of $B \times B$. In addition, a strong Riesz-Fischer theorem exists for $L_2^2(B)$; that is, $\mathbf{u} \in L_2^2(B)$ if and only if

$$\mathbf{u} = \sum_{n=0}^{\infty} a_n \mathbf{u}^{(n)} \text{ with } \sum_{n=0}^{\infty} |a_n|^2 < \infty$$

and

$$a_n = \{\mathbf{u}, \mathbf{u}^{(n)}\} \qquad (n = 0, \cdots).$$

The convergence of each of the component series is uniform and absolute in every closed bounded subdomain of B. The array

$$(98)\quad \mathcal{K}_B(Z;\overline{W}) = \begin{pmatrix} \sum_{n=0}^{\infty} u_1^{(n)}(z)\overline{u_1^{(n)}(w)} & \sum_{n=0}^{\infty} u_1^{(n)}(z)\overline{u_2^{(n)}(w)} \\ \sum_{n=0}^{\infty} u_2^{(n)}(z)\overline{u_1^{(n)}(w)} & \sum_{n=0}^{\infty} u_2^{(n)}(z)\overline{u_2^{(n)}(w)} \end{pmatrix}$$

will be known as a kernel tensor for the space $L_2^2(B)$. Each row of $\mathcal{K}_B$ is, for fixed $W \in B$, a vector element of $L_2^2(B)$ which we shall denote by $\mathcal{K}_B^1(Z,\overline{W})$ and $\mathcal{K}_B^{(2)}(Z,\overline{W})$. Thus,

$$(99)\quad \mathcal{K}_B(Z,\overline{W}) = \begin{pmatrix} \mathcal{K}_B^1(Z;\overline{W}) \\ \mathcal{K}_B^2(Z;\overline{W}) \end{pmatrix}.$$

If

$$\mathbf{u} = (u_1, u_2) \in L_2^2(B),$$

then we have

$$(100)\quad u_i = \{\mathcal{K}_B^i(Z,\overline{W}), \overline{\mathbf{u}(w)}\} \qquad (i = 1,2).$$

Let us consider, for example, the case $i = 1$; then

$$\mathbf{u} = \sum_{n=0}^{\infty} a_n(u_1^{(n)}, u_2^{(n)}),$$

so that

$$\{\mathcal{K}_B^i(Z;W), \overline{\mathbf{u}(W)}\} = \left\{\left(\sum_{n=0}^{\infty} u_1^{(n)}(Z)\overline{u_1^{(n)}(W)}, \sum_{n=0}^{\infty} u_1^{(n)}(Z)\overline{u_2^{(n)}(W)}\right),\right.$$

$$\left.\overline{\left(\sum_{n=0}^{\infty} a_n u_1^{(n)}(W), \sum_{n=0}^{\infty} a_n u_2^{(n)}(W)\right)}\right\}.$$

$$= \sum_{\substack{m=0\\n=0}}^{\infty} a_m u_1^{(n)}(Z) \int_B u_1^{(m)}(W)\overline{u_1^{(n)}(W)} + u_2^{(m)}(W)\overline{u_2^{(n)}(W)}\,d\omega,$$

which by orthonormality reduces to

$$\sum_{n=0}^{\infty} a_n u_1^{(n)}(Z) = u_1(Z).$$

The reproducing property may be written more compactly as

$$\mathbf{u}(Z) = \{\mathbf{u}(W), \mathcal{K}_B(W;\overline{Z})\}. \tag{100′}$$

It can be seen that if $\{u_n\}$ is a complete orthonormal system for $L^2(B)$, then the set of vectors

$$\mathbf{u}_{2n} = (u_n, 0) \qquad (n = 0, 1, \cdots), \tag{101}$$

$$\mathbf{u}_{2n+1} = (0, u_n) \qquad (n = 0, 1, \cdots),$$

is complete and orthonormal for $L_2^2(B)$. With this special selection, we find a kernel tensor of the form

$$\mathcal{K}_B(Z;W) = \begin{pmatrix} K_B(Z;\overline{W}) & 0 \\ 0 & K_B(Z,\overline{W}) \end{pmatrix}, \tag{102}$$

where K_B is the kernel for $L^2(B)$.

We come now to the analogue of Theorem 2.

THEOREM 9. *Let $\{\widetilde{\mathbf{L}}_k\}$ be a set of linear functionals each of which is defined and bounded over $L_2^2(B)$. The set $\{\widetilde{\mathbf{L}}_k\}$ will be assumed independent. Then there exists a set of pairs*

$$\varnothing_k^*(Z) = (\phi_{k,1}^*, \phi_{k,2}^*) \qquad (k = 0, 1, \cdots)$$

and a set of linear functionals $\{\widetilde{\mathbf{L}}_k^*\}$ $(n = 0, 1, \cdots)$ which possess the following properties:

(a) *Each $\varnothing_k^*$ is of class $L_2^2(B)$, and the set is orthonormal:*

$$\{\varnothing_k^*, \varnothing_j^*\} = \delta_{jk}. \tag{103}$$

(b) *Each $\widetilde{\mathbf{L}}_k^*$ is a finite linear combination of the functionals $\widetilde{\mathbf{L}}_k$:*

$$\tilde{\mathbf{L}}_k^* = \sum_{p=0}^{k} a_{kp} \tilde{\mathbf{L}}_p \qquad (k = 0, 1, \cdots) \tag{104}$$

for an appropriate set of constants a_{kp}.

(c) *The sets* $\{\tilde{\mathbf{L}}_k^*\}$ *and* $\{\varnothing_k^*\}$ *are biorthonormal:*

$$\tilde{\mathbf{L}}_j^*(\varnothing_k^*) = \delta_{jk}, \tag{105}$$

(d) *For all* $\mathbf{V} \in L_2^2(B)$ *we have*

$$\tilde{\mathbf{L}}_k^*(\mathbf{V}) = \{\mathbf{V}, \varnothing_k^*\}. \tag{106}$$

(e) *The pairs* $\varnothing_p^*$ *may be obtained by taking the set*

$$\phi_n(Z) = \tilde{\mathbf{L}}_{n,\overline{w}}\, \mathfrak{K}_B(Z, \overline{W}) \qquad (n = 0, 1, \cdots) \tag{107}$$

and orthonormalizing them by the Gram-Schmidt process.

(f) *The set* $\{\varnothing_k\}$ *(or* $\{\varnothing_k^*\}$*) is complete for* $L_2^2(B)$ *if and only if the set* $\{\tilde{\mathbf{L}}_k\}$ *is complete for* $L_2^2(B)$.

By (107) is meant that

$$\phi_{n,i}(Z) = \tilde{L}_{n,\overline{w}}\, \mathfrak{K}_B^{(i)}(Z; \overline{W}) \qquad (i = 1, 2) \tag{108}$$

where

$$\varnothing_n(Z) = (\phi_{n,1}(Z), \phi_{n,2}(Z))$$

Using the specific set of functionals (92), we construct the related biorthonormal sets $\{\tilde{\mathbf{L}}_k^*\}$ and $\{\varnothing_k^*\}$. We then have the following analogue of Theorem 3.

THEOREM 10. *The vector* $\mathbf{g}(z)$ *is a solution of the system* (90) *of class* $L_2^2(B)$ *if and only if there exists an* $\mathbf{f}(Z) \in L_2^2(B)$ *for which*

$$\mathbf{g}(Z) = \mathbf{f}(Z) - \sum_{k=0}^{\infty} \tilde{\mathbf{L}}_k^*(\mathbf{f}) \varnothing_k^*(Z). \tag{109}$$

For each $\mathbf{f} \in L_2^2(B)$, the series in (109) converges uniformly and absolutely in every closed bounded subdomain of B. It is simultaneously a Fourier series and an interpolation series whose terms may be obtained by interpolating to $\mathbf{f}$

by means of $\{\widetilde{\mathbf{L}}_k\}$.

References

1. N. Aronszajn, *Theory of reproducing kernels*, Trans. Amer. Math. Soc. (1950), 337-404.

2. S. Banach, *Théorie des operations linéaires*, Warsaw, 1933.

3. S. Bergman, *Zur Theorie der Funktionen, die eine lineare partielle Differentialgleichung befriedigen*, Recueil Mathematique, nouv. ser. 2 (44) (1937), 1169-1198.

4. ———, *Zur Theorie der Linearen Integral- und Funktionalgleichungen im Komplexen Gebiet*, Bulletin de l'Institute de Mathématiques et Mécanique à l'Université Koubycheff de Tomsk **1** (1937), 242-256.

5. ———, *The kernel function and conformal mapping*, Mathematical Surveys vol. 5, New York, 1950.

6. S. Bergman and M. Schiffer, *Kernel functions in mathematical physics*, New York, 1953.

7. P. Davis, *On the applicability of linear differential operators of infinite order to functions of class $L^2(B)$*, Amer. J. Math. 74 (1952), 475-491.

8. P. Davis and J. L. Walsh, *Representations and extensions of bounded linear functionals defined on classes of analytic functions*, Trans. Amer. Math. Soc. 76 (1954), 190-206.

9. P. Davis and H. Pollak, *Linear functionals and analytic continuation problems*, Pacific J. Math. 3 (1953), 47-72.

10. W. F. Osgood, *Funktionentheorie*, Bd. III, Leipzig-Berlin, 1927.

11. F. Riesz, *Les systèmes d'equations linéaires à une infinité d'inconnues*, Paris, 1913.

12. J. L. Walsh and P. Davis, *Interpolation and orthonormal systems*, J. Analyse Math. 2 (1952), 1-28.

13. A. Wintner, *A criterion for the non-existence of L^2 solutions of a non-oscillating differential equation*, London J. Math. 25 (1950), 347-351.

NATIONAL BUREAU OF STANDARDS
WASHINGTON, D.C.

SETS OF RADIAL CONTINUITY OF ANALYTIC FUNCTIONS

FRITZ HERZOG AND GEORGE PIRANIAN

1. Introduction. A point set E on the unit circle C ($|z| = 1$) will be called a *set of radial continuity* provided there exists a function $f(z)$, regular in the interior of C, with the property that $\lim_{r \to 1} f(re^{i\theta})$ exists if and only if $e^{i\theta}$ is a point of E. From Cauchy's criterion it follows that the set E of radial continuity of a function $f(z)$ is given by the formula

$$E = \prod_{k=1}^{\infty} \sum_{n=1}^{\infty} \prod \mathop{\mathrm{E}}_{e^{i\theta}} \left\{ |f(r_1 e^{i\theta}) - f(r_2 e^{i\theta})| \leq \frac{1}{k} \right\},$$

where the inner intersection on the right is taken over all pairs of real values r_1, r_2 with $1 - 1/n \leq r_1 < r_2 < 1$. From the continuity of analytic functions it thus follows that every set of radial continuity is a set of type $F_{\sigma\delta}$. The main purpose of the present note is to prove the following result.

THEOREM 1. *If E is a set of type F_σ on C, it is a set of radial continuity.*

The theorem will be proved by means of a refinement of a construction which was used by the authors in an earlier paper [**2**] to show that every set of type F_σ on C is the set of convergence of some Taylor series.

2. A special function. That the set consisting of all points of C is a set of radial continuity is trivial. In proving Theorem 1, it may therefore be assumed that the complement of E is not empty. In order to surmount difficulties one at a time, we begin with a new proof of the well-known fact that the empty set is a set of radial continuity (see [**1**, vol. 2, pp. 152-155]).

Let

$$f(z) \equiv \sum_{n=N}^{\infty} C_n(z),$$

where

Received July 28, 1952.

Pacific J. Math., 4 (1954), 533-538

$$
\begin{aligned}
C_n(z) \equiv \frac{z^{k_n}}{n^2}\Big\{ & 1 + z/\omega_{n1} + (z/\omega_{n1})^2 + \cdots + (z/\omega_{n1})^{n^2-1} \\
& + z^{n^2}[1 + z/\omega_{n2} + (z/\omega_{n2})^2 + \cdots + (z/\omega_{n2})^{n^2-1}] \\
& + \cdots \\
& + z^{(n-1)n^2}[1 + z/\omega_{nn} + (z/\omega_{nn})^2 + \cdots + (z/\omega_{nn})^{n^2-1}]\Big\};
\end{aligned} \tag{1}
$$

here

$$\omega_{nj} = e^{2\pi ij/n},$$

and $\{k_n\}$ is a sequence of nonnegative integers which increases rapidly enough so that no two of the polynomials $C_n(z)$ contain terms of like powers of z, and so that a certain other requirement is met; the positive integer N, which is the lower limit of the foregoing series, will be determined later.

If z is one of the points ω_{nj}, then $|C_n(z)| = 1$. On the other hand, let z lie on the unit circle, and let $\Gamma_n(z)$ be any sum of consecutive terms from (1). If z is different from each of the roots of unity ω_{nj} that enter into $\Gamma_n(z)$, and δ denotes the (positive) angular distance between z and the nearest of these ω_{nj}, then

$$|\Gamma_n(z)| < \frac{A_1}{\delta n^2}, \tag{2}$$

where A_1 is a universal constant (see [2, Lemma A]). Now, if

$$z = e^{i\theta}\omega_{nj}, \quad |\theta| < \frac{\pi}{n^2}, \tag{3}$$

and $R_{nj}(z)$ denotes the sum of the terms in the jth row of (1) (including the factor z^{k_n}/n^2), then

$$|R_{nj}(z)| = \frac{\sin(n^2\theta/2)}{n^2 \sin(\theta/2)} > A_2, \tag{4}$$

where A_2 is again a positive universal constant. But if the angular distance between z and ω_{nj} is less than π/n^2, the angular distances between z and the remaining nth roots of unity are all greater than $1/n$, and therefore (3) implies that, for sufficiently large n, by (2) and (4),

$$|C_n(z)| > A_2 - 2A_1/n > 5A_3,$$

where $A_3 = A_2/6$. We now choose N so large that the second of these inequalities holds whenever $n \geq N$.

Let $k_N = 0$; let r_N be a number $(0 < r_N < 1)$ such that

$$|C_N(re^{i\theta}) - C_N(e^{i\theta})| < \frac{A_3}{N!}$$

for $r_N \leq r \leq 1$ and all θ. Next, let k_{N+1} be large enough so that

$$|C_{N+1}(r_N e^{i\theta})| < \frac{A_3}{(N+1)!}$$

for all θ; and let r_{N+1} be greater than r_N, and near enough to 1 so that

$$|C_{N+1}(re^{i\theta}) - C_{N+1}(e^{i\theta})| < \frac{A_3}{(N+1)!}$$

for $r_{N+1} \leq r \leq 1$ and all θ. Let this construction be continued indefinitely.

Now let L be a line segment joining the origin to a point $e^{i\theta}$, and let n be an integer such that $n > N$ and

$$(5) \qquad |C_n(e^{i\theta})| > 5A_3.$$

We then write

$$f(r_n e^{i\theta}) - f(r_{n-1}e^{i\theta}) = C_n(e^{i\theta}) + [C_n(r_n e^{i\theta}) - C_n(e^{i\theta})] - C_n(r_{n-1}e^{i\theta})$$

$$+ \sum_{j=N}^{n-1} \{[C_j(r_n e^{i\theta}) - C_j(e^{i\theta})] - [C_j(r_{n-1}e^{i\theta}) - C_j(e^{i\theta})]\}$$

$$+ \sum_{j=n+1}^{\infty} \{C_j(r_n e^{i\theta}) - C_j(r_{n-1}e^{i\theta})\}$$

and obtain from the inequalities above

$$|f(r_n e^{i\theta}) - f(r_{n-1} e^{i\theta})| > A_3 \left[5 - \frac{1}{n!} - \frac{1}{n!} - 2 \sum_{j=N}^{n-1} \frac{1}{j!} - 2 \sum_{j=n+1}^{\infty} \frac{1}{j!} \right]$$

$$\geq A_3 [5 - 2(e-1)] > A_3 .$$

It follows that, if there exist infinitely many integers n for which (5) is satisfied $f(z)$ does not approach a finite limit as z approaches $e^{i\theta}$ along the line L. But for each real θ there exist infinitely many integers n with the property that, for some integer j_n,

$$\left| \frac{\theta}{2\pi} - \frac{j_n}{n} \right| < \frac{1}{2n^2}$$

(see [3, p. 48, Theorem 14]), so that each z on C admits infinitely many representations (3). It follows that $\lim_{r \to 1} f(re^{i\theta})$ does not exist for any value θ.

3. Closed sets of radial continuity. Let E be a closed set on C, and let G denote its (nonempty) complement. Again, let $f(z)$ be the function defined in § 2, except for the following modification. In the polynomial $C_n(z)$, let ω_{n1}, $\omega_{n2}, \cdots, \omega_{np_n}$ denote those nth roots of unity which lie in G and have the additional property that the angular distance of each one of them from E is greater than $n^{-1/2}$. The exponent of z in the factor outside of the brackets in the last row of the right member of (1) becomes $(p_n - 1)n^2$. And the p_n nth roots of unity ω_{nj} that occur in $C_n(z)$ must be so labelled that their arguments increase as the index j increases, with $\arg \omega_{n1} > 0$ and $\arg \omega_{np_n} \leq 2\pi$. Then every partial sum $\Gamma_n(z)$ of consecutive terms of $C_n(z)$ satisfies the inequality $|\Gamma_n(z)| < A_1 n^{-3/2}$ for all z belonging to E, and therefore the Taylor series of $f(z)$ converges on E. On the other hand, let the exponents k_n in (1) be chosen in a manner similar to that of § 2, and let L be a line segment joining the origin to a point $e^{i\theta}$ in the (open) set G. Then there exist infinitely many integers n for which (5) is satisfied by our newly constructed polynomials $C_n(z)$, and therefore $\lim_{r \to 1} f(re^{i\theta})$ does not exist.

4. The general case. Suppose finally that E is a set of type F_σ on C. Then the complement G of E is of type G_δ; that is, it can be represented as the intersection of open sets $G_1, G_2, \cdots$, with $G_k \supset G_{k+1}$ for all k. In turn, we can represent G_1 as the union of closed intervals I_{1h} in such a way that no two distinct intervals I_{1h} and $I_{1h'}$ contain common interior points, and in such a way that no point of G_1 is a limit point of end points of intervals I_{1h}. Similarly, each set G_k can be represented as the union of closed intervals I_{kh} satisfying similar restrictions.

Let n_0 be any positive integer. Since the denumerable set of all open arcs

$$z = e^{i\theta}, \ |\theta - 2\pi j/n| < \pi/n^2 \qquad (j = 1, 2, \cdots, n, \ \ n > n_0)$$

covers the entire unit circle, there exists a set of finitely many such arcs covering the unit circle. It follows that we can choose a finite number of terms $C_n(z)$ (see (1)), modified as in § 3, such that their sum $f_1(z)$ has the following properties:

i) for each θ in I_{11}, there exist two values ρ' and ρ'', $0 < \rho' < \rho'' < 1$, such that $|f_1(\rho' e^{i\theta}) - f_1(\rho'' e^{i\theta})| > A_3$;

ii) for each point $e^{i\theta}$ outside of I_{11} and outside of the two neighboring intervals I_{1h} and $I_{1h'}$, and for each n for which $C_n(z)$ occurs in $f_1(z)$, the modulus of any sum of consecutive terms of $C_n(e^{i\theta})$ is less than $A_1 \ n^{-3/2}$.

Next we accord a similar treatment to I_{12}, then to I_{21}, I_{13}, I_{22}, I_{31}, I_{14}, and so forth. The sum $f(z)$ of the polynomials $f_1(z)$, $f_2(z)$, $\cdots$ thus constructed has the following properties: if $e^{i\theta}$ lies in E, that is, lies in only finitely many of the intervals I_{kh}, the Taylor series of $f(z)$ converges at $z = e^{i\theta}$; if $e^{i\theta}$ lies in G, there exist pairs of values ρ' and ρ'' arbitrarily near to 1 and such that

$$|f(\rho' e^{i\theta}) - f(\rho'' e^{i\theta})| > A_3 .$$

It follows that E is the set of radial continuity of $f(z)$, and the proof of Theorem 1 is complete.

5. Sets of uniform radial continuity. The following theorem is analogous to Theorem 2 of [2].

THEOREM 2. *If E is a closed set on C, then there exists a function $f(z)$, regular in $|z| < 1$, such that $\lim_{r\to 1} f(re^{i\theta})$ exists uniformly with respect to all $e^{i\theta}$ in E and does not exist for any $e^{i\theta}$ not in E.*

For the proof of Theorem 2, we refer to the function $f(z)$, constructed in § 3. Note that $|\Gamma_n(z)| < A_1 n^{-3/2}$ for all z in E. Hence the Taylor series of $f(z)$ converges uniformly in E. It then follows easily, by the use of Abel's summation, that the convergence

$$\lim_{r\to 1} f(re^{i\theta}) = f(e^{i\theta})$$

is also uniform in E.

6. An unsolved problem. The converse of Theorem 1 is false, since a set of radial continuity can be the complement of a denumerable set which is dense on C. We do not know whether there exist sets of type $F_{\sigma\delta}$ that are not sets of radial continuity.

References

1. Ludwig Bieberbach, *Lehrbuch der Funktionentheorie*, Leipzig-Berlin, 1927.

2. F. Herzog and G. Piranian, *Sets of convergence of Taylor series* I, Duke Math. J., **16** (1949), 529-534.

3. Oskar Perron, *Die Lehre von den Kettenbrüchen*, Leipzig-Berlin, 1929.

MICHIGAN STATE COLLEGE
UNIVERSITY OF MICHIGAN

COMMENTS ON THE PRECEDING PAPER BY HERZOG AND PIRANIAN

P. C. Rosenbloom

1. Our main purpose here is to extract and formulate explicitly the general principle underlying the construction of Herzog and Piranian. The results in this note are implicitly contained in the computations on pp. 535 and 537 of their paper, and the full credit belongs to them.

2. We use the notation $M(r, f) = \max |f(z)|$ on $|z| = r$.

THEOREM 1. *Let f_n be analytic in $|z| \le 1$, let r_n be increasing, $0 < r_n \longrightarrow 1$ as $n \longrightarrow \infty$, let $a_n > 0$,*

$$A = \sum_{n=1}^{\infty} a_n < +\infty,$$

let $R(t) = \sum a_k$ over all k such that $r_k \ge t$, and let $g = \sum_{n=1}^{\infty} f_n$. If

(a) $M(r_n, f_{n+1}) \le a_n$,

and

(b) $M(1, f_n') \le a_n (1 - r_n)^{-1}$

for all n, then g is analytic in $|z| < 1$, and for $|z| \le 1$, $r_{n-1} \le r \le r_n$, we have

$$(1)\qquad \left| g(rz) - \sum_{1}^{n-1} f_k(z) - f_n(rz) \right| \le A(1-r)^{1/2} + R(1-(1-r)^{1/2}),$$

$$(2)\qquad |g(r_n z) - g(r_{n-1} z) - f_n(z)| \le 2A(1-r_{n-1})^{1/2} + 2R(1-(1-r_{n-1})^{1/2}) + R(r_n).$$

Proof. We have

Received April 26, 1954.
Pacific J. Math. 4 (1954), 539-543

$$|f_k(rz) - f_k(z)| \leq a_k(1-r)/(1-r_k)$$

$$\leq \begin{cases} a_k(1-r)^{1/2} & \text{if } r_k \leq 1-(1-r)^{1/2} \\ \\ a_k & \text{if } k \leq n-1, \end{cases}$$

and $|f_k(rz)| \leq a_{k-1}$ for $k > n$. Inequality (1) now follows from

$$g(rz) - \sum_{k=1}^{n-1} f_k(z) - f_n(rz) = \sum_{k=1}^{n-1} (f_k(rz) - f_k(z)) + \sum_{k=n+1}^{\infty} f_k(rz).$$

We now apply (1) with $r = r_n$ and $r = r_{n-1}$ to estimate

$$h(z) = g(r_n z) - g(r_{n-1} z) - f_n(r_n z) + f_n(r_{n-1} z),$$

and obtain (2) from

$$g(r_n z) - g(r_{n-1} z) - f_n(z) = h(z) - f_n(r_{n-1} z) + (f_n(r_n z) - f_n(z)).$$

3. **We denote by** $E(g)$ the set of radial continuity of g.

COROLLARY 1a. *If* $|z_0| = 1$, $\limsup_{n\to\infty} |f_n(z_0)| > 0$, *then* $z_0 \notin E(g)$.

COROLLARY 1b. *If* $|z_0| = 1$, *and* $\lim f_n(rz_0)$ *exists as* $r \longrightarrow 1$ *and* $n \longrightarrow \infty$ *simultaneously,*[1] *then*

$$\lim_{r\to\infty} g(rz_0) \text{ and } \sum_{n=1}^{\infty} f_n(z_0) = g(z_0)$$

either both exist or both do not exist. If $\lim f_n(rz_0) = 0$, *then*

$$\lim_{r\to 1} g(rz_0) = g(z_0)$$

if either exists. Hence if $M(1, f_n) \longrightarrow 0$ *as* $n \longrightarrow \infty$, *then* $E(g)$ *is the set of convergence of* $\sum_{n=1}^{\infty} f_n(z)$ *on* $|z| = 1$.

4. **We now establish:**

[1] The weaker condition that $f_n(rz_0)$ has a limit as $n \longrightarrow +\infty$ and $r \longrightarrow 1$ in such a way that $r_{n-1} \leq r \leq r_n$ for all n is sufficient for this corollary.

THEOREM 2. *If F_n is analytic in $|z| \le 1$, $M(1, F_n) \le M_n$, $M(1, F_n') \le M_n$ for all n, and $a_n > 0$ (all n), $\sum_{n=1}^{\infty} a_n < +\infty$, then there exist sequences r_n and k_n such that $f_n(z) = z^{k_n} F_n(z)$ satisfies* (a) *and* (b) *of Theorem* 1.

Proof. Let $k_1 = 0$ and suppose that $k_2, \cdots, k_n, r_1, \cdots, r_{n-1}$ are defined. Then (b) is satisfied if

$$r_n \ge 1 - \frac{a_n}{M_n(k_n+1)} .$$

Choose any r_n such that

$$1 > r_n > \max\left[r_{n-1}, 1 - \frac{a_n}{M_n(k_n+1)}\right].$$

Then (a) is satisfied if

$$k_{n+1} \ge \frac{\log(a_n/M_{n+1})}{\log r_n} .$$

5. As a consequence, we have:

COROLLARY 2a. *If*

$$\limsup_{n\to\infty} |\alpha_n| > 0, \ \limsup_{n\to\infty} k_n^{-1} \log|\alpha_n| = 0,$$

$$a_n > 0, \ \sum a_n < +\infty, \ \text{and} \ \frac{k_{n+1}}{k_n} \ge \frac{|\alpha_n|}{a_n} \log \frac{|\alpha_{n+1}|}{a_n}$$

for all n, then $E(g) = 0$, where $g(z) = \sum \alpha_n z^{k_n}$.

If $\alpha_n = O(1)$, $\limsup_{n\to\infty} |\alpha_n| > 0$, k_n increasing, and

$$\sum \frac{k_n}{k_{n+1}} \log \frac{k_{n+1}}{k_n} < +\infty,$$

then $E(g) = 0$.

COROLLARY 2b. *Suppose that f is analytic in the circle $|z| \le 1$, $f(1) = 1$, $M(1, f') \le 1$, and that $a_n > 0$ (all n),*

$$\sum_{n=1}^{\infty} a_n < +\infty.$$

Let

$$g(z) = \sum_{n=1}^{\infty} z^{k_n} f(ze^{-i\theta_n}).$$

If

$$\liminf_{n\to\infty} \left[\frac{k_{n+1}}{k_n} + 3\,\frac{\log a_n}{a_n}\right] > 0,$$

then $z = e^{i\theta} \notin E(g)$ *if* $|\theta - \theta_n| \le (\pi/3) - h,\ 0 < h < \pi/3,$ *for infinitely many n. In particular,* $E(g) = 0$ *if the set* $\{\theta_n\}$ *is dense in the interval* $[0, 2\pi]$.

6. The discussion of $C_n(z)$ on pp. 534, 535 of the preceding paper shows that they are constructed essentially in accordance with Theorem 2 above. The gap theorem in Corollary 2a is very crude, and can certainly be improved. The high-indices theorem of Hardy and Littlewood and Tauberian methods (see [2] and [3]) yield much sharper results.

7. The construction on p. 537 of Herzog and Piranian can also be carried out as follows.

LEMMA. *If A and B are disjoint closed sets in the plane and B is bounded and has a simply connected compliment, and* $\epsilon > 0$, *then there is a polynomial* $P(z)$ *such that* $|P(z)| \le \epsilon$ *on B and* $|P(z)| \ge 1$ *on A.*

Proof. Let $T_n(z)$ be the Chebyshev polynomial of degree n for B; that is, T_n is the polynomial of degree n with highest coefficient 1 whose maximum modulus on B is the least possible. Then $T_n(z)^{1/n} \longrightarrow \phi(z)$ in the exterior of B, where $\phi(z)$ is the function which maps the exterior of B onto the exterior of a circle $|w| > c$ and whose Taylor expansion at ∞ begins thus: $\phi(z) = z + \cdots$. Let $c < C < R$ be such that $|\phi(z)| \ge R + \epsilon$ on A. Then there is an n such that

$$|T_n(z)|^{1/n} \ge R \text{ on } A \quad\text{and}\quad |T_n(z)|^{1/n} \le C \text{ on } B.$$

If n is chosen such that $\epsilon(R/C)^n \ge 1$, then $R^{-n}\,T_n$ is a polynomial with the desired properties.

There are, of course, many other ways of constructing such a polynomial.

Now in the construction on p. 537, take a convergent double series $\sum a_{kh}$ with $a_{kh} > C$. Choose $A = I_{kh}$ and let B be the sector $z = re^{i\theta}$ with $0 \leq r \leq 1$ and θ in the closed interval complimentary to I_{kh} and its two adjacent intervals in G_k. Let P_{kh} be a polynomial such that $|P_{kh}(z)| \geq 1$ on I_{kh} and $|P_{kh}(z)| \leq a_{kh}$ on B. Arrange the pairs (k, h) in a sequence by the diagonal process, and apply Theorem 2, then Theorem 1.

8. The polynomials C_n used by Herzog and Piranian are of the desired type for the sets A and B considered in the preceding paragraph. They provide a simple explicit construction and enjoy other interesting properties which seem to be useful in a number of problems. The fact that they are small on the whole set B above follows from the following remark which is surely known:

If

$$f(z) = \sum_{n=0}^{\infty} a_n z^n \quad \text{and} \quad s_n(z) = \sum_{k=0}^{n} a_k z^k,$$

and $0 \leq r \leq 1$, $|z| \leq 1$, then $|f(rz)| \leq \sup_n |s_n(z)|$.

This is a trivial consequence of the identity $f(rz) = O(1-r)\sum_0^\infty r^n s_n(z)$.

References

1. F. Herzog and G. Piranian, *Sets of radial continuity of analytic functions*, Pacific J. Math. 4 (1954), 533-538.

2. N. Levinson, *Gap and density theorems*, Amer. Math. Soc. Colloquium Publications, New York, 1940.

3. N. Wiener, *A Tauberian gap theorem of Hardy and Littlewood*, J. Chinese Math. Soc. 1 (1936), 15.

UNIVERSITY OF MINNESOTA

REMARKS ON SPLITTING EXTENSIONS

D.G. HIGMAN

1. Introduction. If N is a normal subgroup of the finite group G we call G an *extension* of N. Such an extension G over N is said to *split* if there exists a *complement* of N in G, that is, if there exists a subgroup of G which contains exactly one element from each coset of G modulo N. A frequently used criterion for splitting is provided by a theorem of Schur, namely, *if N has order prime to its index in G, then G splits over N.* W. Gaschütz [**1**] has recently given a generalization of this theorem for the case when $N = A$ is abelian, which states that (i) *G splits over A if and only if there is for each prime p a p-Sylow subgroup S of G which splits over $S \cap A$,*[1] *and* (ii) *there exists a subgroup $U < G$ such that $G = AU$ if and only if there exists for some prime p a p-Sylow subgroup S of G and a subgroup V of S such that*

$$S = [S \cap A]V, \text{ and } \mathfrak{N}_G(V \cap A) = S \cap A.^{1}$$

Here $\mathfrak{N}_G(V \cap A)$ denotes the subgroup generated by all the conjugates to $V \cap A$ in G.

In § 2 of this note we apply part (i) of the theorem of Gaschütz to establish a generalization of the theorem of Schur for non-abelian extension. In § 3 we apply (ii) to obtain a characterization of extensions G over N such that N is contained in the Frattini subgroup. The remaining two sections are concerned with the question of conjugacy of complements.

NOTATIONS. Group will always mean finite group unless the contrary is explicitly stated. For H a subgroup of a group G, $[G:H]$ = index of H in G. For Y a set of elements of G, $\{Y\}$ = subgroup generated by the elements of Y. If A and B are groups, $A \times B$ denotes their direct product. $A \leq B$ means A is con-

[1] Since any two p-Sylow subgroups of G are conjugate, this condition is satisfied by all p-Sylow subgroups whenever it is satisfied by any one of them. The condition is automatically satisfied by those p-Sylow subgroups of G for which p does not divide both the order and the index of the normal subgroup.

Received July 8, 1952.

Pacific J. Math. 4 (1954), 545-555

tained in B, while $A < B$ means proper inclusion. $A \cap B$ = set theoretic intersection of A and B.

2. A subgroup C of G is a complement for the extension G over N if and only if $G = NC$ and $1 = N \cap C$.

THEOREM 1. *A subgroup C of a group G is a complement for the extension G over N if and only if C is minimal with respect to the property $G = NC$, and there exists for each prime p a p-Sylow subgroup S of G, and a complement of $N \cap S$ in S which is part of C.*[1]

Proof. Assume that C is a complement of N in G. Then clearly C is minimal with respect to the property $G = NC$. If P is a p-Sylow subgroup of C, and if S is a p-Sylow subgroup of G such that $P \leq S$, then P is a complement of $S \cap N$ in S. For, since $P \leq C$, $N \cap P \leq N \cap C = 1$. And since $P \leq S$, $[S \cap N]P \leq S$. But $S \cap N$ is a p-Sylow subgroup of N, and P is a p-Sylow subgroup of C, from which it follows that $[S \cap N]P$ is a p-Sylow subgroup of G. Hence $[S \cap N]P = S$. We have proved the necessity of the condition of the theorem.

Now assume conversely that this condition is satisfied. Let P be a Sylow subgroup of $M = N \cap C$, x an element of C. Since M is a normal subgroup of C, P^x is also a Sylow subgroup of M for the same prime. Hence there is an element y in M such that $P^{xy} = P$. Then xy is in the normalizer T of P in C, that is x is in MT. Hence $C = MT$, so that $G = NC = NMT = NT$. Hence by the minimality property of C, $T = C$. We have shown that each Sylow subgroup of M is normal in C, that is, that M is nilpotent.[2] We must prove that $M = 1$.

If p is a prime, there exists by our assumption a p-Sylow subgroup S of G, and a complement U of $S \cap N$ in S which is part of C. Since $U \leq S$, U is a p-subgroup of C. If Q is a p-Sylow subgroup of C such that $U \leq Q$, then U is a complement of $M \cap Q$ in Q. For, let P be a p-Sylow subgroup of G such that $Q \leq P$. Then there is an element x in G such that

$$P = S^x = [S \cap N]^x U^x = [P \cap N] U^x.$$

Hence, since $1 = U \cap N$, $P = [P \cap N] U$, so that

$$Q = Q \cap P = [Q \cap N] U = [Q \cap M] U.$$

[2] The condition that C be minimal with respect to the property $G = NC$ is equivalent to the condition $M = N \cap C \leq \phi(C)$. We may infer the nilpotency of M from the nilpotency of $\phi(C)$ (c.f. § 3).

For X a subgroup of G, set $\overline{X} = M'X/M'$, where M' denotes the commutator subgroup of M. Then $\overline{Q}$ is a p-Sylow subgroup of $\overline{C}$, and since $M'U \cap M = M'[U \cap M] = M'$, $\overline{U}$ is a complement of $\overline{Q} \cap \overline{M}$ in C. Hence, since $\overline{M}$ is abelian, there exists by part (i) of the theorem of Gaschütz a complement $\overline{D} = D/M'$ of $\overline{M}$ in $\overline{C}$. But then $C = MD$ and $M' = M \cap D$. Since M is nilpotent, $M \neq 1$ implies $M \cap D = M' < M = M \cap C$, that is $D < C$. Since $G = NC = NMD = ND$, this contradicts the minimality property of C. Hence $M = 1$, which proves the sufficiency of the condition.

COROLLARY (Schur's theorem). *If N has order prime to its index in G, then G splits over N.*

REMARK. Theorem 1 does not, of course, settle the question of the necessity of the hypothesis that N be abelian for the theorem of Gaschütz.[3]

The following example shows that in a splitting extension G over N, not every subgroup C which is minimal with respect to the property $G = NC$ need be a complement, even when N is abelian.

EXAMPLE. Let $M \neq 1$ be an abelian normal subgroup of the group C, and assume that M is contained in the Frattini subgroup $\phi(C)$ of C (*c.f.* § 3). Since $\phi(C)$ is nilpotent it will have a center $\neq 1$; we may take, for instance, $M =$ the center of $\phi(C)$. By a theorem of Artin [2, p. 103] there exists a free abelian group A of finite rank, and an (infinite) group G such that if we set $N = M \times A$, then

1. G is a splitting extension of N.
2. $G = NC$
3. $M = N \cap C$.

By the choice of M and C, no proper subgroup of C satisfies 2.

Let now m be the order of M. Since N is abelian, N^m [= the totality of mth powers of elements of N] is a characteristic subgroup of N, and hence is normal in G. Furthermore, $N^m \cap M = 1$. Since N is abelian of finite rank, and since G/N is finite, as an isomorphic image of the finite group C/M, G/N^m is finite. Set $\overline{G} = G/N^m$, $\overline{N} = N/N^m$ and $\overline{C} = N^m C/N^m$. Since the extension G over N splits, so does $\overline{G}$ over $\overline{N}$. $\overline{C}$ is minimal with respect to the property $\overline{G} = \overline{N}\overline{C}$, but

$$\overline{C} \cap \overline{N} = MN^m/N^m \simeq M \neq 1.$$

THEOREM 2. *For an extension G over N the following five conditions are*

[3] That this hypothesis actually is necessary has been shown by Professor Zassenhaus. See the note at the end of this paper.

equivalent.

(1) *N has order prime to its index in G.*

(2) *if H is a subgroup of G, then*

(a) *there exists a complement of $N \cap H$ in H.*

(b) *if either $H/[N \cap H]$ or $N \cap H$ is solvable, then any two complements of $N \cap H$ in H are conjugate in H.*

(3) *for each prime divisor p of the order of N, there exists a p-Sylow subgroup S of N such that if T denotes the normalizer of S in G,*

(a) *there exists a complement of $N \cap T$ in T.*

(b) *if H is a nilpotent subgroup of T, then any two complements of $N \cap H$ in H are conjugate in H.*

(4) *if H is a nilpotent subgroup of G, then*

(a) *there exists a complement of $N \cap H$ in H.*

(b) *any two complements of $N \cap H$ in H are conjugate in H.*

(5) *if H is a nilpotent subgroup of G, then there exists a subgroup C of H such that for each subgroup U of H, $U = [U \cap N] \times [U \cap C]$.*

Proof. Assume that N has order prime to its index in G, then clearly the same is true of the normal subgroup $N \cap H$ of H, for any subgroup H of G. Hence H splits over $N \cap H$ by the theorem of Schur. Furthermore, by a theorem of Zassenhaus [2, p. 132] if either $H/[N \cap H]$ or $N \cap H$ is solvable then any two complements for this extension are conjugate in H. Thus (2) is a consequence of (1).

Conditions (3) and (4) are immediate consequences of (2).

Next we shall prove that (3) and (4) each imply (1). Assume that the extension G over N satisfies (3), and assume that p is a prime which divides both the order and the index of N. Then by (3), (a) there exists a p-Sylow subgroup S of N, and a subgroup C such that if T denotes the normalizer of S in G, C is a complement of $N \cap T$ in T. But $G = NT$, so that C is a complement of N in G. Thus $[C:1] = [G:N]$, hence since p divides $[G:N]$, there exists an element x in C of order p. Since x is in T, and since p divides the order of N, there exists an element z of order p in $S \cap N$ such that $xz = zx$. Since x is not in N, $H = \{x, z\} = \{x\} \times \{z\}$ and $N \cap H = \{z\}$, whereby it follows from (3), (b), that $\{x\}$ and $\{xz\}$ are conjugate in H. Since this is impossible, (3) implies (1).

Now assume (4), and suppose again that p is a prime which divides both the order and the index of N. If S is a p-Sylow subgroup of G, there exists by

(4), (a), a complement C of $S \cap N$ in S. Since p divides $[G:N]$, there exists an element x in C of order p. Since p divides the order of N, $S \cap N$ is a non-trivial normal subgroup of S. Now a repetition of the construction of the preceding paragraph leads to a contradiction with (4), (b), proving that (4) implies (1). We have proved the equivalence of the first four conditions.

If H is a nilpotent subgroup of G, (2) implies the existence of a complement C of $N \cap H$ in H, and (1) implies that the orders of $N \cap H$ and C are relatively prime. Now (5) is a consequence of a property of nilpotent groups. Thus (5) is implied by the equivalent conditions (1) and (2). Conversely, if S is a p-Sylow subgroup of G, (5) implies the existence of a subgroup C of S such that $U = [U \cap N] \times [U \cap C]$ for each subgroup U of S. But it is well known that this implies that $S \cap N$ and C have relatively prime orders. Hence one of $S \cap N$ and C is trivial. This proves that (5) implies (1), completing the proof of Theorem 2.

3. **The Frattini subgroup** $\phi(G)$ of the group G is the intersection of G with all its maximal subgroups. In this section we shall note a characterization of those normal subgroups N of G which are contained in $\phi(G)$. It is well known that

(a) *$N \leq \phi(G)$ if and only if $G = NC$, C a subgroup of G implies $G = C$.*

Hence part (ii) of the theorem of Gaschütz has an equivalent statement

(b) *the abelian normal subgroup A of G is contained in $\phi(G)$ if and only if for each prime p there is a p-Sylow subgroup S of G such that $S = [S \cap A]V$, V a subgroup, implies $S \cap A = \mathfrak{n}_G(V \cap A)$.*[1]

Using (a) it is easy to verify that

(c) *if M is a normal subgroup of G such that $M \leq N$, then $N \leq \phi(G)$ if and only if $M \leq \phi(G)$ and $N/M \leq \phi(G/M)$.*

Since $\phi(G)$ is nilpotent [2, p. 122; this can be proved using (a) together with the first part of the argument of the sufficiency proof of Theorem 1] it will suffice for the purposes of determining the normal subgroups N which are contained in $\phi(G)$ to consider the case in which N has prime power order.

$N^{(i)}$ denotes the ith derived subgroup of N, $N^{(0)} = N$, $N^{(1)} = N'$. For X a subgroup of G, $\mathfrak{n}_G(X)$ denotes the subgroup generated by all the conjugates to X in G.

THEOREM 3. *Let N be a normal subgroup of the group G, and assume that N has p-power order, p a prime. Then $N \leq \phi(G)$ if and only if there exists a* *p-Sylow subgroup S of G such that for all $i \geq 0$, $S = N^{(i)}V$, V a subgroup, implies*

$$N^{(i)} = N^{(i+1)}\mathfrak{n}_G(V \cap N^{(i)}).^{1}$$

Proof. Assume first that $N \leq \phi(G)$. For X a subgroup of G, write $\overline{X} = N^{(i+1)}X/N^{(i+1)}$. Then $\overline{N^{(i)}}$ is an abelian normal subgroup of G with p-power order. Furthermore, by (c), $\overline{N^{(i)}} \leq \phi(\overline{G})$. Let $\overline{\zeta}$ be the p-Sylow subgroup of $\overline{G}$ whose existence is inferred by (b) (indeed, any p-Sylow subgroup will do[1]). Then $S = \underline{S}/N^{(i+1)}$, S a p-Sylow subgroup of G. If $S = N^{(i)}V$, V a subgroup, then $S = \overline{N^{(i)}}\ \overline{V}$. Hence by (b) it follows that $\overline{N^{(i)}} = \mathfrak{n}_G(\overline{V} \cap \overline{N^{(i)}})$. But

$$\overline{N^{(i)}} \cap \overline{V} = N^{(i)}/N^{(i+1)} \cap N^{(i+1)}V/N^{(i+1)} = N^{(i+1)}[N^{(i)} \cap V]/N^{(i+1)},$$

from which it is easily verified that

$$\mathfrak{n}_{\overline{G}}\,(\overline{V} \cap N^{(i)}) = N^{(i+1)}\mathfrak{n}_G(V \cap N^{(i)})/N^{(i+1)}.$$

Hence $N^{(i)} = N^{(i+1)}\mathfrak{n}_G(V \cap N^{(i)})$. We have proved the necessity of the condition of the theorem.

Assume conversely that this condition is satisfied. We prove $N \leq \phi(G)$ by induction on the order of N. If $N = 1$ there is nothing to prove. Otherwise, since N is a p-group, $N' < N$, and since the condition of the theorem is clearly satisfied by N' whenever it is satisfied by N, it follows from the induction hypothesis that $N' \leq \phi(G)$. $\overline{S} = S/N'$ is a p-Sylow subgroup of $\overline{G} = G/N'$. If $\overline{V} = V/N'$ is a subgroup of $\overline{G}$ such that $\overline{S} = \overline{N}\,\overline{V}$, $\overline{N} = N/N'$, then $S = NV$. Now the condition of the theorem implies

$$\overline{N} = N/N' = N'\mathfrak{n}_G(V \cap N)/N' = \mathfrak{n}_G(V \cap N)/N' = \mathfrak{n}_{\overline{G}}(\overline{V} \cap \overline{N}).$$

Hence, since N is an abelian p-group it follows from (b) that $\overline{N} \leq \phi(\overline{G})$. Hence $N \leq \phi(G)$ by (c).

4. In this section we assume that $N = A$ is abelian, and consider the problem of the conjugacy of complements of A in G. A complement C of A in G is in particular a set of representatives for G over A; C consists of exactly one element $c(X)$ from each coset X in G/A. If D is a second complement, $d(X) = D \cap X$, then the function t from G/A to A defined by $d(X) = t(X)c(X)$ satisfies

$$(1) \qquad 1 = t(Y)^X\, t(XY)^{-1}\, t(X)$$

for all X, Y in G/A. (Since A is abelian, all the elements x in X induce the same automorphism of A. We write $a^X = a^x$ for a in A).

Conversely, if t is any function from G/A to A which satisfies (1), then the totality D of elements $d(X) = t(X)c(X)$ for X in G/A is a complement of A in G. Moreover

(2) *two complements C and D which are related by t are conjugate subgroups of G if and only if there is an element a in A such that $t(X) = a^{1-X}$ for X in G/A.*[3]

Let H be a subgroup of G such that $A \leq H$, and set $m = [G:H]$.

THEOREM 4. *If m is prime to the order of A, if the function t from G/A to A satisfies (1), and if c is an element of A such that $t(Y) = c^{1-Y}$ for all Y in H/A, then there is an element a in A such that $t(X) = a^{1-X}$ for all X in G/A.*[4]

Proof. The function f defined by $f(X) = t(X)c^{X-1}$ satisfies (1), and has the property that $f(Y) = 1$ for all Y in H/A. Choose a system L of left representatives for G/A over H/A so that each X in G/A has (uniquely) the form $X = \overline{X}\underline{X}$, with $\overline{X}$ in L, $\underline{X}$ in H/A. By (1) we have

$$1 = f(\underline{X})^{\overline{X}} f(X)^{-1} f(\overline{X}) = f(X)^{-1} f(\overline{X})$$

that is $f(X) = f(\overline{X})$. Hence

$$f(X) = f(Y)^{-X} f(XY) = f(Y)^{-X} f(\overline{XY}).$$

Taking the product over all Y in L we have

$$f(X)^m = \prod_{Y \in L} f(Y)^{-X} \prod_{Y \in L} f(\overline{XY}). \tag{3}$$

Since m is prime to the order of A, the mapping $\alpha : a \longrightarrow a^m$ is an automorphism of A (which commutes with every other automorphism of A). Hence

---

[3] In terms of the cohomology theory of groups this means that the number of classes of conjugate complements of A in G is the order of the first cohomology group of G/A by A.

[4] This result is a consequence of the 1-dimensional case, whereas (i) of the Gaschutz theorem is a consequence of the 2-dimensional case, of a general theorem in the cohomology theory of groups (see B. Eckmann, *Cohomology groups and transfer*, Ann. of Math., 58 (1953), 481-493.

$$b = \left\{ \prod_{Y \in L} f(Y) \right\} \alpha^{-1}$$

is an element of A. As Y runs through L, so does $\overline{XY}$, hence, applying α^{-1} to (3) we have $f(X) = b^{-X} b = b^{1-X}$. Thus

$$t(X) = f(X) c^{1-X} = b^{1-X} c^{1-X} = (bc)^{1-X}.$$

Theorem 4 is now proved with $a = bc$.

COROLLARY 1. *If m is prime to the order of A, then two complements C and D of A in G are conjugate in G if and only if $C \cap H$ and $D \cap H$ are conjugate in H.*

Proof. Let t be the function relating C and D. The subgroups $C \cap H$ and $D \cap H$ are complements of A in H, and are related by the restriction of t to H/A. If C and D are conjugate in G, then by (2) there exists an element a in A such that $t(X) = a^{1-X}$ for all X in G/A, and hence in particular for X in H/A. Hence by (2), $C \cap H$ and $D \cap H$ are conjugate in H.

If on the other hand $C \cap H$ and $D \cap H$ are conjugate subgroups of H, then it follows by (2) that there is an element c in A such that $t(Y) = c^{1-y}$ for all y in H/A. Hence by Theorem 4 there exists a in A such that $t(X) = a^{1-X}$ for all X in G/A. Hence by (2), C and D are conjugate in G. This proves the corollary.

By part (i) of the theorem of Gaschütz the extension G over A splits if and only if there is for each prime p a p-Sylow subgroup S of G which splits over $S \cap A$. By Theorem 4 we have

COROLLARY 2. *Let G be a splitting extension of A. If for each prime p there is a p-Sylow subgroup S of G such that any two complements of $S \cap A$ in S are conjugate in S, then any two complements of A in G are conjugate in G.*

Proof. We must prove that for each function t satisfying (1) there is an element a in A such that $t(X) = a^{1-X}$ for all X in G/A. Let p_i be the prime divisors of the order of A and let A_i be the corresponding primary components of A $(i = 1, 2, \cdots, k)$. Then $A = A_1 \times \cdots \times A_k$, and each A_i, being characteristic in A, is a normal subgroup of G. For each X in G/A, $t(X)$ has (uniquely) the form

$$t(X) = \prod_{i=1}^{k} t_i(X),$$

with $t_i(X)$ in A_i. Define $T_i(A_i x) = t(Ax)$ for x in G, and let S_i be a p_i-Sylow subgroup of G. We have assumed that S_i may be chosen in such a way that there is an element b_i in A_i with

$$T_i(A_i y) = b_i^{1-A_i y}$$

for all y in S_i (indeed, any p_i-Sylow subgroup will do). By Theorem 4, there exists a_i in A such that

$$T_i(A_i x) = a_i^{1-A_i x}$$

for all x in G. Hence

$$t_i(Ax) = T_i(A_i x) = a_i^{1-A_i x} = a_i^{1-Ax},$$

whereby

$$t(Ax) = \prod_{i=1}^{k} t_i(Ax) = \prod_{i=1}^{k} a_i^{1-Ax} = \left\{\prod_{i=1}^{k} a_i\right\}^{1-Ax}$$

for all x in G, with $a = \prod_{i=1}^{k} a_i$ an element of A.

5. **It has been conjectured that** if N has order prime to its index in G, then any two complements of N in G are conjugate. The following theorem shows that this conjecture is equivalent to

(+) *if G is a group, and Γ a group of automorphisms of G such that the orders of Γ and G are relatively prime, then for each prime p, there exists a p-Sylow subgroup of G which is mapped onto itself by every automorphism in Γ.*

Thus the theorem of Zassenhaus [2, p. 132] suffices to prove (+) in case either G or Γ is solvable.

THEOREM 5. *For an extension G over N such that N has order prime to its index in G, the following are equivalent statements.*

(a) *if C and D are complements of N in G, then they are conjugate in $\{C, D\}$.*

(b) *for each subgroup H of G such that $G = NH$, and for each pair C, D of complements of $N \cap H$ in H, there exists an automorphism α of H such that $C = D^{\alpha}$.*

(c) *for each subgroup H of G such that $G = NH$, for each complement C of $N \cap H$ in H, and for each prime p, there exist a p-Sylow subgroup S of $N \cap H$ such that C is part of the normalizer of S.*

Proof. Clearly (a) implies (b). Assume (b), and let H be a subgroup of G such that $G = NH$. Let P be a p-Sylow subgroup of $N \cap H$, and let T be the normalizer of P in H. Then $H = [N \cap H]T$. Hence, since the order of $N \cap H$ is prime to its index in H, there exists by the theorem of Schur a complement D of $N \cap H$ in H which is part of T, that is, which normalizes P. If now C is any complement of $N \cap H$ in H, there exists by (b) an automorphism of H such that $C = D^{\alpha}$. Hence C normalizes the p-Sylow subgroup $S = P^{\alpha}$ of $N \cap H$. Thus (b) implies (c).

Now assume (c), and let C and D be two complements of N in G. Assume that if $\{U, V\}$ is a pair of complements of N in G such that the order of $\{U, V\}$ is less than the order of $H = \{C, D\}$, then U and V are conjugate in $\{U, V\}$. If $N \cap H$ is nilpotent, since we have assumed that the orders of N and G/N are relatively prime, it follows by the theorem of Zassenhaus that C and D are conjugate in H. Otherwise, there exists a prime p such that the normalizer in H of a p-Sylow subgroup of $N \cap H$ is a proper subgroup of H. By (c) there exist p-Sylow subgroups P and Q of $N \cap H$ which are normalized by C and D respectively. There exists an element x in H such that $P = Q^x$, and the complement $E = D^x$ of N in G normalizes P. Thus, if we let T denote the normalizer of P in H, $\{C, E\} \leq T < H$. Now it follows by the induction hypothesis that there exists an element y in $\{C, E\}$ such that $C = E^y = D^{xy}$. Since both x and y are in $H = \{C, D\}$, so is xy.

Added in proof. The very interesting fact, that the hypothesis that the extension be abelian is indeed necessary for Gaschütz's theorem (i), as stated in the introduction of the present note, is shown by the following example communicated to the author by Professor Zassenhaus:

Let G be the group with generators A_{ik}, B_i, C_i $(i, k = 1, 2)$ and the defining relations

$$A_{11}^2 = A_{12}^2 = (A_{11}A_{12})^2 = A_{21}^2 = A_{22}^2 = (A_{21}A_{22})^2,\ A_{1i}A_{2k} = A_{2k}A_{1i};$$

$$B_i^3 = 1,\ B_i A_{i1} B_i^{-1} = A_{i2},\ B_i A_{i2} B_i^{-1} = A_{i1}A_{i2},\ B_1 A_{2k} = A_{2k} B_1,\ B_2 A_{1k} = A_{1k} B_2,$$

$$B_1 B_2 = B_2 B_1;\ C_i^2 = (C_i A_{i1})^2 = (C_i A_{i1} A_{i2})^2 = A_{i1},\ C_i B_i C_i^{-1} = B_i^{-1},$$

$$C_1 A_{2k} C_1, C_2 A_{1k} = A_{1k} C_2, C_1 B_2 = B_2 C_1, C_2 B_1 = B_1 C_2, C_1 C_2 = C_2 C_1 \quad (i, k = 1, 2).$$

The subgroup N generated by the four elements A_{ik} is the direct product of two quaternion groups with identified centers, thus N is of order 32. The group N is normal in G and the subgroup G_1 of G generated by N, B_1 and B_2 is normal too such that G_1/N is abelian of type (3,3). The factor group G/G_1 is of type (2,2). Thus G is of order 1152.

The group G does not split over its normal subgroup N. But the factor group G_1/N is the 3-Sylow subgroup of the factor group G/N such that G_1 splits over N with the subgroup generated by B_1 and B_2 as representative subgroup. Moreover the factor group generated by N, C_1 and C_2 over N is a 2-Sylow subgroup of G/N such that the subgroup generated by $C_1 A_{21}$ and $C_2 A_{12}$ is a representative subgroup of order 4.

REFERENCES

1. W. Gaschütz, *Zur Erweiterungstheorie der endlichen Gruppen.*, J. Reine Angew. Math. 190 (1952), 93-107.

2. H. Zassenhaus, *The theory of groups.*, Chelsea Publishing Company, New York, 1949.

UNIVERSITY OF ILLINOIS,
MCGILL UNIVERSITY, MONTREAL
MONTANA STATE UNIVERSITY

TRANSFORMATIONS OF SERIES OF THE TYPE ${}_3\Psi_3$

MARGARET JACKSON

1. **Sears** [3] has given relations between series of the type ${}_3\Phi_2$. Generalizations of some of these results are included in, or may be obtained from, the following two formulae established by Slater [4]:

$$\prod_{r=0}^{\infty} \frac{(1-x\xi q^r)(1-q^{r+1}/x\xi)(1-b_1q^r)\cdots(1-b_Mq^r)}{(1-a_1q^r)\cdots(1-a_Mq^r)}$$

$$\times \frac{(1-q^{r+1}/a_{M+2})\cdots(1-q^{r+1}/a_{2M+1})}{(1-q^{r+1}/a_1)\cdots(1-q^{r+1}/a_M)} \;{}_M\Psi_M\left[\begin{matrix} a_{M+2},\cdots,a_{2M+1}\,;\,x \\ b_1,\cdots,b_M \end{matrix}\right]$$

$$= q/a_1 \prod_{r=0}^{\infty}\left[\frac{(1-a_1x\xi q^{r-1})(1-q^{r+2}/a_1x\xi)(1-b_1q^{r+1}/a_1)\cdots}{(1-a_1q^r)(1-q^{r+1}/a_1)(1-a_1q^r/a_2)\cdots}\right.$$

$$\text{(1.1)} \qquad \left.\times\cdots\frac{(1-b_Mq^{r+1}/a_1)(1-a_1q^r/a_{M+2})\cdots(1-a_1q^r/a_{2M+1})}{(1-a_1q^r/a_M)(1-a_2q^{r+1}/a_1)\cdots(1-a_Mq^{r+1}/a_1)}\right]$$

$$\times\;{}_M\Psi_M\left[\begin{matrix} qa_{M+2}/a_1,\cdots,qa_{2M+1}/a_1\,;\,x \\ qb_1/a_1,\cdots,qb_M/a_1 \end{matrix}\right]$$

$+ (M-1)$ similar terms obtained by interchanging a_1 with $a_2, a_3,\cdots,a_M$,

$$= q/a_1 \prod_{r=0}^{\infty}\left[\frac{(1-a_1x\xi q^{r-1})(1-q^{r+2}/a_1x\xi)(1-b_1q^{r+1}/a_1)\cdots}{(1-a_1q^r)(1-q^{r+1}/a_1)(1-a_1q^r/a_2)\cdots}\right.$$

Received February 12, 1953. This work was included in a doctoral thesis approved by the University of Nottingham.

Pacific J. Math. 4 (1954), 557-562

$$\times \frac{(1-b_Mq^{r+1}/a_1)(1-a_1q^r/a_{M+2})\cdots(1-a_1q^r/a_{2M+1})}{(1-a_1q^r/a_M)(1-a_2q^{r+1}/a_1)\cdots(1-a_Mq^{r+1}/a_1)}\Bigg]$$

(1.2)

$$\times {}_M\Psi_M\left[\begin{matrix} a_1/b_1, \cdots, a_1/b_M; \\ a_1/a_{M+2}, \cdots, a_1/a_{2M+1} \end{matrix} \quad \frac{b_1\cdots b_M}{xa_{M+2}\cdots a_{2M+1}}\right]$$

$+ \ (M-1)$ similar terms obtained as in (1.1),

where

$$M \geq 1, \ \xi = \frac{a_{M+2}\cdots a_{2M+1}}{a_1\cdots a_M}, \ |x| < 1, \text{ and } |q| < 1.$$

In particular we see that (1.2), with $M = 3$, is a generalization of the basic analogue of the fundamental three-term relation [3, § 10, result IVa] for ${}_3F_2$ to which it reduces if we take $a_1 = aq$, $a_2 = bq$, $a_3 = cq$, $a_5 = a$, $a_6 = b$, $a_7 = c$, $b_1 = q$, $b_2 = e$, $b_3 = f$, and $x = ef/abc$. Similarly, (1.1) and (1.2) may be used to obtain many more of the relations given by Sears. It will be noted, however, that the parameters occurring in the Ψ series in (1.1) and (1.2) are related in a very symmetrical way, and consequently these formulae can only be expected to provide generalizations of the two-, three-, and four-term relations between ${}_3\Phi_2$ which are of a symmetrical nature; in particular, they do not provide a generalization of the basic analogue of the fundamental two-term relation [3, § 10, 1]. In this paper, one such generalization is obtained which, when used in conjunction with (1.1), will yield generalizations of all Sears' formulae and provide basic analogues of known transformations [2] of ${}_3H_3$.

2. **To obtain** the required generalization, we establish the basic analogue of the formula [2, § 2.1] which was used to obtain the generalization of the fundamental two-term relation between ${}_3F_2$. The method by which this result can be obtained has been indicated by Bailey [1], who obtained a particular case of the following formula (2.1). We use the fact that a basic bilateral series ${}_8\Psi_8$ which terminates below can be expressed in terms of an ${}_8\Phi_7$, which can in turn be transformed into two series ${}_4\Phi_3$, one of which can be replaced by a ${}_4\Psi_4$ which terminates below. Then, proceeding to the limit, we obtain a transformation which can be restated in the form (2.1). The analysis is straightforward, though rather lengthy, so we just state the result:

$$(2.1)\qquad \sum_{n=-\infty}^{\infty}\left[\frac{(q\sqrt{a})_n(-q\sqrt{a})_n(b)_n(c)_n(d)_n}{(\sqrt{a})_n(-\sqrt{a})_n(aq/b)_n(aq/c)_n(aq/d)_n}\times\frac{(e)_n(f)_n(-1)^n q^{n^2/2+n}}{(aq/e)_n(aq/f)_n}\left(\frac{a^3}{bcdef}\right)^n\right]$$

$$=\prod_{r=0}^{\infty}\frac{(1-aq^{r+1})(1-q^{r+1}/a)(1-aq^{r+1}/bc)}{(1-q^{r+1}/d)(1-q^{r+1}/e)(1-q^{r+1}/f)(1-aq^{r+1}/b)(1-aq^{r+1}/c)}$$

$$\times\left\{\prod_{r=0}^{\infty}\frac{(1-aq^{r+1}/de)(1-aq^{r+1}/ef)(1-aq^{r+1}/df)}{(1-a^2q^{r+1}/def)}\right.$$

$$\times\ {}_3\Psi_3\left[\begin{matrix}b,\ c,\ a^2q/def; \\ aq/d, aq/e, aq/f\end{matrix}\ \frac{aq}{bc}\right]$$

$$+\prod_{r=0}^{\infty}\frac{(1-dq^r/a)(1-eq^r/a)(1-fq^r/a)}{(1-q^{r+1}/b)(1-q^{r+1}/c)}$$

$$\times\frac{(1-a^2q^{r+2}/bdef)(1-a^2q^{r+2}/cdef)(1-q^{r+1})}{(1-a^2q^{r+2}/def)(1-defq^{r-1}/a^2)}$$

$$\left.\times\ {}_3\Phi_2\left[\begin{matrix}aq/ef,\ aq/df,\ aq/de\ ;\ q \\ a^2q^2/bdef,\ a^2q^2/cdef\end{matrix}\right]\right\}.$$

We obtain a generalization of the basic analogue of the fundamental two-term relation by interchanging both b and d and c and e in (2.1), then replacing a by def/aq^2, d by ef/aq, e by df/aq, f by de/aq, leaving b and c unaltered, and replacing $def/abcq$ by σ, we obtain:

$$\prod_{r=0}^{\infty}\frac{(1-\sigma q^r)}{(1-aq^{r+2}/ef)(1-aq^{r+2}/df)(1-\sigma cq^r)(1-\sigma bq^r)}$$

$$\times\left\{\prod_{r=0}^{\infty}\frac{(1-aq^{r+1}/d)(1-aq^{r+1}/e)(1-aq^{r+1}/f)}{(1-aq^{r})}\ {}_3\Psi_3\left[\begin{matrix}a, b, c; \\ d, e, f\end{matrix}\ \frac{def}{abcq}\right]\right.$$

$$+\prod_{r=0}^{\infty}\frac{(1-q^{r+1}/d)(1-q^{r+1}/e)(1-q^{r+1}/f)}{(1-q^{r+1}/b)(1-q^{r+1}/c)}$$

$$\left.\times\frac{(1-aq^{r}/b)(1-aq^{r}/c)(1-q^{r+1})}{(1-aq^{r+1})(1-q^{r}/a)}\ {}_3\Phi_2\left[\begin{matrix}aq/d, aq/e, aq/f; \ q \\ aq/b, aq/c\end{matrix}\right]\right\}$$

$$=\prod_{r=0}^{\infty}\frac{(1-aq^{r+1}/f)}{(1-q^{r+1}/b)(1-q^{r+1}/c)(1-dq^{r})(1-eq^{r})} \tag{2.2}$$

$$\times\left\{\prod_{r=0}^{\infty}\frac{(1-\sigma q^{r})(1-fq^{r}/b)(1-fq^{r}/c)}{(1-f\sigma q^{r-1})}\ {}_3\Psi_3\left[\begin{matrix}ef/aq, df/aq, f/q; \\ b, \quad c, \quad f\end{matrix}\ \frac{aq}{f}\right]\right.$$

$$+\prod_{r=0}^{\infty}\frac{(1-q^{r+1}/c\sigma)(1-q^{r+1}/b\sigma)(1-q^{r+1})}{(1-aq^{r+2}/ef)(1-aq^{r+2}/df)}$$

$$\left.\times\frac{(1-q^{r+1}/f)(1-dfq^{r}/bc)(1-efq^{r}/bc)}{(1-\sigma fq^{r})(1-q^{r+1}/f\sigma)}\ {}_3\Phi_2\left[\begin{matrix}f/c, f/b, \sigma; \ q \\ df/bc, ef/bc\end{matrix}\right]\right\}.$$

The two ${}_3\Phi_2$ which occur in this formula are not connected by a two-term relation, and it would appear therefore that (2.2) is probably the simplest generalization of the fundamental two-term relation for ${}_3\Phi_2$ to which it reduces when $f=q$. This is the only relation between ${}_3\Phi_2$ which can be obtained from (2.2).

There are some relations involving ${}_3\Psi_3$, which generalize more than one ${}_3\Phi_2$ transformation. Such a formula can be obtained from (2.1) by interchanging the parameters b and d, then replacing a by def/aq^2, d by ef/aq, e by df/aq, f by de/aq, but leaving b and c unaltered:

$${}_3\Psi_3\left[\begin{matrix}a, b, c; \\ d, e, f\end{matrix}\ \frac{def}{abcq}\right]$$

$$
(2.3) \qquad = \prod_{r=0}^{\infty} \frac{(1-aq^r)(1-aq^{r+2}/ef)(1-\sigma c q^r)}{(1-q^{r+1}/b)(1-dq^r)(1-\sigma q^r)}
$$

$$
\times \frac{(1-dq^r/c)(1-eq^r/b)(1-fq^r/b)}{(1-aq^{r+1}/e)(1-aq^{r+1}/f)(1-efq^{r-1}/b)} \; {}_3\Psi_3\left[\begin{matrix} c,\ ef/aq,\ ef/bq\ ; & \dfrac{d}{c} \\ \sigma c,\ e,\ f & \end{matrix}\right]
$$

$$
+ \prod_{r=0}^{\infty} \frac{(1-q^{r+1}/e)(1-q^{r+1}/f)(1-aq^{r+1}/b)}{(1-aq^{r+1}/d)(1-q^{r+1}/b)(1-q^{r+1}/c)}
$$

$$
\times \frac{(1-q^{r+1})(1-aq^r)(1-c\sigma q^r)}{(1-aq^{r+1}/e)(1-aq^{r+1}/f)(1-\sigma q^r)}
$$

$$
\times \left\{ \prod_{r=0}^{\infty} \frac{(1-q^{r+1}/c\sigma)(1-efq^r/bc)(1-dq^r/c)}{(1-efq^r/b)(1-bq^{r+1}/ef)(1-dq^r)} \; {}_3\Phi_2\left[\begin{matrix} aq/d,\ f/b,\ e/b\ ; & q \\ aq/b,\ ef/bc & \end{matrix}\right] \right.
$$

$$
\left. - \prod_{r=0}^{\infty} \frac{(1-\sigma q^r)(1-q^{r+1}/d)(1-aq^{r+1}/c)}{(1-c\sigma q^r)(1-aq^{r+1})(1-q^r/a)} \; {}_3\Phi_2\left[\begin{matrix} aq/d,\ aq/e,\ aq/f\ ; & q \\ aq/b,\ aq/c & \end{matrix}\right] \right\}.
$$

If e (or f) $= q$, (2.3) reduces to a two-term relation; but it reduces to a four-term relation between $_3\Phi_2$ when $c = 1$. This particular result is not stated explicitly by Sears but can be deduced from his results.

It will be seen that the $_3\Psi_3$ transformations are more complicated than the analogous $_3H_3$ transformations. For this reason, no more such results are given, but they can all be obtained from (1.1) and (2.2).

3. Corrigenda. In (2.3) and (2.4) of [2], the terms $\Gamma(1+b-\sigma)$, $\Gamma(1+c-\sigma)$ should be $\Gamma(1-b-\sigma)$, $\Gamma(1-c-\sigma)$, in (5.1) the factor $\Gamma(d-c)$ on the left should be in the denominator of the first term on the right, and there should be a factor $\Gamma(d)$ in the denominator on the left.

References

1. W. N. Bailey, *Series of hypergeometric type which are infinite in both directions*, Quart. J. Math., Oxford Ser., 7 (1936), 105-115.

2. M. Jackson, *Transformations of series of the types $_3H_3$ with unit arguments*, J. London Math. Soc. 27 (1952), 116-123.

3. D. B. Sears, *On the transformation theory of basic hypergeometric functions*, Proc. London Math. Soc. (2) 53 (1951), 158-180.

4. L. J. Slater, *General transformations of bilateral series*, Quart. J. Math., Oxford Ser. (2) 9 (1952), 73-80.

THE UNIVERSITY, NOTTINGHAM

TRANSFORMATIONS OF SYSTEMS OF RELATIVISTIC PARTICLE MECHANICS

HERMAN RUBIN AND PATRICK SUPPES

1. Introduction. In [7] the axiomatic foundations of classical particle mechanics were investigated; and in [8] the transformations which carry systems of classical particle mechanics into systems of classical particle mechanics were determined. The purpose of the present paper is a similar investigation of relativistic particle mechanics (in the sense of the special theory of relativity). Some remarks on the general orientation of these studies are to be found in [7, § 1] and in [9].

In regard to our axiomatization of relativisitic particle mechanics, we want to emphasize that we have in no sense attempted to use primitive notions which are logically or epistemologically simple. Investigations with these latter aims are to be found in [11], [12], [13], and [14]; but these studies are incomplete in the sense that they do not give axioms adequate for relativisitic particle mechanics as it is ordinarily conceived by physicists. We have attempted to present such a complete set of axioms in a mathematically clear way.

The main result of the present paper is the determination under a certain weak hypothesis of the set of transformations which always carry systems of relativistic particle mechanics into systems of relativistic particle mechanics. Although this set of transformations is not a group (under the usual operation) we are able to show that it is essentially a Brandt groupoid. It is difficult precisely to compare our results with those in [6], but our results seem to represent an improvement in three respects: (i) we work within an explicit axiomatic framework; (ii) we consider transformations of the units of mass and force as well as position and time; (iii) we consider transformations from one value for the velocity of light to another.

We briefly summarize the mathematical notations we use, most of which are standard. We denote the ordered n-tuple whose first member is a_1, whose second member is a_2, and so on, by

Received May 7, 1953. The authors are grateful to Professor J.C.C. McKinsey for a large number of helpful suggestions and criticisms. This work was supported in part by a grant from the Office of Naval Research.

Pacific J. Math. 4 (1954), 563-601

$$\langle a_1, \cdots, a_n \rangle .$$

By an n-dimensional vector we mean an ordered n-tuple of real numbers. Operations on vectors are defined in the usual way. We use the symbol "0" to denote the real number zero, the n-dimensional vector all of whose components are zero, and the matrix all of whose elements are zero. If $A = \langle a_1, \cdots, a_n \rangle$ is any vector, the length $|A|$ of A is defined by

$$|A| = \sqrt{a_1^2 + \cdots + a_n^2} \ ;$$

and by $[A]_{i_1, i_2, \cdots, i_r}$ we mean the r-dimensional vector $\langle a_{i_1}, a_{i_2}, \cdots, a_{i_r} \rangle$. Thus if $A = \langle 4, 7, 5 \rangle$, then $[A]_{2,3} = \langle 7, 5 \rangle$. If A is a vector, we sometimes write "A^2" for "$|A|^2$." If $\mathfrak{A}$ is a matrix, we denote the transpose of $\mathfrak{A}$ by "$\mathfrak{A}^*$," and the determinant of $\mathfrak{A}$ by "$|\mathfrak{A}|$." We denote the identity matrix by "$\mathfrak{I}$." Although we treat vectors as one-rowed matrices, if A is a vector we always mean by $|A|$ the length of A and not the determinant of A: the meaning should be clear from the context. We use both matrix notation and usual vector notation for the inner product of two vectors A and B. Thus we sometimes write: AB^*, and sometimes: $A \cdot B$, whichever is more convenient.

We use Menger's notation for derivatives (see [**10**]). If f is a function, then $D(f)$ is the derivative of f. Thus, for example,

$$D(\sin) = \cos, \ \ [D(\sin)](x) = \cos x, \ \text{and} \ [D^2(\sin)](x) = -\sin x .$$

In this connection, we use the standard notation for sums, products, quotients, square roots, and so on, of functions. Thus, for example, if f and g are functions of a real variable, by $f + g$ we mean the function h such that for every real number x

$$h(x) = f(x) + g(x).$$

If f is a one-to-one function, f^{-1} is the inverse function of f. It is also convenient to introduce a special symbol for the *composition* of two functions: if f and g are functions of a real variable, by $g \circ f$ we mean the function h such that for every real number x

$$h(x) = g(f(x)).$$

To make some of our equations involving derivatives more perspicuous in relation to the notation ordinarily used in physics, we introduce formally the following two symbols: if f and g are functions of a real variable, then the function df/dg is defined by the following equation (for all real numbers x)

$$\frac{df}{dg}(x) = \left(\frac{Df}{Dg}\right)(x),$$

and the function d^2f/dg^2 by the equation

$$\frac{d^2f}{dg^2}(x) = \left(\frac{D[Df/Dg]}{Dg}\right)(x).$$

Finally, we also use the following notation: I is the set of all positive integers, R is the set of all real numbers, R^+ is the set of all positive real numbers, and E_n is the set of all n-dimensional vectors. We sometimes use geometrical language, referring to vectors in E_n as points in n-dimensional Euclidean space, and so on.

2. Primitive notions. Our axioms for relativistic particle mechanics are based on six primitive notions: P, $\mathfrak{I}$, m, s, f, and c. P is a set, $\mathfrak{I}$ and m are unary functions, s is a binary function, f is a ternary function, and c is a constant.

The intended physical interpretation of P is as the set of particles. For every p in P, $\mathfrak{I}(p)$ is to be interpreted physically as a set of real numbers measuring elapsed times (in terms of some unit of time and measured from some origin of time). There is a good physical reason for assigning (possibly) different sets of real numbers to different particles, instead of having one set of elapsed times for the whole system, as in [7]: two particles which have a simultaneous "life-span" with respect to one inertial frame of reference may have life-spans which do not even overlap with respect to another inertial frame.

For every p in P, $m(p)$ is to be interpreted physically as the numerical value of the rest mass of p. For every p in P and t in $\mathfrak{I}(p)$, $s(p, t)$ is a vector, to be thought of physically as giving the position of p at time t. Thus the primitive s fixes the choice of a coordinate system. It is also possible to take as a primitive the set of all admissible (that is, inertial) coordinate systems; this procedure is followed in [3]. We remark that for a fixed p in P, it is usually convenient to use in place of s the function s_p, which is defined on $\mathfrak{I}(p)$ and is such that, for every t in $\mathfrak{I}(p)$,

$$s_p(t) = s(p, t).$$

For every p in P and t in $\mathfrak{I}(p)$, and for i any positive integer, $f(p, t, i)$ is a vector giving the components (parallel to the axes of the coordinate system) of the ith force acting on p at time t. For further discussion of this primitive, applicable to relativistic as well as classical particle mechanics, see [7].

Our primitive constant c is to be interpreted as the numerical value of the velocity of light.

3. Axioms. Using the six primitive notions just described, we now give our axioms for relativistic particle mechanics.

An ordered sextuple $\Gamma = \langle P, \mathfrak{I}, m, s, f, c\rangle$ which satisfies the following Axioms A1-A7 is called an *n-dimensional system of relativistic particle mechanics* (or sometimes, simply a *system of relativistic particle mechanics*, for abbreviation, S.R.P.M.):

KINEMATICAL AXIOMS

A1. *P is a nonempty, finite set.*

A2. *If $p \in P$, then $\mathfrak{I}(p)$ is an interval of real numbers.*

A3. *If $p \in P$ and $t \in \mathfrak{I}(p)$, then $s_p(t)$ is an n-dimensional vector; and, moreover, the second derivative of s_p exists throughout the interval $\mathfrak{I}(p)$.*

A4. *The constant c is a positive real number such that for every p in P and t in $\mathfrak{I}(p)$,*

$$|(Ds_p)(t)| < c.$$

DYNAMICAL AXIOMS

A5. *If $p \in P$, then $m(p)$ is a positive real number.*

A6. *If $p \in P$ and $t \in \mathfrak{I}(p)$, then $f(p, t, 1)$, $f(p, t, 2), \cdots$ are n-dimensional vectors such that the series*

$$\sum_{i=1}^{\infty} f(p, t, i)$$

is absolutely convergent.

A7. *If $p \in P$ and $t \in \mathfrak{I}(p)$, then*

$$m(p)\left[D\,\frac{Ds_p}{\sqrt{1-|Ds_p|^2/c^2}}\right](t) = \sqrt{1-\frac{|(Ds_p)(t)|^2}{c^2}}\cdot\sum_{i=1}^{\infty} f(p, t, i).$$

Since this set of axioms is similar in many ways to that given for classical mechanics in [7], a large number of remarks to be found in § 3 of that paper are also applicable here and will not be repeated. From Axiom A7 it is clear that the force concept we are using is that of Minkowski. In the solution of special problems this concept is not always the most useful one, but the relative simplicity of its transformation properties more than justifies its use here. Some readers may feel that there are good physical grounds for taking the notion of relativistic mass as primitive instead of that of rest mass; however, it is easy to define the notion of relativistic mass in terms of the notion of rest mass and our other primitives, and the use of the notion of rest mass as a primitive emphasizes the considerable formal similarity between our axioms for relativistic mechanics and the axioms in [7] for classical mechanics.

For p in P, $\mathfrak{J}(p)$ is a time interval for the particle p (with respect to the *frame of reference fixed by our choice of primitives*). *It may seem that it would* have been simpler to take $\mathfrak{J}(p)$ as the interval of proper time of the particle p. However, this approach would complicate the treatment of *systems* of particles. In the main, the notion of proper time is most convenient in discussions restricted to the consideration of a single particle. From the remark in the previous section it is clear that it is not reasonable to require that the intervals $\mathfrak{J}(p)$ be overlapping. A second argument against such an assumption is the prominence in modern physics of elementary particles with very short life-spans.[1] We note, however, that in studying certain special problems, such as that of defining a reasonable notion of center of mass of a S.R.P.M., it is desirable to restrict the discussion to systems in which $\mathfrak{J}(p) = (-\infty, +\infty)$ for every p in P.

If (i) "c" is replaced by "$1/k$" in the inequality of Axiom A4 and the equation of Axiom A7, (ii) k is treated as a primitive replacing c, and (iii) Axiom A4 is modified to read: "The constant k is a nonnegative real number such that⋯," then, by adding appropriate further axioms, we can get either classical or relativistic particle mechanics. Thus an additional axiom asserting that $k = 0$ gives us classical mechanics; and the assertion that $k > 0$ gives us relativistic mechanics.

[1] Paulette Destouches-Février [2, pp. 5-6] advocates the use of a three-valued logic to describe the creation and annihilation of elementary particles. Actually, the situation is easily handled by the simple device of introducing the function $\mathfrak{J}$ defined on P instead of a fixed interval T for the whole system. Indeed, to our mind, her drastic proposal cannot be taken seriously until we know a great deal more about the mathematics which goes with a multi-valued logic. Even if such a body of mathematics existed (as it does not—we do not have even the general outlines of elementary set theory in three-valued logic), it would be reasonable to adopt such a proposal only after every feasible alternative in standard mathematics had been explored.

We close this section with a number of definitions which will be useful later.

For p in P and t in $\mathfrak{J}(p)$, we set

$$v_p(t) = (Ds_p)(t);$$

$v_p(t)$ is, of course, the *velocity* of p at time t. With respect to a fixed element t_0 in $\mathfrak{J}(p)$, we define the function τ_{t_0} (for p in P and t in $\mathfrak{J}(p)$) as follows:

$$\tau_{t_0}(p, t) = \int_{t_0}^{t} \sqrt{1 - \frac{|v_p(t)|^2}{c^2}}\, dt;$$

$\tau_{t_0}(p, t)$ is the *proper time* of p. Since we are interested only in the derivative of this function with respect to t, and since the derivative is independent of t_0, we shall usually drop the subscript.

For p in P and t in $\mathfrak{J}(p)$, we define the function q as follows:

$$q(p, t) = \langle s(p, t), t\rangle.$$

It is natural to call q the *space-time* function.

For p in P, t in $\mathfrak{J}(p)$, and i any positive integer, we define what we call the *relativistic force* function f^{rel} as follows:

$$f^{\mathrm{rel}}(p, t, i) = \left\langle f(p, t, i), \frac{f(p, t, i) \cdot v_p(t)}{c^2}\right\rangle$$

Although it is not usual to adopt a special name for this function, the function itself is used frequently in textbook treatments of relativity.

By a *c-particle path* (for any positive number c) we mean a set $\mathcal{S}$ of points (that is, vectors) in E_{n+1} for which there exists a S.R.P.M. $\langle \{1\}, \mathfrak{J}, m, s, f, c\rangle$ such that for every point X of E_{n+1}, X is in $\mathcal{S}$ if and only if there exists a t in $\mathfrak{J}(1)$ such that $X = \langle s(1, t), t\rangle$.[2] It is obvious that if g is any twice-differentiable function defined on an interval T of real numbers and taking vectors in E_n as values, then the set of vectors $\langle g(t), t\rangle$ for t in T is a c-particle path, provided that $|(Dg)(t)| < c$ for all t in T.

By the *slope* of a line α in E_{n+1}, whose projection on the $(n+1)$st-axis

[2] The intuitive interpretation of E_{n+1} is as the space-time manifold of special relativity with the $(n+1)$st coordinate representing the time coordinate. Thus, if $\langle Z, x\rangle$ is a point of E_{n+1}, then under the intended interpretation, the n-dimensional vector Z gives the spatial coordinates of the point and x its time coordinate.

is a nondegenerate segment, we mean the n-dimensional vector $\mathbb{W}$ such that for any two distinct points $\langle Z_1, x_1\rangle$ and $\langle Z_2, x_2\rangle$ of α,

$$\frac{Z_1 - Z_2}{x_1 - x_2} = \mathbb{W}.$$

By the speed of α we mean the nonnegative number $|\mathbb{W}|$. By a *c-inertial path* we mean a line in E_{n+1} whose speed is less than c. We note that every segment of a c-inertial path is a c-particle path, but is not necessarily a c-inertial path (since a c-inertial path must be a whole line). By a *c-line* we mean a line in E_{n+1} whose speed is equal to c. The notion of a c-line corresponds to the intuitive notion of a light line.

If we want to refer to a S.R.P.M. Γ with numerical constant c, we shall write: S.R.P.M. Γ_c.

4. Transformation theorems. We begin by defining the notion of a *generalized Lorentz matrix*. An intuitive discussion of such matrices follows Theorem 1.

DEFINITION 1. Let c, c', and λ be positive real numbers. Then a matrix $\mathfrak{A}$ of order $n+1$ is said to be a *generalized Lorentz matrix with respect to* $\langle c, c', \lambda\rangle$ if and only if there exist numbers δ and β, an n-dimensional vector U, and an orthogonal matrix $\mathcal{E}$ of order n, such that

$$\delta^2 = 1, \qquad \beta^2\left(1 - \frac{U^2}{c'^2}\right) = 1,$$

and

$$\text{(i)} \qquad \mathfrak{A} = \lambda \begin{pmatrix} \mathfrak{J} & 0 \\ 0 & \dfrac{c}{c'} \end{pmatrix} \begin{pmatrix} \mathcal{E} & 0 \\ 0 & \delta \end{pmatrix} \begin{pmatrix} \mathfrak{J} + \dfrac{\beta - 1}{U^2}\, U^*U & -\dfrac{\beta U^*}{c'^2} \\ -\beta U & \beta \end{pmatrix}.$$

The following two lemmas simplify the statement and proof of Theorem 1.

LEMMA 1. *Let* $\langle \{1\}, \mathfrak{I}, m, s, f, c\rangle$ *be a* S.R.P.M., *let* c' *and* λ *be positive real numbers, and let* $\mathfrak{A}$ *be a generalized Lorentz matrix with respect to* $\langle c, c', \lambda\rangle$. *Let the function* h *be defined by the equation (for every* t *in* $\mathfrak{I}(1)$*):*

$$h(t) = [\langle s_1(t), t\rangle \mathfrak{A}]_{n+1}.$$

Then the function Dh exists; its values are either always positive or always negative; and the function h is one-to-one.

Proof. From Definition 1 and the hypothesis of the lemma we see that there are numbers δ and β, an n-dimensional vector U, and an orthogonal matrix $\mathcal{E}$, such that

$$\delta^2 = 1, \qquad \beta^2\left(1 - \frac{U^2}{c'^2}\right) = 1,$$

and

$$\mathcal{A} = \begin{pmatrix} \lambda\left(\mathcal{E} + \dfrac{(\beta-1)\mathcal{E}U^*U}{U^2}\right) & -\dfrac{\lambda\beta\mathcal{E}U^*}{c'^2} \\ -\dfrac{\lambda c\delta\beta U}{c'} & \dfrac{\lambda c\delta\beta}{c'} \end{pmatrix}.$$

Thus

$$h(t) = \frac{\lambda c\delta\beta t}{c'} - \frac{\lambda\beta s_1(t)\mathcal{E}U^*}{c'^2} = \left(\frac{\lambda c\delta\beta}{c'}\right)\left(t - \frac{\delta s_1(t)\mathcal{E}U^*}{cc'}\right).$$

Hence

$$\frac{c'(Dh)(t)}{\delta\lambda\beta c} = 1 - \frac{v_1(t)\delta\mathcal{E}U^*}{cc'} \geq 1 - \frac{|v_1(t)\mathcal{E}|\,|U|}{cc'}.$$

Using Axiom A4 and the fact that $\mathcal{E}$ is orthogonal, we have

$$\frac{c'(Dh)(t)}{\delta\lambda\beta c} \geq 1 - \frac{|U|}{c'}.$$

Since $|U| < c'$, the function Dh is bounded away from zero, and it thus follows from Rolle's theorem that h is one-to-one.

The following lemma is a theorem of matrix theory.

LEMMA 2. *Let c, c', and λ be positive real numbers. Then a matrix $\mathcal{A}$ of order $n+1$ is a generalized Lorentz matrix with respect to $\langle c, c', \lambda\rangle$ if and only if*

(i) $$\mathfrak{A}\begin{pmatrix} \mathfrak{I} & 0 \\ 0 & -c'^2 \end{pmatrix}\mathfrak{A}^* = \lambda^2\begin{pmatrix} \mathfrak{I} & 0 \\ 0 & -c^2 \end{pmatrix}.$$

Proof. The proof of necessity is obtained by direct application of Definition 1.

For the proof of sufficiency, let

$$\mathfrak{A} = \begin{pmatrix} \mathfrak{H} & K^* \\ L & m \end{pmatrix},$$

where $\mathfrak{H}$ is a matrix of order n, K and L are n-dimensional vectors, and m is a real number. From (i) we obtain at once:

(1) $$\mathfrak{H}\mathfrak{H}^* - c'^2 K^*K = \lambda^2\mathfrak{I},$$

(2) $$\mathfrak{H}L^* - c'^2 mK^* = 0,$$

(3) $$LL^* - c'^2 m = -\lambda^2 c^2.$$

From (3) it follows that

(4) $$m \neq 0.$$

We define:

(5) $$\beta = \frac{c'|m|}{c\lambda},$$

(6) $$\delta = \frac{m}{|m|},$$

(7) $$U = -\frac{L}{m},$$

(8) $$\mathfrak{E} = \frac{1}{\lambda}\left(\mathfrak{H} + \frac{(\beta - 1)c'^2 K^*U}{\beta U^2}\right).$$

Since the right member of equation (i) of Definition 1 can be written

$$\begin{pmatrix} \lambda\left(\mathcal{E} + \dfrac{(\beta-1)\mathcal{E}U^*U}{U^2}\right) & -\dfrac{\lambda\beta\mathcal{E}U^*}{c'^2} \\ -\dfrac{\lambda c\delta\beta U}{c'} & \dfrac{\lambda c\delta\beta}{c'} \end{pmatrix},$$

in order to complete the proof it suffices to show that

$$\text{(I)} \qquad \lambda\left(\mathcal{E} + \frac{(\beta-1)\mathcal{E}U^*U}{U^2}\right) = \mathfrak{H},$$

$$\text{(II)} \qquad -\frac{\lambda\beta\mathcal{E}U^*}{c'^2} = K^*,$$

$$\text{(III)} \qquad -\frac{\lambda c\delta\beta U}{c'} = L,$$

$$\text{(IV)} \qquad \frac{\lambda c\delta\beta}{c'} = m,$$

$$\text{(V)} \qquad \delta^2 = 1,$$

$$\text{(VI)} \qquad \beta^2\left(1 - \frac{U^2}{c'^2}\right) = 1,$$

$$\text{(VII)} \qquad \mathcal{E}\mathcal{E}^* = \mathfrak{I}.$$

Equation (III) follows immediately from (5), (6), and (7), equation (IV) from (5) and (6), equation (V) from (6), and equation (VI) from (3), (5), and (7).

From (2) and (7) we get

$$\text{(9)} \qquad \mathfrak{H}U^* = -c'^2K^*,$$

and then from (8) and (9) we have

$$\text{(10)} \qquad \mathcal{E}\mathcal{E}^* = \frac{1}{\lambda^2}\left(\mathfrak{H} + \frac{(\beta-1)c'^2K^*U}{\beta U^2}\right)\left(\mathfrak{H}^* + \frac{(\beta-1)c'^2U^*K}{\beta U^2}\right)$$

$$= \frac{1}{\lambda^2}\left\{\mathfrak{H}\mathfrak{H}^* + \frac{(\beta-1)c'^2}{\beta^2U^2}\left[\beta(-c'^2K^*)K + \beta K^*(-c'^2K) + (\beta-1)c'^2K^*K\right]\right\}$$

$$= \frac{1}{\lambda^2}\left[\mathfrak{H}\mathfrak{H}^* + \frac{(\beta-1)c'^4}{\beta^2 U^2} K^*K(-2\beta+\beta-1)\right]$$

$$= \frac{1}{\lambda^2}\left[\mathfrak{H}\mathfrak{H}^* - \frac{(\beta^2-1)c'^4}{\beta^2 U^2} K^*K\right].$$

From (VI), (1), and (10) we conclude that

$$\mathcal{E}\mathcal{E}^* = \mathfrak{I},$$

which establishes equation (VII). Multiplying both sides of (8) on the right by $-\lambda\beta U^*/c'^2$, and using (9), we get equation (II). Equation (I) follows from (8) and (II), completing the proof of the lemma.

The following theorem is a generalization of the well-known result that the relativistic equation of motion is covariant under a Lorentz transformation.

THEOREM 1. *Let* $\langle P, \mathfrak{T}, m, s, f, c\rangle$ *be an n-dimensional* S.R.P.M. *Let* c', γ, *and* λ *be positive real numbers, let* B *be an* $(n+1)$*-dimensional vector, and let* $\mathfrak{A}$ *be a generalized Lorentz matrix with respect to* $\langle c, c', \lambda\rangle$. *For each* p *in* P *let the function* h_p *be defined as follows (for all* t *in* $\mathfrak{T}(\mathrm{p})$*):*

$$h_p(t) = [\langle s_p(t), t\rangle\, \mathfrak{A} + B]_{n+1}.$$

(By Lemma 1 *the inverse function* h_p^{-1} *exists.) Let the function* $\mathfrak{T}'$ *be defined as follows: for* p *in* P, $\mathfrak{T}'(\mathrm{p})$ *is the range of the function* h_p; *and let the functions* m', s', *and* f' *be defined by the following equations (for* p *in* P, t' *in* $\mathfrak{T}'(\mathrm{p})$ *and* i *in* I*):*

$$m'(\mathrm{p}) = \gamma m(\mathrm{p}),$$

$$s'(\mathrm{p}, t') = [\langle s(\mathrm{p}, h_p^{-1}(t')), h_p^{-1}(t')\rangle\, \mathfrak{A} + B]_{1,\dots,n},$$

$$f'(\mathrm{p}, t', i) = \frac{\gamma c'^2}{\lambda^2 c^2}\left[\left\langle f(\mathrm{p}, h_p^{-1}(t'), i), \frac{f(\mathrm{p}, h_p^{-1}(t'), i)\cdot v_p(h_p^{-1}(t'))}{c^2}\right\rangle \mathfrak{A}\right]_{1,\dots,n}.$$

Then $\Gamma' = \langle P, \mathfrak{T}', m', s', f', c'\rangle$ *is an n-dimensional* S.R.P.M.

Proof. It will suffice to show that Γ' satisfies Axioms A4 and A7, since the proof for the other axioms is trivial. Let

$$\mathfrak{A} = \begin{pmatrix} \mathfrak{D} & E^* \\ F & g \end{pmatrix}.$$

It is easy to show that for p in P, and t' in $\mathfrak{T}'(p)$,

$$(1) \qquad v_p'(t') = \frac{\langle v_p(h_p^{-1}(t'), 1\rangle \begin{pmatrix} \mathfrak{D} \\ F \end{pmatrix}}{\langle v_p(h_p^{-1}(t'), 1\rangle \begin{pmatrix} E^* \\ g \end{pmatrix}},$$

with the denominator of the right member of (1) always unequal to zero. (Since in this proof we always consider a fixed particle p, we drop the subscript "p" from this point on.)

We have, from Axiom A4,

$$(2) \qquad \lambda^2(|v(h^{-1}(t'))|^2 - c^2) < 0;$$

but

$$\lambda^2(|v(h^{-1}(t'))|^2 - c^2) = \lambda^2 \langle v(h^{-1}(t')), 1\rangle \begin{pmatrix} \mathfrak{I} & 0 \\ 0 & -c^2 \end{pmatrix} \langle v(h^{-1}(t')), 1\rangle^*.$$

Then by Lemma 2 we have

$$(3) \quad \lambda^2(|v(h^{-1}(t'))|^2 - c^2) = \langle v(h^{-1}(t')), 1\rangle \, \mathfrak{A} \begin{pmatrix} \mathfrak{I} & 0 \\ 0 & -c'^2 \end{pmatrix} \mathfrak{A}^* \langle v(h^{-1}(t')), 1\rangle^*.$$

The right member of (3) is equal to

$$(4) \qquad \langle v(h^{-1}(t')), 1\rangle \begin{pmatrix} \mathfrak{D} \\ F \end{pmatrix} (\mathfrak{D}^* F^*) \langle v(h^{-1}(t')), 1\rangle^*$$
$$- c'^2 \langle v(h^{-1}(t')), 1\rangle \begin{pmatrix} E^* \\ g \end{pmatrix} (E g) \langle v(h^{-1}(t')), 1\rangle^*,$$

and using (1) we see that (4) is equal to

$$(5) \qquad \left(\langle v(h^{-1}(t')), 1\rangle \begin{pmatrix} E^* \\ g \end{pmatrix} v'(t')\right)^2 - c'^2 \left(\langle v(h^{-1}(t'), 1\rangle \begin{pmatrix} E^* \\ g \end{pmatrix}\right)^2.$$

From (2), (3), (4), and (5) we conclude that

$$|v'(t')|^2 - c'^2 < 0,$$

which verifies Axiom A4 for Γ'.

It is not difficult to show that from Axiom A7 we have

$$\frac{m(\mathrm{p})}{(1-|v(t)|^2/c^2)^{1/2}} D\left[\frac{Dq}{(1-|v|^2/c^2)^{1/2}}\right](t) = \sum_{i=1}^{\infty} f^{\mathrm{rel}}(\mathrm{p}, t, i). \tag{6}$$

Setting $q'(t') = \langle s'(t'), t'\rangle$ for all t' in $\mathfrak{I}'(\mathrm{p})$, we conclude from the hypothesis of our theorem that

$$q'(h(t)) = q(t)\mathfrak{A} + B,$$

and thus

$$((Dq')\circ h)(t)(Dh)(t) = (Dq)(t)\mathfrak{A}. \tag{7}$$

Directly from the definition of q and q' we obtain

$$(Dq)(t)\begin{pmatrix} \mathfrak{I} & 0 \\ 0 & -c^2 \end{pmatrix}((Dq)(t))^* = |v(t)|^2 - c^2, \tag{8}$$

and

$$((Dq')\circ h)(t)\begin{pmatrix} \mathfrak{I} & 0 \\ 0 & -c'^2 \end{pmatrix}(((Dq')\circ h)(t))^* = |(v'\circ h)(t)|^2 - c'^2. \tag{9}$$

Using Lemma 1, Lemma 2, and (7) we obtain, from (8) and (9),

$$|(v'\circ h)(t)|^2 - c'^2 = \frac{\lambda^2}{((Dh)(t))^2}(|v(t)|^2 - c^2),$$

and thus

$$1 - \frac{|(v'\circ h)(t)|^2}{c'^2} = \frac{\lambda^2 c^2}{c'^2((Dh)(t))^2}\left(1 - \frac{|v(t)|^2}{c^2}\right). \tag{10}$$

By Lemma 1, $(Dh)(t)$ is either always positive or always negative; the remainder of our proof is analogous in the two cases, so that we shall only consider the case where it is always positive. We then have, from (7) and (10),

$$\frac{[(Dq')\circ h](t)}{(1-|(v'\circ h)(t)|^2/c'^2)^{1/2}} = \frac{c'(Dq)(t)\mathfrak{a}}{\lambda c(1-|v(t)|^2/c^2)^{1/2}},$$

and hence

$$(11)\qquad \left(\left[\frac{Dq'}{(1-|v'|^2/c'^2)^{1/2}}\right]\circ h\right)(t) = \frac{c'(Dq)(t)\mathfrak{a}}{\lambda c(1-|v(t)|^2/c^2)^{1/2}}.$$

Differentiating both sides of (11), and using (6), we obtain

$$(12)(Dh)(t)\left(\left[D\left(\frac{Dq'}{(1-|v'|^2/c'^2)^{1/2}}\right)\right]\circ h\right)(t) = \frac{c'}{\lambda c} D\left(\frac{Dq}{(1-|v|^2/c^2)^{1/2}}\right)(t)\mathfrak{a}$$

$$= \left(\sqrt{1-\frac{|v(t)|^2}{c^2}}\right)\left(\frac{c'}{\lambda cm(p)}\right)\left(\sum_{i=1}^{\infty} f^{\mathrm{rel}}(p,t,i)\mathfrak{a}\right).$$

From (10), (12), and the hypothesis of our theorem, we infer that

$$(13)\qquad \gamma m(p)\left(\left[D\left(\frac{Dq'}{(1-|v'|^2/c'^2)^{1/2}}\right)\right]\circ h\right)(t)$$

$$= \sqrt{1-\frac{|(v'\circ h)(t)|^2}{c'^2}}\ \sum_{i=1}^{\infty} f^{\mathrm{rel}\prime}(p,h(t),i),$$

and from (13) we conclude immediately that Axiom A7 holds for Γ'.[3]

REMARK 1. All the transformations mentioned in Definition 1 and Theorem 1 have a clear intuitive interpretation if we consider $\langle P, \mathfrak{I}, m, s, f, c\rangle$ as a physical system whose mechanical properties are observed and measured with respect to some (inertial) frame of reference and some set of units of measurement, and $\langle P, \mathfrak{I}', m', s', f', c'\rangle$ as the same physical system observed and measured with respect to some other (inertial) frame of reference and some other set of units of measurement. Thus, c is the old and c' the new velocity of light. The introduction of the number γ amounts to changing the unit of mass by an amount $1/\gamma$, and the vector B corresponds to shifting the origin of the spatial frame of reference by $-[B]_{1,\ldots,n}$, and the origin of time by an amount

[3]Readers familiar with the standard treatments of relativistic mechanics will note that (in the interests of rigor and explicitness) we have replaced "t'" by "$h_p(t)$."

$-[B]_{n+1}$. The number λ represents a uniform stretch of space and time. When $\delta = -1$, we have a reversal of the direction of time. The matrix $\mathcal{E}$ represents (for $n \leq 3$) a rotation of the spatial coordinates—or a rotation followed by a reflection. The vector U represents the relative velocity of the two inertial frames of reference, and the number β, which is determined by U and c', is the well-known Lorentz contraction factor. Finally, it is easy to check that the last matrix in the factorization of the matrix $\mathcal{A}$ yields the ordinary Lorentz transformations. We note that the rather complicated transformation of the forces is the velocity-dependent transformation to be expected in relativistic mechanics.

REMARK 2. Theorem 2, our main result, is a sort of converse of Theorem 1: roughly speaking, we show that the transformations described in Theorem 1 are the only transformations which always take systems of relativistic particle mechanics into systems of relativistic particle mechanics. To facilitate the formulation and proof of Theorem 2, an additional lemma and some definitions will be useful.

LEMMA 3. *Let $X_1 = \langle Z_1, x_1 \rangle$, $X_2 = \langle Z_2, x_2 \rangle$, and $X_3 = \langle Z_3, x_3 \rangle$ be any three points in E_{n+1} such that* (i) $x_1 < x_2 < x_3$, (ii) *there is a c-inertial path through X_1 and X_2, and* (iii) *there is a c-inertial path through X_2 and X_3. Then there is a c-particle path through X_1, X_2, and X_3.*

Proof. In view of the remark near the end of § 3, it will suffice to construct a function g which: (a) is defined on the closed interval $[x_1, x_3]$; (b) takes vectors in E_n as values; (c) is twice differentiable; (d) is such that for every t in $[x_1, x_3]$, $|(Dg)(t)| < c$; and (e) is such that

$$g(x_1) = Z_1, \quad g(x_2) = Z_2, \quad \text{and} \quad g(x_3) = Z_3.$$

Let

$$a = x_2 - x_1, \qquad b = x_3 - x_2,$$

$$V = \frac{Z_2 - Z_1}{a}, \qquad W = \frac{Z_3 - Z_2}{b},$$

$$\gamma = \frac{2c \log 2}{(c - \max(|V|, |W|)) \min(a, b)}, \qquad A = \frac{aV \log\cosh\gamma b + bW \log\cosh\gamma a}{a \log\cosh\gamma b + b \log\cosh\gamma a},$$

$$B = \frac{ab\gamma(W - V)}{a \log\cosh\gamma b + b \log\cosh\gamma a}.$$

The reader may verify that the function g defined by the following equation (for t in $[x_1, x_3]$) has properties $(a)-(e)$:

$$g(t) = Z_2 + (t - x_2)A + \frac{B}{\gamma}[\log\cosh\gamma(t - x_2)].$$

DEFINITION 2. Let ϕ_1 be a function mapping R^+ into R^+; let ϕ_2 be a function which is a one-to-one mapping of E_{n+1} into itself; and let ϕ_3 be a function mapping E_{2n} into E_n. Then we call the ordered triple $\langle \phi_1, \phi_2, \phi_3 \rangle$ an *eligible transformation.*

DEFINITION 3. Let $\Phi = \langle \phi_1, \phi_2, \phi_3 \rangle$ be an eligible transformation; let $\Gamma = \langle P, \mathfrak{J}, m, s, f, c \rangle$ be a S.R.P.M.; and for each p in P let the function H_p be defined as follows (for every t in $\mathfrak{J}(p)$):

$$H_p(t) = [\phi_2(s(p, t), t)]_{n+1}.$$

Then by the Φ-*transform* of Γ (which we also write: $\Phi(\Gamma)$), we mean the ordered quintuple $\langle P, \mathfrak{J}', m', s', f' \rangle$, where for p in P:

$$m'(p) = \phi_1(m(p));$$

$\mathfrak{J}'(p)$ is the range of the function H_p; and s' and f' are defined by the following equations for t' in $\mathfrak{J}'(p)$, if the pre-image $H_p^{-1}(t')$ of t' under H_p is unique, and otherwise they are undefined:

$$s'(p, t') = [\phi_2(s(p, H_p^{-1}(t'), H_p^{-1}(t')]_{1,\ldots,n}$$

$$f'(p, t', i) = \phi_3(f(p, H_p^{-1}(t'), i)v(p, H_p^{-1}(t'))),$$

for $i \geq 1$.

We are now in a position to state and prove the main result of this paper.

THEOREM 2.[4] *Let* $\Phi = \langle \phi_1, \phi_2, \phi_3 \rangle$ *be an eligible transformation, and let*

[4] The statement of Theorem 2 would be made more symmetrical to Theorem 1 if ϕ_2 were replaced by two functions ϕ' and ϕ'' such that

$$\phi'(Z, x) = [\phi_2(Z, x)]_{n+1} \quad \text{and} \quad \phi''(Z, x) = [\phi_2(Z, x)]_{1,\ldots,n}.$$

This procedure was followed in [8] for classical mechanics; but in relativistic mechanics, it is natural to introduce the single transformation ϕ_2 for the space-time manifold.

c and c' be positive real numbers such that (i) *for every n-dimensional system of relativistic particle mechanics* Γ_c, $\langle \Phi(\Gamma_c), c' \rangle$ *is a system of relativistic particle mechanics, and* (ii) ϕ_2 *carries no c-line into a c'-particle path. Then there exist positive real numbers* γ *and* λ, *an* $(n+1)$*-dimensional vector* B, *and a generalized Lorentz matrix* $\mathfrak{A}$ *with respect to* $\langle c, c', \lambda \rangle$, *such that, for any vectors* Z_1 *and* Z_2 *in* E_n *with* $|Z_2| < c$, *every* x *in* R, *and* y *in* R^+,

$$\phi_1(y) = \gamma y,$$

$$\phi_2(Z_1, x) = \langle Z_1, x \rangle \mathfrak{A} + B,$$

$$\phi_3(Z_1, Z_2) = \frac{\gamma c'^2}{\lambda^2 c^2} \left[\left\langle Z_1, \frac{Z_1 \cdot Z_2}{c^2} \right\rangle \mathfrak{A} \right]_{1,\dots,n}.$$

Proof. We first want to show that if Z is any vector in E_n such that $|Z| < c$, then

$$\phi_3(0, Z) = 0.$$

Setting $P = \{1\}$, $\mathfrak{I}(1) = (-\infty, \infty)$, $m(1) = 1$, and, for t in $\mathfrak{I}(1)$,

$$s(1, t) = Zt,$$

$$f(1, t, i) = 0 \quad \text{for} \quad i \geq 1,$$

we see that $\langle P, \mathfrak{I}, m, s, f, c \rangle$ is a S.R.P.M. Since for every t in $\mathfrak{I}(1)$, $Z = v(1, t)$, we conclude from the hypothesis of our theorem, Definition 2, and Axiom A6 that the series

$$\phi_3(0, Z) + \phi_3(0, Z) + \cdots$$

is absolutely convergent. Hence

$$\phi_3(0, Z) = 0. \tag{1}$$

For every segment $\mathcal{S}$ of a c-inertial path there exists a one-particle S.R.P.M. $\langle \{1\}, \mathfrak{I}, m, s, f, c \rangle$ such that, for every t in $\mathfrak{I}(1)$,

$$f(p, t, i) = 0 \quad \text{for} \quad i \geq 1,$$

and for every vector X in E_{n+1}, X is in $\mathcal{S}$ if and only if there is a t in $\mathfrak{I}(1)$ such that

$$X = \langle s(1,t), t \rangle .$$

Hence it follows immediately from (1) and the hypothesis of our theorem that:

(2) ϕ_2 carries segments of c-inertial paths into segments of c'-inertial paths.

Let $\Gamma_c = \langle P, \mathfrak{I}, m, s, f, c' \rangle$ be any S.R.P.M. with constant c. By hypothesis, $\langle \Phi(\Gamma_c), c \rangle$ is a S.R.P.M. For any p in P, if t_1 and t_2 are in $\mathfrak{I}(p)$ and $t_1 \neq t_2$, then

$$\phi_2(s(p, t_1), t_1) \neq \phi_2(s(p, t_2), t_2),$$

since ϕ_2 is one-to-one. Suppose now that

$$[\phi_2(s(p, t_1), t_1)]_{n+1} = [\phi_2(s(p, t_2), t_2)]_{n+1} .$$

Then we must have

$$[\phi_2(s(p, t_1), t_1)]_{1,\dots,n} \neq [\phi_2(s(p, t_2), t_2)]_{1,\dots,n} ;$$

but then $\langle \Phi(\Gamma_c), c' \rangle$ is not a S.R.P.M., for p is required to be in two places at the same time, which violates Axiom A3. We thus conclude:

(3) ϕ_2 is one-to-one in the last coordinate along the space-time path of any particle of a S.R.P.M. also, Γ_c, and thus the pre-image under ϕ_2 of any point t' in $\mathfrak{I}'(p)$, is unique.

Furthermore, since by hypothesis ϕ_2 takes the interval $\mathfrak{I}(p)$ into an interval $\mathfrak{I}'(p)$, we have:

(4) ϕ_2 is continuous in the last coordinate along the space-time path of any particle of a S.R.P.M.

From (4) and the fact that any two points $\langle Z, x \rangle$ and $\langle Z, y \rangle$ lie on a c-inertial path, we obtain:

(5) For any point $\langle Z, x \rangle$ and any $\epsilon > 0$, there exists a $\delta > 0$ such that for any point $\langle Z, y \rangle$ if $|x - y| < \delta$, then $|x' - y'| < \epsilon$, where

$$x' = [\phi_2(Z, x)]_{n+1} \quad \text{and} \quad y' = [\phi_2(Z, y)]_{n+1} .$$

We next show that

(I) ϕ_2 is continuous.

Let $\langle Z_1, x_1\rangle$ be any point of E_{n+1}, and let ϵ be any positive number. Let $\epsilon^* = \epsilon/[2(1+c')]$. Using (5), let δ^* be a positive number such that if $|x_1 - y| < \delta^*$ then $|x_1' - y'| < \epsilon^*$, where

$$x_1' = [\phi_2(Z_1, x_1)]_{n+1} \quad \text{and} \quad y' = [\phi_2(Z_1, y)]_{n+1} ;$$

and let $\delta = c\delta^*/(3c+2)$. We shall show that if $\langle Z_2, x_2\rangle$ is any point of E_{n+1} such that

$$(6) \qquad |\langle Z_1, x_1\rangle - \langle Z_2, x_2\rangle| < \delta ,$$

then

$$|\phi_2(Z_1, x_1) - \phi_2(Z_2, x_2)| < \epsilon .$$

Suppose for definiteness that

$$(7) \qquad x_1 \geq x_2 .$$

We may choose x_0 and x_3 so that

$$(8) \qquad x_2 - \frac{|Z_2 - Z_1|}{c} - \delta < x_0 < x_2 - \frac{|Z_2 - Z_1|}{c}$$

and

$$(9) \qquad x_1 + \frac{|Z_2 - Z_1|}{c} < x_3 < x_1 + \frac{|Z_2 - Z_1|}{c} + \delta$$

From (7), (8), and (9), we obtain

$$(10) \qquad |x_3 - x_0| < |x_2 - x_1| + 2\,\frac{|Z_2 - Z_1|}{c} + 2\delta ;$$

and from (6) and (10) we then infer that

$$(11) \qquad |x_3 - x_0| < \frac{(3c+2)\delta}{c} = \delta^* .$$

Since from (7), (8), and (9) we have

$$(12) \qquad x_0 < x_2 \leq x_1 < x_3 ,$$

we obtain from (11)

$$|x_3 - x_1| < \delta^*,$$

$$|x_1 - x_0| < \delta^*.$$

Consequently, by (5),

$$|x_3' - x_1'| + |x_1' - x_0'| < 2\epsilon^*,$$

and thus, by the triangle inequality,

$$|x_3' - x_0'| < 2\epsilon^* , \tag{13}$$

where

$$x_0' = [\phi_2(Z_1, x_0)]_{n+1},$$

$$x_3' = [\phi_2(Z_1, x_3)]_{n+1}.$$

From the second part of (8) it follows that there is a c-inertial path through $\langle Z_1, x_0\rangle$ and $\langle Z_2, x_2\rangle$; and from (7) and the first part of (9) it follows that there is a c-inertial path through $\langle Z_2, x_2\rangle$ and $\langle Z_1, x_3\rangle$. We thus conclude from Lemma 3 that there exists a c-particle path through $\langle Z_1, x_0\rangle$, $\langle Z_2, x_2\rangle$, and $\langle Z_1, x_3\rangle$. As before, for abbreviation, we set

$$x_2' = [\phi_2(Z_2, x_2)]_{n+1},$$

$$Z_2' = [\phi_2(Z_2, x_2)]_{1,\dots,n},$$

$$Z_{1i}' = [\phi_2(Z_1, x_i)]_{1,\dots,n} \qquad (i = 0, 1, 3).$$

Since ϕ_2 is one-to-one and continuous in the last coordinate along any c-particle path, it is monotone in the last coordinate along any c-particle path, and we thus have: either

$$\begin{cases} x_0' < x_1' < x_3', \\ x_0' < x_2' < x_3'; \\ \text{or} \\ x_3' < x_1' < x_0', \\ x_3' < x_2' < x_0'. \end{cases} \tag{14}$$

Also, since segments of c-inertial paths are carried by ϕ_2 into segments of c'-inertial paths, we have:

$$(15)\qquad |Z'_{11} - Z'_2| \leq |Z'_{11} - Z'_{13}| + |Z'_{13} - Z'_2| < c'|x'_1 - x'_3| + c'|x'_3 - x'_2|$$

and

$$(16)\qquad |Z'_{11} - Z'_2| \leq |Z'_{11} - Z'_{10}| + |Z'_{10} - Z'_2| < c'|x'_1 - x'_0| + c'|x'_0 - x'_2|.$$

We obtain from (14), (15), and (16):

$$(17)\qquad 2|Z'_{11} - Z'_2| < c'[|x'_3 - x'_1| + |x'_1 - x'_0| + |x'_3 - x'_2| + |x'_2 - x'_0|]$$

$$< 2c'|x'_3 - x'_0|.$$

Thus from (13) and (17) we conclude that

$$|Z'_{11} - Z'_2| < 2c'\epsilon^*,$$

and from (13) and (14) that

$$|x'_1 - x'_2| < 2\epsilon^*;$$

and since $\epsilon^* = \epsilon/[2(1 + c')]$, we infer that

$$|\phi_2(Z_1, x_1) - \phi_2(Z_2, x_2)| < \epsilon,$$

which establishes (I).

We now establish:

(II) ϕ_2 carries parallel segments of c-inertial paths into parallel segments of c'-inertial paths.

It is clearly sufficient to show that ϕ_2 carries parallel c-inertial paths into parallel segments of c'-inertial paths. Let η_1 and η_2 be two parallel c-inertial paths, and let η_3 be a c-inertial path which intersects η_1 and η_2 in the points A_1 and A_2, respectively (obviously such a c-inertial path η_3 exists). (See Figure 1, on following page.) As previously, we use a prime to designate the image under ϕ_2 of a point, line, and so on. We may construct a fourth c-inertial path which intersects η_3 between A_1 and A_2 and which intersects η_1 and η_2 at

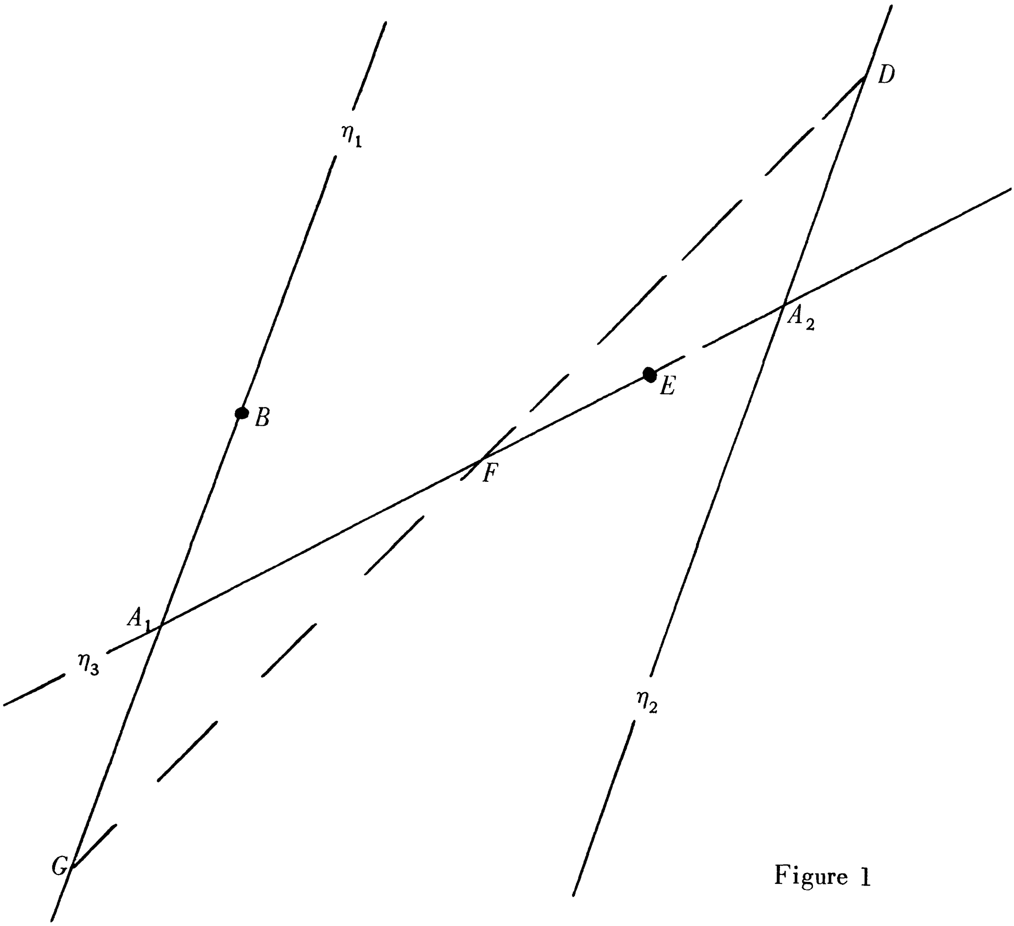

Figure 1

points distinct from A_1 and A_2. Consequently, we infer from (2) that the segments η_1' and η_2' lie in the same plane in the image space of ϕ_2.(See Figure 2 on following page.) Suppose now that η_1' and η_2' are not parallel. We extend (if necessary) η_1' and η_2' to their point of intersection, say J'. We next select B' on η_1' between J' and A_1' (we use "between" in such a way that B' must be distinct from J' and A_1'); similarly, we select D' on η_2' between J' and A_2'. We now consider the pre-images, B and D, of B' and D'. Since ϕ_2 is one-to-one and continuous, it is clear that B and D must be on the same side of η_3; that is, the segment BD does not intersect η_3. Let E be a point on η_3 between A_1 and A_2. Then, since η_3 is a c-inertial path, one of the numbers $[A_2]_{n+1} - [E]_{n+1}$ and $[A_1]_{n+1} - [E]_{n+1}$ is positive, and the other is negative. Since η_1 and η_2 are parallel, $[B]_{n+1} - [A_1]_{n+1}$ and $[D]_{n+1} - [A_2]_{n+1}$ have the same sign. We then construct a line through D' parallel to η_1' or through B' parallel

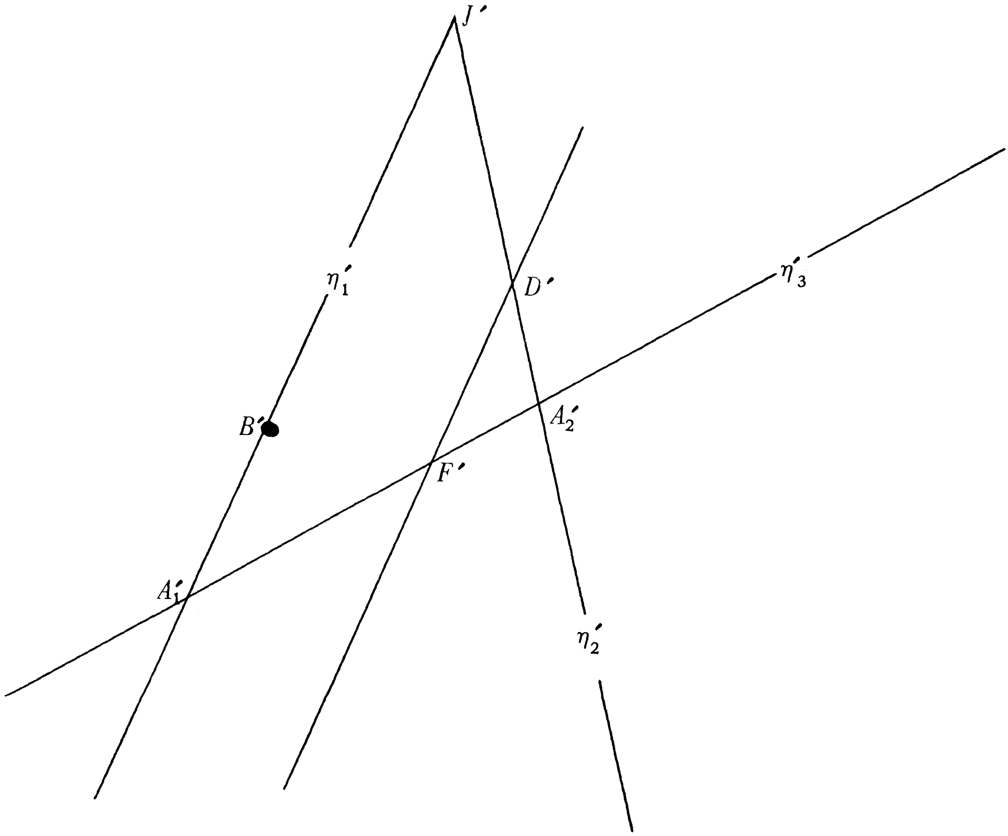

Figure 2

to η_2', according to whether $[A_2]_{n+1} - [E]_{n+1}$ or $[A_1]_{n+1} - [E]_{n+1}$ agrees in sign with $[B]_{n+1} - [A]_{n+1}$. Suppose, for definiteness, (see Figure 1) that $[A_2]_{n+1} - [E]_{n+1}$ agrees in sign and that this sign is positive. Let F' be the point of intersection of η_3' with the line through D' parallel to η_1'. By construction F' is between A_1' and A_2', and thus F is between A_1 and A_2.

We then have:

$$|[D]_{1,\dots,n} - [F]_{1,\dots,n}| \leq |[D]_{1,\dots,n} - [A_2]_{1,\dots,n}| + |[A_2]_{1,\dots,n} - [F]_{1,\dots,n}|$$

$$< c([D]_{n+1} - [A_2]_{n+1}) + c([A_2]_{n+1} - [F]_{n+1})$$

$$< c([D]_{n+1} - [F]_{n+1}).$$

Hence the line through D and F is a c-inertial path. This line intersects η_1 at a point, say G, and, furthermore, by construction DFG is a segment of a c-inertial path, and hence the image $D'F'G'$ is a segment of a c'-inertial path. But $D'F'$ is parallel to η_1', and the image of G does not lie on the extension of $D'F'$, which is a contradiction. Thus η_1 and η_2 are parallel, and the proof of (II) is complete.

We next show that

(III) ϕ_2 carries the midpoint of any finite segment α of a c-inertial path into the midpoint of α'.

We consider a fixed plane containing α and a line parallel to the t-axis (the $(n+1)$st-coordinate axis). In this plane we construct, with α as a diagonal, a parallelogram whose sides and other diagonal are segments of c-inertial paths. Let the speed of the c-inertial path containing α be k. It is clear that through any point of our fixed plane there are exactly two lines with speed l, for every positive number l. Obviously, we may construct a parallelogram $\mathfrak{P}$ with α as one diagonal, with the other diagonal a segment of a c-inertial path with speed $(1/4)(3k+c)$, and with one side a segment of a c-inertial path with speed $(1/2)(k+c)$. The other side of the parallelogram $\mathfrak{P}$ is then a segment of a c-inertial path with speed $(1/6)(5k+c)$. We conclude from (II) that $\mathfrak{P}$ is carried by ϕ_2 into a parallelogram $\mathfrak{P}'$, and the diagonals of $\mathfrak{P}$ are carried into the diagonals of $\mathfrak{P}'$. Hence the midpoint of α is carried into the midpoint of α' and (III) is established.

We next show that

(IV) ϕ_2 carries arbitrary lines into lines.

Let α be an arbitrary line in E_{n+1}, and let $\langle Z_1, x_1\rangle$ and $\langle Z_2, x_2\rangle$ be any two points on α. We now construct an "inertial" parallelogram through these two points. For definiteness, we assume:

$$x_1 \geq x_2 .$$

We set

$$Z_0 = \frac{Z_1 + Z_2}{2},$$

and we choose x_0 and x_3 so that:

$$x_0 < x_2 - \frac{|Z_1 - Z_2|}{2c}, \qquad x_3 > x_1 + \frac{|Z_1 - Z_2|}{2c},$$

$$|\langle Z_1, x_1\rangle - \langle Z_0, x_3\rangle| = |\langle Z_0, x_0\rangle - \langle Z_2, x_2\rangle|,$$

$$|\langle Z_1, x_1\rangle - \langle Z_0, x_0\rangle| = |\langle Z_0, x_3\rangle - \langle Z_2, x_2\rangle|.$$

Let (see Figure 3)

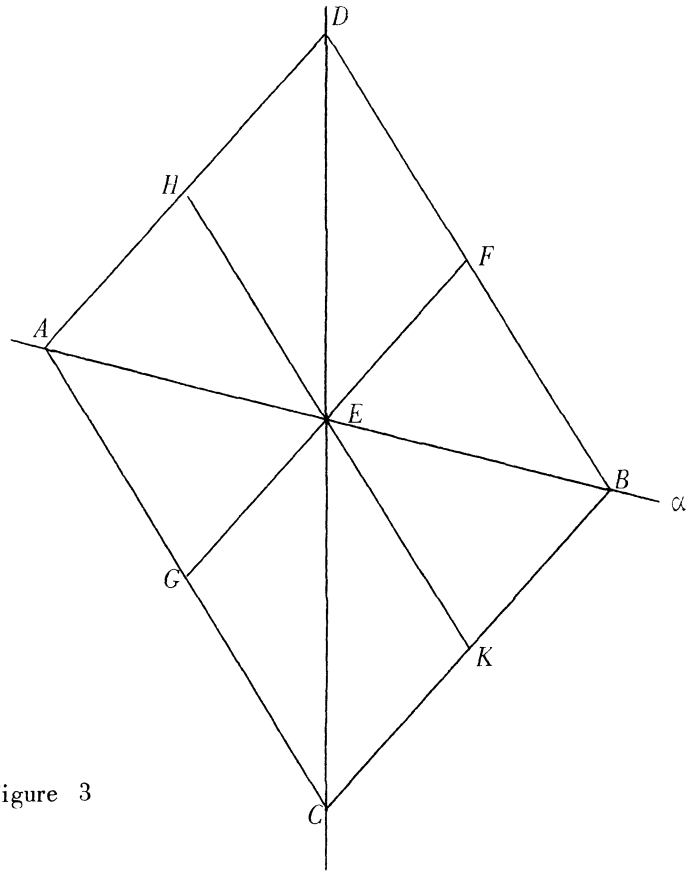

Figure 3

$$A = \langle Z_1, x_1\rangle, \qquad B = \langle Z_2, x_2\rangle, \qquad C = \langle Z_0, x_0\rangle, \qquad D = \langle Z_0, x_3\rangle,$$

$$E = \frac{A+B}{2}, \qquad F = \frac{B+D}{2}, \qquad G = \frac{A+C}{2}, \qquad H = \frac{A+D}{2}, \qquad K = \frac{B+C}{2}.$$

Since the sides of the parallelogram $ACBD$ are by construction segments of c-inertial paths, we conclude from (II) that $A'C'B'D'$ is a parallelogram, where $A' = \phi_2(A)$, and so on, and that the sides of $A'C'B'D'$ are segments of c'-inertial paths. Moreover, it is clear that by construction CED, FEG, and HEK are segments of c-inertial paths, and consequently $C'E'D'$, $F'E'G'$, and

$H'E'K'$ are segments of c'-inertial paths. Hence, by (III), F', G', H', and K' are the midpoints of the respective sides of $A'C'B'D'$. Thus E', the point of intersection of the segments $F'G'$ and $H'K'$, is the point of intersection of the diagonals of $A'C'B'D'$. Consequently, E' is the midpoint of the segment $A'B'$. Since midpoints of finite segments are carried into midpoints of finite segments, and ϕ_2 is continuous, the proof of (IV) is complete.

From (IV) and the fact that ϕ_2 is one-to-one and continuous, we immediately infer that ϕ_2 is a projective transformation, and since it takes no finite point into a point at infinity, we conclude that

(18) ϕ_2 is a nonsingular affine transformation; that is, for every point $\langle Z, x\rangle$ in E_{n+1},

(19) $$\phi_2(Z,x) = \langle Z, x\rangle\, \mathfrak{A} + B,$$

where $\mathfrak{A}$ is a nonsingular matrix of order $n+1$ and B is an $(n+1)$-dimensional vector.

Now let

(20) $$\mathfrak{A} = \begin{pmatrix} \mathfrak{D} & E^* \\ F & g \end{pmatrix} \quad \text{and} \quad B = \langle B_1, b\rangle,$$

where $\mathfrak{D}$ is a matrix of order n; B_1, E, and F are n-dimensional vectors; and b and g are real numbers. Then

(21) $$\phi_2(Z,x) = \langle Z\mathfrak{D} + xF + B_1,\ ZE^* + gx + b\rangle.$$

Let α be a c-line such that, for any two distinct points $\langle Z_1, x_1\rangle$ and $\langle Z_2, x_2\rangle$ of α,

$$\frac{Z_1 - Z_2}{x_1 - x_2} = W.$$

Obviously, $|W| = c$. Now α is carried by ϕ_2 into a line α'. We want to show that α is carried into a c'-line. From (21) it follows that the slope W' of α' is given by

(22) $$W' = \frac{W\mathfrak{D} + F}{WE^* + g}.$$

By the hypothesis of our theorem,

$$|\mathbb{W}'| \geq c'. \tag{23}$$

Consider now a sequence of c-inertial lines $\alpha_1, \alpha_2, \cdots$, whose slopes $\mathbb{W}_1, \mathbb{W}_2, \cdots$ are such that

$$\lim_{i \to \infty} \mathbb{W}_i = \mathbb{W}.$$

From (21) and the hypothesis of our theorem we have

$$|\mathbb{W}_i'| = \left| \frac{\mathbb{W}_i \mathbb{D} + F}{\mathbb{W}_i E^* + g} \right| < c'.$$

Hence, if $\mathbb{W}E^* + g \neq 0$, then

$$|\mathbb{W}'| = \left| \frac{\mathbb{W}\mathbb{D} + F}{\mathbb{W}E^* + g} \right| = \left| \lim_{i \to \infty} \frac{\mathbb{W}_i \mathbb{D} + F}{\mathbb{W}_i E^* + g} \right| \leq c'. \tag{24}$$

Suppose now that $\mathbb{W}E^* + g = 0$. Then

$$\lim_{i \to \infty} (\mathbb{W}_i E^* + g) = 0,$$

and therefore

$$\lim_{i \to \infty} (\mathbb{W}_i \mathbb{D} + F) = 0.$$

Hence

$$\mathbb{W}\mathbb{D} + F = 0,$$

but then

$$\langle \mathbb{W}, 1 \rangle \mathfrak{A} = 0,$$

which is impossible, since $\mathfrak{A}$ is nonsingular. Thus we have

$$|\mathbb{W}'| = c'. \tag{25}$$

For subsequent use we observe that for any S.R.P.M., $\langle P, \mathfrak{T}, m, s, f, c \rangle$, and any p in P and t in $\mathfrak{T}(p)$,

(26) $$v(p, t)E^* + g \neq 0.$$

For $v(p, t) \neq 0$, the argument is the same as above; in case $v(p, t) = 0$ for some t, on the supposition that $v(p, t)E^* + g = 0$, we must have $g = 0$ and $F = 0$, which again contradicts the nonsingularity of $\mathfrak{A}$.

From (22) and (25) we get

$$\frac{\mathbb{W}\mathbb{D}\mathbb{D}^*\mathbb{W}^* + 2\mathbb{W}\mathbb{D}F^* + |F|^2}{(\mathbb{W}E^* + g)^2} = c'^2,$$

and hence

(27) $$\mathbb{W}(\mathbb{D}\mathbb{D}^* - c'^2E^*E)\mathbb{W}^* + 2\mathbb{W}(\mathbb{D}F^* - c'^2E^*g) + |F|^2 - c'^2g^2 = 0.$$

Since (27) holds for an arbitrary c-line, we may replace $\mathbb{W}$ by $-\mathbb{W}$, and thus conclude that

$$\mathbb{W}(\mathbb{D}F^* - c'^2E^*g) = 0.$$

Therefore, since the direction of $\mathbb{W}$ is arbitrary,

(28) $$\mathbb{D}F^* = c'^2E^*g.$$

In view of the fact that (26) holds for $v(p, t) = 0$, we have

$$g \neq 0,$$

and we may then obtain, from (28),

(29) $$E^* = \frac{\mathbb{D}F^*}{c'^2g}.$$

Using (20) and (29), we obtain

$$|\mathfrak{A}| = g\left|\mathbb{D}\left(\mathcal{I} - \frac{F^*F}{c'^2g^2}\right)\right|;$$

and since $\mathfrak{A}$ is nonsingular, we have

(30) $$|F|^2 - c'^2g^2 \neq 0.$$

From (27), (28), and (30) it follows that

$$\mathfrak{D}\mathfrak{D}^* - c'^2 E^* E \neq 0.$$

Thus, from (27), we have

$$\mathfrak{W}(\mathfrak{D}\mathfrak{D}^* - c'^2 E^* E)\mathfrak{W}^* = \frac{c'^2 g^2 - |F|^2}{c^2}\,|\mathfrak{W}|^2.$$

Using again the fact that the direction of $\mathfrak{W}$ is arbitrary, we infer that

$$\mathfrak{D}\mathfrak{D}^* - c'^2 E^* E = \left(\frac{c'^2 g^2 - |F|^2}{c^2}\right)\mathfrak{I} = \mu\mathfrak{I}, \tag{31}$$

where $\mu = [c'^2 g^2 - |F|^2]/c^2$. From (28) and (31) we obtain

$$\begin{aligned} \mathfrak{A}\begin{pmatrix} \mathfrak{I} & 0 \\ 0 & -c'^2 \end{pmatrix}\mathfrak{A}^* &= \begin{pmatrix} \mathfrak{D}\mathfrak{D}^* - c'^2 E^* E & \mathfrak{D}F^* - c'^2 E^* g \\ (\mathfrak{D}F^* - c'^2 E^* g)^* & FF^* - c'^2 g^2 \end{pmatrix} \\ &= \begin{pmatrix} \mu\mathfrak{I} & 0 \\ 0 & -\mu c^2 \end{pmatrix} = \mu\begin{pmatrix} \mathfrak{I} & 0 \\ 0 & -c^2 \end{pmatrix}. \end{aligned} \tag{32}$$

We next want to show that μ is positive. Let $\langle Z_1, x_1\rangle$ and $\langle Z_2, x_2\rangle$ be two points in E_{n+1} such that

$$\frac{|Z_1 - Z_2|}{|x_1 - x_2|} < c,$$

and let

$$V = \langle Z_1, x_1\rangle - \langle Z_2, x_2\rangle$$

and

$$V' = V\mathfrak{A}.$$

From (32) we obtain

$$V\mathfrak{A}\begin{pmatrix} \mathfrak{I} & 0 \\ 0 & -c'^2 \end{pmatrix}\mathfrak{A}^* V^* = \mu V\begin{pmatrix} \mathfrak{I} & 0 \\ 0 & -c^2 \end{pmatrix} V^*.$$

Hence

$$\mu = \frac{V'\begin{pmatrix} \mathcal{I} & 0 \\ 0 & -c'^2 \end{pmatrix}(V')^*}{V\begin{pmatrix} \mathcal{I} & 0 \\ 0 & -c^2 \end{pmatrix}V^*}.$$

By the hypothesis on V,

$$V\begin{pmatrix} \mathcal{I} & 0 \\ 0 & -c^2 \end{pmatrix}V^* < 0,$$

and from the fact that c-inertial paths are carried into c'-inertial paths, we have

$$V'\begin{pmatrix} \mathcal{I} & 0 \\ 0 & -c'^2 \end{pmatrix}(V')^* < 0.$$

Thus μ is positive since it is the ratio of two negative numbers. We set

(33) $$\lambda = \sqrt{\mu}\ .$$

We then conclude from (32), (33), Definition 1, and Lemma 2 that

(34) $\mathfrak{A}$ is a generalized Lorentz matrix with respect to $\langle c, c', \lambda \rangle$.

We now turn to the function ϕ_3 which transforms the forces. In deducing the form of ϕ_3 it will be convenient to make use of the functions τ, q, and f^{rel} defined in § 3 (in the course of the present proof we obtain their transformation properties). It is also useful to introduce the function H defined by the following equation for every p in P and t in $\mathfrak{T}(p)$,

$$H(p, t) = [\phi_2(s(p, t), t)]_{n+1}.$$

We thus have that, for t' the element in $\mathfrak{T}'(p)$ corresponding to t in $\mathfrak{T}(p)$,

$$H_p(t) = t'.$$

We obtain, from (21),

(35) $$D(H_p)(t) = v_p(t)E^* + g.$$

For any S.R.P.M. $\Gamma = \langle P, \mathfrak{T}, m, s, f, c \rangle$, the following equation is a direct consequence of Axiom A7 and the appropriate definitions (for any p in P and t in $\mathfrak{T}(p)$):

$$(36) \qquad m(p)\frac{d^2 q_p}{d\tau_p^2}(t) = \sum_{i=1}^{\infty} f^{\mathrm{rel}}(p, t, i);$$

and also, under the hypothesis of our theorem,

$$(37) \qquad \phi_1(m(p))\frac{d^2 q_p'}{d\tau_p'^2}(H_p(t)) = \sum_{i=1}^{\infty} f^{\mathrm{rel}'}(p, H_p(t), i).$$

We now obtain the relationship between

$$\frac{d^2 q_p'}{d\tau_p'^2}(H_p(t)) \text{ and } \frac{d^2 q_p}{d\tau_p^2}(t).$$

Using (35), we obtain

$$(38) \qquad \frac{d(\tau' \circ H_p)}{d\tau_p}(t) = \left(\frac{D(\tau_p' \circ H_p)}{D\tau_p}\right)(t)$$

$$= \frac{[(D\tau_p')(H_p(t)][DH_p)(t)]}{(D\tau_p)(t)}$$

$$= \frac{\sqrt{1 - |v_p'(H_p(t))|^2/c'^2}\,(v_p(t)E^* + g)}{\sqrt{1 - |v_p(t))|^2/c^2}}.$$

It is easy to show that

$$(39) \qquad v_p'(H_p(t)) = \frac{v_p(t)\mathfrak{D} + F}{v_p(t)E^* + g};$$

hence, using (39) and squaring (38), we get

$$(40) \qquad \left[\left(\frac{D(\tau_p' \circ H_p)}{D\tau_p}\right)(t)\right]^2 = \frac{c^2}{c'^2}\left[\frac{c'^2(v_p(t)E^* + g)^2 - (v_p(t)\mathfrak{D} + F)^2}{c^2 - |v_p(t)|^2}\right].$$

Using (27) to give us the expansion of the right member of (40) and then using (32) to simplify the result, we obtain

$$(41)\quad \left[\frac{D(\tau_p'\circ H_p)}{D\tau_p}(t)\right]^2 = \frac{c^2}{c'^2}\left[\frac{-v_p(t)\lambda^2\,\mathfrak{A}(v_p(t))^* + \lambda^2 c^2}{c^2 - |v_p(t)|^2}\right] = \frac{c^2}{c'^2}\lambda^2\,;$$

hence

$$(42)\quad D(\tau_p'\circ H_p)(t) = \frac{\delta c\lambda}{c'}(D\tau_p)(t),$$

where $\delta^2 = 1$. We have, from (21) and Definition 3,

$$q_p'(H_p(t)) = q_p(t)\,\mathfrak{A} + B\,,$$

and thus

$$(43)\quad [D(q_p'\circ H_p)(t)] = (Dq_p)(t)\,\mathfrak{A}\,.$$

Since

$$\left[\frac{D(q_p'\circ H_p)}{D(\tau_p'\circ H_p)}\right](t) = \frac{(Dq_p')(H_p(t))(DH_p)(t)}{(D\tau_p')(H_p(t))(DH_p)(t)} = \left(\frac{Dq_p'}{D\tau_p'}\right)(H_p(t)),$$

it is easily shown that

$$(44)\quad \frac{d^2q_p'}{d\tau_p'^2}(H_p(t)) = \frac{D[D(q_p'\circ H_p)/D(\tau_p'\circ H_p)]}{D(\tau_p'\circ H_p)}(t).$$

From (42), (43), and (44) we infer that

$$(45)\quad \frac{d^2q_p'}{d\tau_p'^2}(H_p(t)) = \frac{D[(Dq_p)\,\mathfrak{A}]}{(\delta c\lambda/c')(D\tau_p)}(t) = \frac{c'^2}{c^2\lambda^2}\,\frac{d^2q_p}{d\tau_p^2}(t)\,\mathfrak{A}\,.$$

Now let X and Y be any two vectors in E_n with $X \neq 0$ and $|Y| < c$. Then we set:

$$P = \{1\},\ m(1) = 1,\ Z = \left(1 - \frac{Y^2}{c^2}\right)\left(X - \frac{(X\cdot Y)Y}{c^2}\right),$$

$$\mathfrak{T}(1) = \left(-\frac{c - |Y|}{|Z|},\ \frac{c - |Y|}{|Z|}\right),$$

and for all t in $\mathfrak{I}(1)$,

$$s_1(t) = tY + \frac{1}{2} t^2 Z,$$

$$f(1, t, 1) = \frac{Z}{1 - |v_1(t)|^2/c^2} + \frac{v_1(t)(Z \cdot v_1(t))}{c^2(1 - |v_1(t)|^2/c^2)^2},$$

$$f(1, t, i) = 0 \quad \text{for} \quad i > 1.$$

It is easy to verify that $\Gamma_{XY} = \langle P, \mathfrak{I}, m, s, f, c\rangle$ is a S.R.P.M., and consequently so is $\langle \Phi(\Gamma_{XY}), c'\rangle$. Thus there is a positive number γ such that

$$\phi_1(m(1)) = \gamma.$$

We note next that at $t = 0$:

$$s_1(0) = 0, \quad (Ds_1)(0) = Y, \quad (D^2 s_1)(0) = Z, \quad \text{and} \quad f(1, 0, 1) = X.$$

We thus have from (37), for $t = 0$,

$$\gamma \left[\frac{d^2 q_1'}{d\tau_1'^2} (H_1(0)) \right] = \phi_3(X, Y),$$

and thus, from (45),

$$\phi_3(X, Y) = \frac{\gamma c'^2}{c^2 \lambda^2} \left[\frac{d^2 q_1}{d\tau_1^2} (0) \mathfrak{A} \right]_{1, \cdots, n};$$

hence, from (36),

$$\phi_3(X, Y) = \frac{\gamma c'^2}{c^2 \lambda^2} \left[\left\langle X, \frac{X \cdot Y}{c^2} \right\rangle \mathfrak{A} \right]_{1, \cdots, n}. \tag{46}$$

In view of (1), (46) also holds for $X = 0$.

Now let x be any positive real number. Then we set:

$$\delta = \langle 1, \cdots, 1\rangle, \quad P = \{1\}, \quad m(1) = x, \quad \mathfrak{I}(1) = (-c, c);$$

for t in $\mathfrak{I}(1)$,

$$s_1(t) = \frac{1}{2} t^2 \delta,$$

$$f(1, t, 1) = \frac{x\delta}{1 - |v_1(t)|^2/c^2} + \frac{x v_1(t)(\delta \cdot v_1(t))}{c^2(1 - |v_1(t)|^2/c^2)^2},$$

$$f(1, t, i) = 0 \text{ for } i > 1$$

We easily verify that $\Gamma_x = \langle P, \mathfrak{T}, m, s, f, c \rangle$ is a S.R.P.M. such that for all t in $\mathfrak{T}(1)$,

$$\sum_{i=1}^{\infty} f^{\mathrm{rel}}(1, t, i) \neq 0.$$

Furthermore, we infer from (36), (37), and (45) that, for every t in $\mathfrak{T}(1)$,

(47) $$\sum_{i=1}^{\infty} f^{\mathrm{rel}\prime}(1, H_1(t), i) = \frac{\phi_1(x)}{x} \frac{c'^2}{c^2 \lambda^2} \sum_{i=1}^{\infty} f^{\mathrm{rel}}(1, t, i) \mathfrak{A};$$

hence, from (46),

(48) $$\phi_1(x) = \gamma x.$$

Our theorem now follows from (19), (33), (34), (46), and (48).

REMARK 3. We want to emphasize the physically reasonable nature of the hypothesis of the theorem just proved. We have assumed that systems of relativistic mechanics are carried by our transformations into systems of relativistic mechanics and that light lines are not carried into particle paths. No assumptions concerning the continuity of either ϕ_1, ϕ_2, or ϕ_3 have been made. Our assumption that ϕ_2 is one-to-one may be justified physically by the argument that any two space-time positions of a particle distinct with respect to one observer must be distinct with respect to every observer.

The standard presentations of the special theory of relativity vary a good deal in their "derivations" of the Lorentz transformations. Almost without exception, however, the assumptions underlying these derivations are not clearly and completely stated. For the physicist who wants to begin with a set of axioms for relativistic particle mechanics with respect to a fixed coordinate system, our Theorem 2 provides a rigorous approach to the derivation of the

Lorentz transformations. The transformations we obtain in Theorem 2 are, of course, more general than the Lorentz transformations, but it is obvious how the hypothesis of Theorem 2 may be strengthened so as to obtain just the ordinary Lorentz transformations.

Theorem 2 is also pertient to discussions of the relativity of size, (see, for example, [4]), since the determination of ϕ_1, ϕ_2, and ϕ_3 tells us exactly how the system of units of measurement may be changed in passing from one inertial frame of reference to another.

It is interesting to note that the set of transformations admissible (that is, satisfying the hypothesis of Theorem 2) in relativistic particle mechanics differs sharply from the set of those admissible (see the hypothesis of Theorem 3 of [8]) in classical particle mechanics: in the latter case, but not in the former, admissible transformations can change the unit of distance differently along different coordinates (with correspondingly different changes in the unit of force). Thus, although classical mechanics can in a certain sense be regarded as a limiting case of relativistic mechanics, the set of transformations admissible in classical mechanics is in no sense a limit of the set of transformations admissible in relativistic mechanics.

5. Algebraic structure of the set of admissible transformations. Let Φ be an eligible transformation which satisfies the hypothesis of Theorem 2 with respect to the positive real numbers c and c'. We then call the ordered triple $\langle \Phi, c, c' \rangle$ an *admissible triple*; and, corresponding to the informal usage at the end of the previous section, we call an eligible transformation an *admissible transformation* if it is the first element of some admissible triple. Since the set of admissible transformations is not a group under the obvious operation of composition, it is natural to ask what is its algebraic structure. We shall show that the structure of the set of admissible triples is that of a Brandt groupoid (formally defined below). Roughly speaking, the main difference between Brandt groupoids and groups is that a Brandt groupoid is not assumed to be closed under the binary operation corresponding to the group operation. Consequently, a Brandt groupoid may contain many identity elements, that is, many elements e such that $x * e = x = e * x$ whenever x, $x * e$, and $e * x$ are in the groupoid. If there is an e in the groupoid such that, for all x in the groupoid, $e * x = x = x * e$, then the groupoid is also a group. For this reason, we introduced the notion of an admissible triple: the admissible transformation which carries every S.R.P.M. into itself is an identity element whose composition with every admissible transformation is defined; consequently, the set of admissible transformations is neither a group nor a Brandt groupoid.

The notion of a Brandt groupoid was first defined in [**1**]; we use the formal definition given in [**5**].

DEFINITION 4. An algebraic system $\mathcal{G} = \langle G, *, J, {}^{-1} \rangle$ (where $*$ is an operation on a subset of $U \times U$ to U, J is a subset of U and ${}^{-1}$ is an operation on U to U) is called a *Brandt groupoid* if and only if the following conditions are satisfied:

(i) For x, y, z in G, if $x * y \in G$ and $y * z \in G$, then $(x * y) * z \in G$ and $(x * y) * z = x * (y * z)$.

(ii) For x, y, z in G, if $x * y \in G$ and $x * y = x * z$, then $y = z$.

(iii) For x, y, z in G, if $x * z \in G$ and $x * z = y * z$, then $x = y$.

(iv) For x in J, $x * x = x$.

(v) For x in G, $x^{-1} * x \in J$ and $x * x^{-1} \in J$.

(vi) For x, z in J, there exists a y in G such that $x * y \in G$ and $y * z \in G$.

Rather than deal directly with admissible triples, we find it somewhat simpler to use the following representation. From Theorem 2 we conclude that to each admissible triple there corresponds a unique ordered sextuple $\langle \mathfrak{A}, B, \gamma, \lambda, c, c' \rangle$, where B is an $(n+1)$-dimensional vector, γ, λ, c, and c' are positive real numbers, and $\mathfrak{A}$ is a generalized Lorentz matrix (of order $n+1$) with respect to $\langle c, c', \lambda \rangle$. Such an ordered sextuple $\langle \mathfrak{A}, B, \gamma, \lambda, c, c' \rangle$ we shall call a *carrier*. From Theorem 1, together with Theorem 2, it then follows that there is a one-to-one correspondence between the set of carriers and the set of admissible triples.

We say that the carrier $\langle \mathfrak{A}', B', \gamma', \lambda', c_1, c_2 \rangle$ is *left-conformable* to the carrier $\langle \mathfrak{A}, B, \gamma, \lambda, c_3, c_4 \rangle$ if and only if $c_1 = c_4$. By *the conformable subset* $\mathfrak{D}$ of $K \times K$ we mean the set of ordered pairs of elements of K such that the first element is left-conformable to the second.

We now define what we call *the carrier system*.

DEFINITION 5. By *the carrier system* we mean the ordered quadruple $\mathcal{K} = \langle K, *, J, {}^{-1} \rangle$, where:

(i) K is the set of all carriers;

(ii) $*$ is the operation on $\mathfrak{D}$ to K such that if the carrier $\langle \mathfrak{A}', B', \gamma', \lambda', c_1, c_2 \rangle$ is left-conformable to the carrier $\langle \mathfrak{A}, B, \gamma, \lambda, c_3, c_4 \rangle$ then

$$\langle \mathfrak{A}', B', \gamma', \lambda', c_1, c_2 \rangle * \langle \mathfrak{A}, B, \gamma, \lambda, c_3, c_4 \rangle = \langle \mathfrak{A}\mathfrak{A}', B\mathfrak{A}' + B', \gamma\gamma', \lambda\lambda', c_3, c_2 \rangle ;$$

(iii) J is the set of carriers of the form $\langle \mathcal{I}, 0, 1, 1, c, c \rangle$, where $\mathcal{I}$ is the identity matrix of order $n+1$; and

(iv) $^{-1}$ is the operation on K to K such that if $\langle \mathcal{A}, B, \gamma, \lambda, c, c' \rangle \in K$ then

$$\langle \mathcal{A}, B, \gamma, \lambda, c, c' \rangle^{-1} = \langle \mathcal{A}^{-1}, -B\mathcal{A}^{-1}, 1/\gamma, 1/\lambda, c', c \rangle.$$

We have then the following theorem, the proof of which we omit.

THEOREM 3. *The carrier system is a Brandt groupoid.*

We remark first that the operation $*$ of the carrier system corresponds to the composition of admissible triples; that is, if $\langle \Phi, c, c' \rangle$ corresponds to $\langle \mathcal{A}, B, \gamma, \lambda, c, c' \rangle$, and $\langle \Psi, c', c'' \rangle$ corresponds to $\langle \mathcal{A}', B', \gamma', \lambda', c', c'' \rangle$, then $\langle \mathcal{A}\mathcal{A}', B\mathcal{A}' + B', \gamma\gamma', \lambda\lambda', c, c'' \rangle$ corresponds to $\langle \theta, c, c'' \rangle$, where $\langle \theta, c, c'' \rangle$ is is the admissible triple such that, for any S.R.P.M. Γ_c,

$$\langle \Psi(\langle \Phi(\Gamma_c), c' \rangle), c'' \rangle = \langle \theta(\Gamma_c), c'' \rangle.$$

Similarly, the inverse operation $^{-1}$ of the carrier system corresponds to the natural inverse operation on admissible triples; that is, if $\langle \Phi, c, c' \rangle$ corresponds to $\langle \mathcal{A}, B, \gamma, \lambda, c, c' \rangle$, and $\langle \Psi, c', c \rangle$ corresponds to $\langle \mathcal{A}^{-1}, -B\mathcal{A}^{-1}, 1/\gamma, 1/\lambda, c', c \rangle$, then, for any S.R.P.M. Γ_c,

$$\langle \Psi(\langle \Phi(\Gamma_c), c' \rangle), c \rangle = \Gamma_c .$$

It thus follows as a corollary to Theorem 3 that the set of admissible triples is a Brandt groupoid under the natural operations of composition and formation of inverses.

It is natural to ask how the hypothesis of Theorem 2 may be strengthened so that the set of eligible transformations satisfying it form a group. We state without proof some results concerning this question.

THEOREM 4. *Let $\Phi = \langle \phi_1, \phi_2, \phi_3 \rangle$ be an eligible transformation which carries every system of relativistic particle mechanics into a system of relativistic particle mechanics. Then there exist positive real numbers δ, γ, λ, and ρ, an $(n+1)$-dimensional vector B, an orthogonal matrix $\mathcal{E}$ of order n, and a matrix $\mathcal{A}$ of order $n+1$, such that*

$$\delta^2 = 1, \qquad \mathcal{A} = \lambda \begin{pmatrix} \mathcal{I} & 0 \\ 0 & \rho \end{pmatrix} \begin{pmatrix} \mathcal{E} & 0 \\ 0 & \delta \end{pmatrix};$$

and for any vectors Z_1 *and* Z_2 *in* E_n, *any* x *in* R, *and* y *in* R^+,

$$\phi_1(y) = \gamma y,$$

$$\phi_2(Z_1, x) = \langle Z_1, x\rangle \, \mathcal{A} + B,$$

$$\phi_3(Z_1, Z_2) = \frac{\gamma \rho^2 Z_1 \mathcal{E}}{\lambda}.$$

The interpretation of δ, γ, λ, B, and $\mathcal{E}$ is the same as that stated in Remark 1. The number ρ is the ratio c/c' of the absolute values of the old and new velocities of light. The matrix $\mathcal{A}$ is a generalized Lorentz matrix with $U = 0$, which intuitively means that the old and new spatial frames of reference are at rest with respect to each other. The fact that the hypothesis of Theorem 4 thus excludes the possibility of transforming from one inertial frame of reference to another moving with respect to it is sufficient reason to regard this hypothesis as unnecessarily strong from the point of view of our intended physical interpretation. On the other hand, it is, of course, clear that the set of transformations satisfying this hypothesis constitute a group under the obvious operations.

References

1. H. Brandt, *Über eine Verallgemeinerung des Gruppenbegriffes*, Math. Ann. 96 (1926-1927), 360-366.

2. P. Destouches-Février, *La structure des théories physiques*, Paris, 1951.

3. H. Hermes, *Eine Axiomatisierung des allgemeinen Mechanik*, Heft 3, Forschungen zur Logik und zur Grundlegung der Exakten Wissenschaften, Neue Folge, Leipzig, 1938.

4. B. Hoffmann, *The relativity of size*, Physical Rev. 89 (1953), 49-52.

5. B. Jónsson and A. Tarski, *Boolean algebras with operators, Part* II, Amer. J. Math. 74 (1952), 127-162.

6. L. A. MacColl, *Geometric aspects of relativistic dynamics*, Trans. Amer. Math. Soc. 46 (1939), 328-347.

7. J. C. C. McKinsey, A. C. Sugar, and P. Suppes, *Axiomatic foundations of classical particle mechanics*, J. Rational Mech. Anal. 2 (1953), 253-272.

8. J. C. C. McKinsey and P. Suppes, *Transformations of systems of classical particle mechanics*, J. Rational Mech. Anal. 2 (1953), 273-289.

9. ————, *Philosophy and the axiomatic foundations of physics*, Proceedings of the XIth International Congress of Philosophy, vol. VI (1953), 49-54.

10. K. Menger, *Are variables necessary in calculus*, Amer. Math. Monthly 56 (1949), 609-620.

11. H. Reichenbach, *Axiomatik der relativistischen Raum-Zeit-Lehre,* Braunschweig, 1924.

12. A. A. Robb, *Geometry of time and space*, Cambridge, 1936.

13. K. Schnell, *Eine Topologie der Zeit in logistischer Darstellung*, Dissertation, Münster, 1937.

14. A. G. Walker, *Foundations of relativity* I, II, Proc. Royal Society of Edinburgh, Section A, 62 (1948), 319-335.

STANFORD UNIVERSITY

ON THE DIMENSION THEORY OF RINGS (II)

A. SEIDENBERG

1. Introduction. As in [3], we shall say that an integral domain O is n-dimensional if in O there is a proper chain

$$(0) \subset P_1 \subset \cdots \subset P_n \subset (1)$$

of prime ideals, but no such chain

$$(0) \subset P_1' \subset \cdots \subset P_{n+1}' \subset (1).$$

In Theorem 2 of [3] it was shown that if O is n-dimensional, then $O[x]$ is at least $(n+1)$-dimensional and at most $(2n+1)$-dimensional: here, as throughout, x is an indeterminate. After preparatory constructions in Theorems 1 and 2 below, this theorem is completed in Theorem 3 by showing that for any integers m and n with $n+1 \leq m \leq 2n+1$, there exist n-dimensional rings O such that $O[x]$ is m-dimensional. The other theorems mainly concern 1-dimensional rings. Such rings O can be divided into those for which $O[x]$ is 2-dimensional and those for which this condition fails, the so-called F-rings. The paper [3] was concerned with the existence of F-rings and showed [3, Theorem 8] that the 1-dimensional ring O is not an F-ring if and only if every quotient ring of the integral closure of O is a valuation ring. Below, in Theorem 5, we show more generally that if O is 1-dimensional but not an F-ring, then $O[x_1, \cdots, x_n]$ is $(n+1)$-dimensional, where the x_i are indeterminates: this theorem depends on the essentially more general Theorem 4, which says that if O is an m-dimensional multiplication-ring, then $O[x_1, \cdots, x_n]$ is $(m+n)$-dimensional. In the case that the x_i are not indeterminates, one can still say (Theorem 10) that

$$\dim O[x_1, \cdots, x_n] = 1 + \text{degree of transcendency of } O[x_1, \cdots, x_n]/O,$$

provided that the intersection of the prime ideals ($\neq (0)$) in O is $= (0)$, where O is a 1-dimensional ring such that $O[x]$ is 2-dimensional. For F-rings O, Theorem 6 shows that

$$n+2 \leq \dim O[x_1, \cdots, x_n] \leq 2n+1,$$

Received April 10, 1953.
Pacific J. Math. 4 (1954), 603-614

where the x_i are indeterminates, while Theorem 7 constructs for any N and n with

$$n + 2 \leq N \leq 2n + 1$$

an F-ring O such that $O[x_1, \cdots, x_n]$ is N-dimensional. Similar results for rings of dimension greater than 1 would be interesting if one could get them.

2. Simple extensions. Let us call the integral domain O *of type* (n, m) if

$$\dim O = n \text{ and } \dim O[x] = m.$$

THEOREM 1. *Let O be integrally closed and of type (n, m), let K be its quotient field, and let K' be a proper extension of K in which K is algebraically closed. Let Σ be any field having a discrete rank 1 valuation with K' as residue field. Let O^* be the set of elements whose residues are finite and in O. Then O^* is integrally closed and of type $(n + 1, m + 2)$.*

Proof. Let $\alpha \in \Sigma$, with α integral over O^*,

$$\alpha^s + a_1 \alpha^{s-1} + \cdots + a_s = 0 \qquad (a_i \in O^*),$$

an equation of integral dependence. Dividing this equation by α^s and supposing $1/\alpha$ to have residue 0, we get the contradiction $1 = 0$. So α has finite residue, and

$$\bar{\alpha}^s + \bar{a}_1 \bar{\alpha}^{s-1} + \cdots + \bar{a}_s = 0,$$

where the bars indicate residues. Since K is algebraically closed in K', we have $\bar{\alpha} \in K$; and $\bar{\alpha} \in O$, since O is integrally closed. Hence O^* is integrally closed.

Let P be the set of $\alpha \in O^*$ having residue 0. Then P is a prime ideal. From the definitions one obtains

$$O^*/P \simeq 0,$$

whence O^* is at least $(n+1)$-dimensional. If P' is a prime ideal in O^*, $P' \neq (0)$, then $P' \supseteq P$. In fact, let $g \in P'$; since g is O^*, we have $v(g) = s \geq 0$, where v is the given valuation (and the group of integers is the valuation group). Then the $(s + 1)$th power of any element in P is divisible by g, whence $P \subseteq P'$. From this it follows that O^* is at most $(n + 1)$-dimensional.

The quotient ring O_P^* is integrally closed and has only one prime ideal

($\neq (0)$). Moreover it is not a valuation ring. In fact, let $\alpha \in \Sigma$ be an element having residue in K' but not in K. Since α can clearly be written as a quotient of two elements of positive value, we have that α is in the quotient field of O_P^*; but neither α nor $1/\alpha$ has residue in K, so neither α nor $1/\alpha$ is in O_P^*. Thus O_P^* is not a valuation ring, and hence is an F-ring, by [3, Theorem 8]. It follows at once that $O^*[x] \cdot P$ is not minimal in $O^*[x]$. Now

$$O^*[x]/O^*[x] \cdot P \simeq O^*/P[x] \simeq O[x],$$

so $O^*[x]$ is at least $(m+2)$-dimensional.

Finally, let $(0) \subset P_1 \subset P_2 \subset \cdots \subset P_s \subset (1)$ be a chain of prime ideals in $O^*[x]$. Let P_1 be minimal; then $P_1 \cap O^* = (0)$, as otherwise

$$P_1 \cap O \supseteq P \quad \text{and} \quad P_1 \supseteq O^*[x] \cdot P.$$

Similarly one concludes that if no chain of prime ideals $P_1' \subset P_1''$ can be inserted between (0) and P_2, then

$$P_2 \cap O^* = P \quad \text{and} \quad P_2 = O^*[x] \cdot P$$

(by [3, Theorem 1], P_2 cannot contract in O to (0)). From this it follows at once that $O^*[x]$ is at most $(m+2)$-dimensional, and the proof is complete.

Remark. The above construction stems from an example of Krull showing that an integrally closed integral domain with only one proper prime ideal need not be a valuation ring; see [2, p. 670f].

Theorem 2. *Let O, K, K', Σ, O^* be as in Theorem 1 except that we assume $K = K'$. Then O^* is integrally closed and of type $(n+1, m+1)$.*

Proof. The proof follows exactly the lines of the proof of Theorem 1, except that here O_P^* *is* a valuation ring, as one easily sees.

Theorem 3. *For every n and m such that $n+1 \leq m \leq 2n+1$ there exist integrally closed rings of type (n, m).*

Proof. Any field is of type $(0, 1)$. Theorem 1 gives us an integrally closed ring of type $(1, 3)$, and Theorem 2 gives us one of type $(1, 2)$—the required valuations obviously exist. Suppose now by induction that for some n and each m, $n+1 \leq m \leq 2n+1$, we have an integrally closed ring of type (n, m). If $n+3 \leq m \leq 2n+3$, then $n+1 \leq m-2 \leq 2n+1$, and from an integrally closed ring of type $(n, m-2)$ we get, by Theorem 1, an integrally closed ring of type

$(n+1, m)$. If $m = n+2$, we apply Theorem 2 similarly to get an integrally closed ring of type $(n+1, m)$.

As for simple algebraic extensions $O[\alpha]$ of an n-dimensional ring O, it is clear that dim $O[\alpha] \leq 2n$. On the other hand, let O be an integrally closed ring of type (n, m) and let O^* be a ring constructed as in Theorem 1; also let Σ and P be as in Theorem 1. Let

$$\alpha \in \Sigma, \; \alpha \notin O^*_P, \; 1/\alpha \notin O^*_P.$$

Then

$$O^*[\alpha]/O^*[\alpha] \cdot P \simeq O^*/P[x] \simeq O[x],$$

by [3, Theorem 7], so $O^*[\alpha]$ is at least $(m+1)$-dimensional; it is also at most $(m+1)$-dimensional, since $O^*[x]$ is $(m+2)$-dimensional. Hence

$$(n+1)+1 \leq \dim O^*[\alpha] \leq 2(n+1).$$

It is thus clear that for any $n' > 0$ and m' with $n'+1 \leq m' \leq 2n'$, there exists an n'-dimensional ring O^* such that for some α in the quotient-field of O^* we have dim $O^*[\alpha] = m'$. — Also

$$\dim O[\alpha] < \dim O$$

is possible. In fact, let O be a valuation ring of rank n, $(0) \subset p_1 \subset \cdots \subset p_n \subset (1)$, the chain of prime ideals in O. Let $c \in p_{i+1}$, $c \notin p_i$; then dim $O[1/c] = 1$. *In short,* dim $O[\alpha]$ *covers precisely the range from* 0 *to* $2n$ *as* O *varies over the* n*-dimensional rings* O.

3. Multiple transcendental extensions. We recall that a *multiplication-ring* may be defined as an integral domain O such that O_p is a valuation ring for each prime ideal p in O (see [2, p. 554]).

THEOREM 4. *If O is an m-dimensional multiplication-ring, then $O[x_1, \cdots, x_n]$ is $(m+n)$-dimensional, where the x_i are indeterminates.*

Proof. To facilitate the proof, we define the *dimension of a prime ideal* P in an extension $O' = O[\alpha_1, \cdots, \alpha_n]$ of a finite-dimensional ring O (relative to O) as follows:

$$\dim P = \text{d.t.}\,(O'/P)/(O/P) + \dim O/p,$$

where $p = P \cap O$ (and "d.t." abbreviates "degree of transcendence"). The following points (a), (b) do not assume O to be a multiplication-ring.

(a) Let $\overline{O}'$, $\overline{O}$, $\overline{P}$, $\overline{\mathfrak{p}}$ be the images of O', O, P, $\mathfrak{p}$, respectively, under a homomorphism with kernel contained in P. Then $\dim \overline{P} = \dim P$.

In fact, $\overline{O}'/\overline{P} = O'/P$ and $\overline{O}/\overline{\mathfrak{p}} = O/\mathfrak{p}$; also $\overline{P} \cap \overline{O} = \overline{\mathfrak{p}}$.

(b) Let M be a nonempty multiplicatively closed system in O not meeting $\mathfrak{p}$,

$$O_M = \{\alpha \mid \alpha = a/b,\ a \in O,\ b \in M\},$$

$$O'_M = \{\alpha \mid \alpha = a/b,\ a \in O',\ b \in M\}.$$

Then

$$\dim P - \dim O/\mathfrak{p} = \dim O'_M \cdot P - \dim O_M/O_M \cdot \mathfrak{p}.$$

In fact, the rings O'/P and $O'_M/O'_M \cdot P$ have the same quotient field, as do the rings $O/\mathfrak{p}$ and $O_M/O_M \cdot \mathfrak{p}$. Note also that $O'_M \cdot P \cap O_M = O_M \cdot \mathfrak{p}$, whence the required equality follows.

Let P_1, P_2 be two prime ideals in O', $P_1 \subset P_2$, $\mathfrak{p}_i = P_i \cap O$, $i = 1, 2$. We want to compare $\dim P_1$ with $\dim P_2$. If $\mathfrak{p}_1 = \mathfrak{p}_2$, then, passing to a residue class ring, we may assume $\mathfrak{p}_1 = \mathfrak{p}_2 = (0)$. Taking $M = O - (0)$, we pass to the quotient-ring O'_M, which is a finite integral domain. Thus $\dim P_1 > \dim P_2$ if $\mathfrak{p}_1 = \mathfrak{p}_2$. This conclusion holds also if $\mathfrak{p}_1 \subset \mathfrak{p}_2$ provided O is a multiplication-ring.

(c) If P_1 and P_2 are prime ideals in $O[x_1, \cdots, x_n]$ and $P_1 \subset P_2$, then

$$\dim P_1 > \dim P_2;$$

also

$$\dim P_1 - \dim P_2 \geq \dim O/\mathfrak{p}_1 - \dim O/\mathfrak{p}_2,$$

provided that O is a multiplication-ring.

In fact, we may suppose $\mathfrak{p}_1 \subset \mathfrak{p}_2$, and have only to prove the second point. Also, by (b), we may pass to any quotient-ring O_M, where M does not meet $\mathfrak{p}_2$. Taking $M = O - \mathfrak{p}_2$, we may assume that O is a valuation-ring and that $\mathfrak{p}_2$ is its ideal of non-units. Let $z_1, \cdots, z_r$ be elements of O' which are algebraically dependent mod P_1 over O. Then they are also dependent mod P_2. In fact, let

$$f(z_1, \cdots, z_r) \equiv 0 (P_1),$$

where the coefficients of the polynomial f are in O but not all in p_1. Dividing by a coefficient of least value, we may suppose f to have a coefficient equal to unity. But then we have a relation mod P_2. This proves that

$$\text{d.t. } (O'/P_2)/(O/p_2) \leq \text{d.t. } (O'/P_1)/(O/p_1),$$

that is, (c) is proved.

The theorem now follows from (c) since $\dim (0) = m + n$.

COROLLARY. *If O is an m-dimensional multiplication-ring then*

$$\dim O[\alpha_1, \cdots, \alpha_n] \leq m + r,$$

where

$$r = \text{d.t. } O[\alpha_1, \cdots, \alpha_n]/O.$$

Proof. The foregoing proof shows that

$$\dim O[x_1, \cdots, x_n] \leq \dim (0) = m + \text{d.t. } O[x_1, \cdots, x_n]/O,$$

and in doing so makes no use of the fact that the x_i are indeterminates; this fact is used only to get that

$$\dim O[x_1, \cdots, x_n] \geq m + n.$$

THEOREM 5. *If O is a* 1*-dimensional ring such that $O[x]$ is* 2*-dimensional, then*

$$\dim O[\alpha_1, \cdots, \alpha_n] \leq 1 + \text{d.t. } O[\alpha_1, \cdots, \alpha_n]/O;$$

if the α_i are indeterminates, then

$$\dim O[\alpha_1, \cdots, \alpha_n] = 1 + n.$$

Proof. We may suppose O to be integrally closed. In that event, O is a multiplication-ring, by [3, Theorem 8]. The present theorem now follows immediately from the preceding corollary.

THEOREM 6. *If O is* 1*-dimensional, then $O[x_1, \cdots, x_n]$ is at most* $(2n+1)$*-dimensional, where the x_i are indeterminates.*

Proof. Let $(0) \subset p_1 \subset p_2 \subset \cdots \subset p_s \subset (1)$ be a chain of prime ideals in

$O[x_1, \cdots, x_n]$. Let K = quotient field of O. If $p_s \cap O = (0)$, then the above chain extends to a chain of s prime ideals in $K[x_1, \cdots, x_n]$, so $s \leq n$. Suppose, then, that

$$p_i \cap O = (0), \quad p_{i+1} \cap O = p \neq (0),$$

whence also $p_{i+k} \cap O = p$, since O is 1-dimensional. Passing to $K[x_1, \cdots, x_n]$, we see that $i \leq n$; and passing to

$$O[x_1, \cdots, x_n]/p_{i+1} = O/p\,[\bar{x}_1, \cdots, \bar{x}_n],$$

we have $s - (i + 1) \leq n$, since O/p is a field. Hence $s \leq 2n + 1$.

THEOREM 7. *If O is an F-ring, then $O[x_1, \cdots, x_n]$ is at least $(n+2)$-dimensional and at most $(2n+1)$-dimensional. For any N, $n+2 \leq N \leq 2n+1$, there is an F-ring O such that $O[x_1, \cdots, x_n]$ is N-dimensional, where the x_i are indeterminates.*

Proof. Let K be a field, x, $y_1, \cdots, y_m$ indeterminates. Let

$$K' = K(y_1, \cdots, y_m), \quad \Sigma = K'(x),$$

and let v be the discrete rank 1 valuation of Σ obtained by placing

$$v(a_i x^i + a_{i+1} x^{i+1} + \cdots + a_s x^s) = i,$$

where $a_j \in K'$, $a_i \neq 0$. Let O^* be the set of elements whose residues are finite and in K. The ring O^* consists of the elements in $K(x, y_1, \cdots, y_m)$ which can be written in the form

$$\alpha(x, y_1, \cdots, y_m)/\beta(x, y_1, \cdots, y_m),$$

where

$$\alpha, \beta \in K[x, y_1, \cdots, y_m], \ \beta(0, y_1, \cdots, y_m) \neq 0,$$

$$\alpha(0, y_1, \cdots, y_m)/\beta(0, y_1, \cdots, y_m) \in K.$$

By Theorem 1, O^* is an F-ring; and O^* contains only one proper prime ideal, namely the ideal P consisting of the elements α/β with

$$\alpha(0, x_1, \cdots, x_m)/\beta(0, x_1, \cdots, x_n) = 0.$$

We shall prove that for $m \leq n$, $O^*[x_1, \cdots, x_n]$ is $(m+n+1)$-dimensional. In $O^*[x_1, \cdots, x_n]$ let P_m be the ideal of polynomials which vanish for $x_i = y_i$ $(i = 1, \cdots, m)$. We claim this ideal is in

$$O^*[x_1, \cdots, x_n] \cdot P = P'.$$

In fact, let

$$\sum a_{i_1 \cdots i_n} x_1^{i_1} \cdots x_n^{i_n} \in O^*[x_1, \cdots, x_n]$$

be in P_m, and write

$$\sum a_{i_1 \cdots i_n} x_1^{i_1} \cdots x_n^{i_n} = \sum c_{i_1 \cdots i_n} x_1^{i_1} \cdots x_n^{i_n} + \sum d_{i_1 \cdots i_n} x_1^{i_1} \cdots x_n^{i_n},$$

where $c_{i_1 \cdots i_n} \in K$, $d_{i_1 \cdots i_n} \in P$. This polynomial vanishes for $x_i = y_i$, $i = 1, \cdots, m$; hence also for $x_i = y_i$, $i = 1, \cdots, m$, $x = 0$. Hence

$$\sum c_{i_1 \cdots i_n} x_1^{i_1} \cdots x_n^{i_n}$$

vanishes for $x_i = y_i$, $i = 1, \cdots, m$, whence

$$\sum c_{i_1 \cdots i_n} x_1^{i_1} \cdots x_n^{i_n} = 0,$$

and

$$\sum a_{i_1 \cdots i_n} x_1^{i_1} \cdots x_n^{i_n} = \sum d_{i_1 \cdots i_n} x_1^{i_1} \cdots x_n^{i_n} \in O^*[x_1, \cdots, x_n] \cdot P = P'.$$

Let P_j be the ideal of elements in $O^*[x_1, \cdots, x_n]$ which vanish for $x_i = y_i$, $i = 1, \cdots, j$. Then P_j is prime and $(0) \subset P_1 \subset \cdots \subset P_m \subset P'$. Since any chain of n prime ideals in $O^*/P[x_1, \cdots, x_n]$ gives rise to such a chain in $O^*[x_1, \cdots, x_n]$ containing P', we see that $O^*[x_1, \cdots, x_n]$ is at least $(m+n+1)$-dimensional. On the other hand, $O^*[x_1, \cdots, x_n]$ is of degree of transcendency $m+n+1$ over K, and so $O^*[x_1, \cdots, x_n]$ is at most $(m+n+1)$-dimensional. This last point follows from the following lemma, the proof of which is exactly as in the well-known case that O is a valuation ring.

LEMMA. *Let O be an arbitrary integral domain containing a field K, and let O be of degree of transcendency r over K. Then O is at most r-dimensional.*

Proof. This follows at once if we can show that the degree of transcendency of O/P over K is less than r for any proper prime ideal P in O. If $\theta_1, \theta_2, \cdots, \theta_s \in O$ map into (given) algebraically independent elements in O/P, and $\theta \in P$, $\theta \neq 0$, then $\theta, \theta_1, \cdots, \theta_s$ are algebraically independent over K. Hence

$$\text{d.t. } O/K > \text{d.t. } (O/P)/K.$$

4. Arbitrary finite extensions. Let O be an arbitrary integral domain which is not a field. It is certainly possible, for appropriate O, that some simple ring extension $O[\alpha]$ of O will be a field. In fact, let O be such that the intersection of all its prime ideals ($\neq (0)$) is not the ideal (0); for example, any integral domain with a finite, positive number of prime ideals ($\neq (0)$) will do. If $c (\neq 0)$ is an element in all the prime ideals, then $O[1/c]$ is a field; for if P is a prime ideal in $O[1/c]$, $P \neq (0)$, then

$$P \cap O = p \neq (0)$$

and

$$1 = (1/c) \cdot c \in O[1/c] \cdot p \subseteq P.$$

We also have the converse.

THEOREM 8. *Given an integral domain O, there exists a field F which is a simple ring extension of O if and only if the intersection of all the prime ideals $(\neq (0))$ in O is $\neq (0)$.*

Proof. Let $F = O[\alpha]$. Here α must be algebraic over O, say

$$c_0 \alpha^m + c_1 \alpha^{m-1} + \cdots + c_m = 0, \quad c_i \in O, \quad c_0 \neq 0.$$

Then $c_0\alpha$ is integral over O, as is the ring $O_1 = O[c_0\alpha]$. Let $F_1 = O_1[\alpha]$; then F_1 is a field [**1**, p. 253]. Over every prime ideal in O there lies a prime ideal in O_1; since O_1 is algebraic over O, if the intersection of the prime ideals $(\neq (0))$ in O_1 is $\neq (0)$, then the like is true in O. Hence we may assume that $O = O_1$, that is, that α is in the quotient field of O. By a similar reasoning we may suppose O is integrally closed. From the fact that $1/\alpha \in O[\alpha]$, one finds that $1/\alpha$ is integral over O, hence in O. Thus $\alpha = 1/b$, $b \in O$. The element b must be in every prime ideal $p (\neq (0))$; in fact, if $b \notin p$, then $O[1/b] \subseteq O_p$, whence $O[1/b] = O[\alpha]$ is not a field. This completes the proof. —This theorem has been previously proved in [**4**, p. 76].

A study of algebraic extensions of O must therefore separate the cases that the intersection of the prime ideals ($\neq (0)$) is $= (0)$ or is $\neq (0)$.

THEOREM 9. *If O is a 1-dimensional ring such that $O[x]$ is 2-dimensional, and the intersection of the prime ideals* ($\neq (0)$) *in O is* $= (0)$, *then the like is true of any simple algebraic ring extension $O[\alpha]$ of O (where it is assumed, of course, that $O[\alpha]$ is an integral domain).*

Proof. By Theorem 5, we know that $O[\alpha]$ is 0- or 1-dimensional, and the previous theorem excludes the first alternative. Also $O[\alpha, x]$ is 2-dimensional, for otherwise $O[y, x]$, y an indeterminate, would be of dimension more than 3, contradicting Theorem 5. Thus it remains to prove that the intersection of the prime ideals ($\neq (0)$) in $O[\alpha]$ is $= (0)$. Let

$$c_0 \alpha^n + c_1 \alpha^{n-1} + \cdots + c_n = 0, \quad c_i \in O, \quad c_0 \neq 0,$$

and let $S = \{p\}$ be the set of prime ideals ($\neq (0)$) in O which do not contain c_0; S is not empty. Then $\cap p = (0)$, for if $d \in \cap p$, $d \neq 0$, then $c_0 d$ is in every prime ideal ($\neq (0)$) of O. Over every prime ideal $p \in S$ there lies a prime ideal P in $O[\alpha]$. If $T = \{P\}$ is the set of prime ideals in $O[\alpha]$ contracting to prime ideals in S, then one concludes immediately that $\cap P = (0)$. A fortiori the intersection of all prime ideals ($\neq (0)$) in $O[\alpha]$ is $= (0)$. This completes the proof.

If O is an integral domain in which the intersection ($\cap p$) of the prime ideals ($\neq (0)$) is $\neq (0)$, then for every r it is possible to define a finite extension $O[\alpha_1, \cdots, \alpha_n]$ of O such that

$$\dim O[\alpha_1, \cdots, \alpha_n] = r$$

and

$$\text{d.t. } O[\alpha_1, \cdots, \alpha_n]/O = r;$$

namely, we adjoin to O an element $1/c$, $c \in \cap p$, so that $O[1/c]$ is the quotient field of O, and thereupon adjoin r indeterminates. The situation is different for a 1-dimensional ring which is not an F-ring and in which the intersection of the prime ideals ($\neq (0)$) is $= (0)$.

THEOREM 10. *Let O be a 1-dimensional ring such that $O[x]$ is 2-dimensional, and let the intersection of the prime ideals* ($\neq (0)$) *in O be* $= (0)$. *Then for any integral domain $O[\alpha_1, \cdots, \alpha_n]$,*

$$\dim O[\alpha_1, \cdots, \alpha_n] = 1 + \text{d.t. } O[\alpha_1, \cdots, \alpha_n]/O.$$

Proof. Let

$$K = \text{quotient field of } O, \quad r = \text{d.t. } O[\alpha_1, \cdots, \alpha_n]/O.$$

Then $K[\alpha_1, \cdots, \alpha_n]$ is r-dimensional and a chain $(0) \subset P_1 \subset \cdots \subset P_r \subset (1)$ of prime ideals in $K[\alpha_1, \cdots, \alpha_n]$ contracts to a chain

$$(0) \subset p_1 \subset \cdots \subset p_r \subset (1), \text{ and } p_i \cap O = (0), \; i = 1, \cdots, r.$$

Moreover, p_r is not maximal, for if it were, then

$$O[\alpha_1, \cdots, \alpha_n]/p_r = O[\bar{\alpha}_1, \cdots, \bar{\alpha}_n]$$

would be a field; hence also $K[\bar{\alpha}_1, \cdots, \bar{\alpha}_n]$ would be a field, whence the $\bar{\alpha}_i$ would be algebraic over K, therefore also over O. This contradicts the previous theorem. Hence

$$\dim O[\alpha_1, \cdots, \alpha_n] \geq 1 + \text{d.t. } O[\alpha_1, \cdots, \alpha_n]/O,$$

and we have already seen the reverse inequality.

Since the theory of 1-dimensional rings must separate the cases that the intersections of prime ideals $(\neq (0))$ is $= (0)$ or $\neq (0)$, it may be of interest to have an example of a 1-dimensional ring, not an F-ring, *with infinitely many prime ideals* $(\neq (0))$ having intersection $\neq (0)$. We construct such a ring O as follows. Let K be a field containing all roots of unity, x an indeterminate, L the algebraic closure of $K(x)$, S the integral closure in L of $K[x]$, and O, the quotient-ring of S with respect to the multiplicatively closed system of polynomials in $K[x]$ which are not divisible by x. Infinitely many prime ideals in S lie over (x) in $K[x]$; to see this, let n be any integer not divisible by the characteristic of K, $a_1, \cdots, a_n$ the nth roots of unity, $y = \sqrt[n]{1+x}$. In $K[x, y]$ there lie n prime ideals over (x), namely $(x, y - a_i)$, since $(0, a_i)$ is a point of $y^n = 1 + x$. Going up to S, we see that there exist at least n prime ideals over (x). Every prime ideal in O which differs from (0) contains x, and there are infinitely many such ideals. We now verify immediately that O is a ring of the required type.

REFERENCES

1. I.S. Cohen and A. Seidenberg, *Prime ideals and integral dependence*, Bull. Amer. Math. Soc. 52 (1946), 252-261.

2. W. Krull, *Beiträge zur Arithmetik kommutativer Integritätsbereiche*, I. *Multiplikationsringe, ausgezeichnete Idealsysteme, Kroneckersche Funktionalringe*, Math. Z. **41** (1936), 545-577. II. *v-Ideale und vollständige ganz abgeschossene Integritätsbereiche*, Math. Z. **41** (1936), 665-679.

3. A. Seidenberg, *A note on the dimension theory of rings*, Pacific J. Math. 3 (1953), 505-512.

4. E. Artin and J. T. Tate, *A note on finite ring extensions*, J. Math. Soc. Japan, 3 (1951), 74-77.

UNIVERSITY OF CALIFORNIA, BERKELEY

DIFFERENCE ALGEBRAS OF LINEAR TRANSFORMATIONS ON A BANACH SPACE

Bertram Yood

1. Introduction. Let $\mathfrak{E}(\mathfrak{X})$ be the Banach algebra of all bounded linear transformations defined on an infinite-dimensional Banach space $\mathfrak{X}$ and with range in $\mathfrak{X}$. Let $\mathfrak{K}(\mathfrak{X})$ be the set of completely continuous transformations contained in $\mathfrak{E}(\mathfrak{X})$. It is well known that $\mathfrak{K}(\mathfrak{X})$ is a closed two-sided ideal in $\mathfrak{E}(\mathfrak{X})$. Thus, under the usual definitions, the difference algebra $\mathfrak{E}(\mathfrak{X}) - \mathfrak{K}(\mathfrak{X})$ is again a Banach algebra. Let π be the canonical homomorphism of $\mathfrak{E}(\mathfrak{X})$ onto $\mathfrak{E}(\mathfrak{X}) - \mathfrak{K}(\mathfrak{X})$.

The algebraic nature of $\mathfrak{E}(\mathfrak{X}) - \mathfrak{K}(\mathfrak{X})$ differs from that of $\mathfrak{E}(\mathfrak{X})$. In particular $\mathfrak{E}(\mathfrak{X})$ is semi-simple while $\mathfrak{E}(\mathfrak{X}) - \mathfrak{K}(\mathfrak{X})$ need not be semi-simple. An example of this is provided by taking for $\mathfrak{X}$ the Banach space $L(S)$ of Lebesgue-integrable numerical functions defined on, say, the unit interval S. If T and U are in $\mathfrak{E}(\mathfrak{X})$ and are weakly completely continuous then TU is completely continuous as shown by Dunford and Pettis [**5**, p. 370]. From this it follows readily that the image of the set of weakly completely continuous transformations in $\mathfrak{E}(\mathfrak{X})$ under π is contained in the radical $\mathfrak{P}_1$ of $\mathfrak{E}(\mathfrak{X}) - \mathfrak{K}(\mathfrak{X})$. Hence $\mathfrak{E}(\mathfrak{X}) - \mathfrak{K}(\mathfrak{X})$ is not semi-simple for this $\mathfrak{X}$. On the other hand if $\mathfrak{X}$ is (separable) Hilbert space, then $\mathfrak{E}(\mathfrak{X}) - \mathfrak{K}(\mathfrak{X})$ is semi-simple.

In this paper we begin an investigation of the algebra $\mathfrak{E}(\mathfrak{X}) - \mathfrak{K}(\mathfrak{X})$. In particular its radical and its set of regular elements are examined. This turns out to be useful in the study of certain properties of transformations in $\mathfrak{E}(\mathfrak{X})$.

In § 3 the inverse image $\pi^{-1}(\mathfrak{P}_1)$ of the radical is characterized. One formulation for this is that $\pi^{-1}(\mathfrak{P}_1)$ is the set of all $U \in \mathfrak{E}(\mathfrak{X})$ such that $(T+U)(\mathfrak{X})$ is closed and $(T+U)^{-1}(0)$ is finite-dimensional for all T which are regular in $\mathfrak{E}(\mathfrak{X})$.

A well-known result of Schauder [**13**] asserts that if I is the identity in $\mathfrak{E}(\mathfrak{X})$, and $U \in \mathfrak{K}(\mathfrak{X})$, then $I+U$ and its adjoint $I^* + U^*$ have the same (finite) nullity. In § 4 we obtain a generalization of this result as a reflection of the internal structure of $\mathfrak{E}(\mathfrak{X}) - \pi^{-1}(\mathfrak{P}_1)$. Let $\mathfrak{C}$ be any subset of $\mathfrak{E}(\mathfrak{X})$ containing

Received May 7, 1953. Presented to the American Mathematical Society, February 23, 1952.

Pacific J. Math. 4 (1954), 615-636

I such that (1) $\pi(\mathfrak{C})$ is a multiplicative group and (2) the closure of the component of $\mathfrak{C}$ containing I intersects $\pi^{-1}(\mathfrak{P}_1)$. Then there is a subring $\mathfrak{C}_1$ of $\mathfrak{E}(\mathfrak{X})$ where the images of $\mathfrak{C}$ and $\mathfrak{C}_1$ are geometrically related in $\mathfrak{E}(\mathfrak{X}) - \pi^{-1}(\mathfrak{P}_1)$ such that (a) $\mathfrak{G}_1 \supset \pi^{-1}(\mathfrak{P}_1)$, (b) $\pi(\mathfrak{G}_1)$ is a group under the circle operation (see § 4) where for each $T \in \mathfrak{G}$, $U \in \mathfrak{G}_1$ the quantities nul T, nul T^*, and nul $(T + U)$, nul $(T^* + U^*)$ are all finite and

$$\text{nul}\,(T^* + U^*) - \text{nul}\,(T + U) = \text{nul}\,(T^*) - \text{nul}\,(T).$$

For $\mathfrak{G}$ the set of nonzero scalar multiples of I this result already improves Schauder's, for there

$$\mathfrak{G}_1 = \pi^{-1}(\mathfrak{P}_1);$$

and since

$$\text{nul}\,(I) = \text{nul}\,(I^*) = 0$$

we have

$$\text{nul}\,(I^* + U^*) = \text{nul}\,(I + U)$$

for every $U \in \pi^{-1}(\mathfrak{P}_1)$.

Let

$$f(T) = \text{nul}\,(T^*) - \text{nul}\,(T).$$

This is known [**1**, **15**] to be defined (finite) for the inverse image under π of the set of regular elements of $\mathfrak{E}(\mathfrak{X}) - \mathfrak{K}(\mathfrak{X})$. Atkinson [**1**] has shown that the equation $f(TU) = f(T) + f(U)$ is satisfied. In § 5 this is obtained as an application of the theory of functionals on an abstract semi-group. These considerations lead in § 6 to a detailed study of the relation of the sets in $\mathfrak{E}(\mathfrak{X})$ of elements with a one-sided or two-sided inverse to the corresponding sets, in $\mathfrak{E}(\mathfrak{X}) - \mathfrak{K}(\mathfrak{X})$.

2. Notation and preliminaries. Let $\mathfrak{X}$ be an infinite-dimensional Banach space and let $\mathfrak{E}(\mathfrak{X})$ be the algebra of all bounded linear transformations defined on $\mathfrak{X}$ into $\mathfrak{X}$ made into a Banach algebra by the usual definition of the norm of a transformation [**7**, p. 32] and with identity I. Let $\mathfrak{K}(\mathfrak{X})$ be the subset of $\mathfrak{E}(\mathfrak{X})$ consisting of the completely continuous transformations in $\mathfrak{E}(\mathfrak{X})$. It is well known [**2**, p. 96] that $\mathfrak{K}(\mathfrak{X})$ is a closed two-sided ideal in $\mathfrak{E}(\mathfrak{X})$. Thus under the usual definitions [**7**, p. 472] the difference algebra $\mathfrak{E}(\mathfrak{X}) - \mathfrak{K}(\mathfrak{X})$ is a

Banach algebra. Let π be the canonical homomorphism of $\mathfrak{E}(\mathfrak{X})$ into $\mathfrak{E}(\mathfrak{X})-\mathfrak{K}(\mathfrak{X})$. Let $\mathfrak{P}_1$ be the radical of $\mathfrak{E}(\mathfrak{X})-\mathfrak{K}(\mathfrak{X})$ [7, p. 476], and let $\mathfrak{W}(\mathfrak{X})$ be any closed two-sided ideal of $\mathfrak{E}(\mathfrak{X})$ contained in $\pi^{-1}(\mathfrak{P}_1)$ and containing $\mathfrak{K}(\mathfrak{X})$. Let τ be the canonical homomorphism of $\mathfrak{E}(\mathfrak{X})$ onto $\mathfrak{E}(\mathfrak{X})-\mathfrak{W}(\mathfrak{X})$.

2.1. LEMMA. *$T \in \mathfrak{E}(\mathfrak{X})$ has a left (right) inverse modulo $\mathfrak{W}(\mathfrak{X})$ if and only if T has a left (right) inverse modulo $\mathfrak{K}(\mathfrak{X})$.*

Proof. Suppose that T has a left inverse modulo $\mathfrak{W}(\mathfrak{X})$. Thus there exists $U \in \mathfrak{E}(\mathfrak{X})$, $V \in \mathfrak{W}(\mathfrak{X})$ such that $UT = I + V$. Now $V \in \pi^{-1}(\mathfrak{P}_1)$ so that $I + V$ has a two-sided inverse W modulo $\mathfrak{K}(\mathfrak{X})$. Hence WU is the desired left inverse of T modulo $\mathfrak{K}(\mathfrak{X})$.

It may be noted that since $\mathfrak{P}_1$ is closed in $\mathfrak{E}(\mathfrak{X})-\mathfrak{K}(\mathfrak{X})$ then $\pi^{-1}(\mathfrak{P}_1)$ is a closed two-sided ideal in $\mathfrak{E}(\mathfrak{X})$.

2.2. LEMMA. *$T \in \mathfrak{E}(\mathfrak{X})$ has the properties that $T(\mathfrak{X})$ is closed and its null-space is finite-dimensional if and only if T takes each bounded set which is not conditionally compact onto a set which is not conditionally compact.*

Lemma 2.2 is a rewording of [**15**, Lemma 3.1].

If the null-space of T is finite-dimensional, its dimension is designated by nul T. A transformation with the properties of Lemma 2.2 is said in [**15**] to have *property A*.

2.3. LEMMA. *$T \in \mathfrak{E}(\mathfrak{X})$ has a two-sided inverse modulo $\mathfrak{W}(\mathfrak{X})$ if and only if both T and T^* have property A.*

Proof. By Lemma 2.1 we may take $\mathfrak{K}(\mathfrak{X})$ for $\mathfrak{W}(\mathfrak{X})$. The result then follows immediately from the results of [**15**, § 5] (see also [**1**, Theorem 1] and [**6**]).

If both T and T^* have property A we define

$$f(T) = \text{nul } T^* - \text{nul } T.$$

Here T^* is the adjoint of T Let $\mathfrak{H}$ be the set of all such transformations. By Lemma 2.3, $\mathfrak{H}$ is a semi-group.

2.4. LEMMA. *The function $f(T)$ is a continuous function on $\mathfrak{H}$. If T and U lie in the same component of $\mathfrak{H}$, then $f(T) = f(U)$.*

Proof. The continuity of f follows from the work of Dieudonne′ [**4**, proposition 4]; see also [**15**, Theorem 3.8] and [**1**, Theorem 4]. Since f is integer-valued, the second statement follows.

2.5. LEMMA. *If $T \in \mathfrak{H}$ and if $U - T \in \mathfrak{B}(\mathfrak{X})$ then $U \in \mathfrak{H}$ and $f(T) = f(U)$.*

Proof. It is clear that U has a two-sided inverse modulo $\mathfrak{F}(\mathfrak{X})$ if T does, by Lemma 2.1. That $f(T) = f(U)$ follows from Lemma 2.4 since the set $T + \mathfrak{B}(\mathfrak{X})$ is a connected subset of $\mathfrak{H}$.

We adopt the following notation used by Rickart [**12**] for a Banach algebra. An element is *left* (*right*) *regular* provided that it possesses a left (right) inverse in the algebra. If the element is both left and right regular then it possesses a unique two-sided inverse and is said to be regular. For $\mathfrak{E}(\mathfrak{X})$ we designate the sets of left regular, right regular, and regular elements by $\mathfrak{G}^l$, $\mathfrak{G}^r$, and $\mathfrak{G}$, respectively. The corresponding sets in $\mathfrak{E}(\mathfrak{X}) - \mathfrak{B}(\mathfrak{X})$ are designated by $\mathfrak{G}_1^l$, $\mathfrak{G}_1^r$, and $\mathfrak{G}_1$, respectively. In the foregoing notation, $\mathfrak{H} = \tau^{-1}(\mathfrak{G}_1)$.

Thus, by Lemmas 2.3 and 2.4, f defines a mapping of $\mathfrak{G}_1$ into the set of integers. This mapping will also be designated by f.

2.6. LEMMA. *Let $T \in \mathfrak{H}$, $f(T) = 0$. Then T can be expressed as the sum $U + V$ where $U \in \mathfrak{G}$, $V \in \mathfrak{F}(\mathfrak{X})$.*

Proof. This is given in [**15**, Corollary 3.11].

3. On the radical of $\mathfrak{E}(\mathfrak{X}) - \mathfrak{B}(\mathfrak{X})$.

In view of Lemma 2.1 and the definition of the radical of $\mathfrak{E}(\mathfrak{X}) - \mathfrak{B}(\mathfrak{X})$, the inverse image under τ of the radical of $\mathfrak{E}(\mathfrak{X}) - \mathfrak{B}(\mathfrak{X})$ is the same set as $\pi^{-1}(\mathfrak{P}_1)$, where $\mathfrak{P}_1$ is the radical of $\mathfrak{E}(\mathfrak{X}) - \mathfrak{F}(\mathfrak{X})$. In this section we determine the nature of $\pi^{-1}(\mathfrak{P}_1)$.

3.1. LEMMA. *Let $T \in \mathfrak{E}(\mathfrak{X})$ be an isomorphism between $\mathfrak{X}$ and a proper closed linear manifold of $\mathfrak{X}$. Then there exists a sphere in $\mathfrak{E}(\mathfrak{X})$ with center T each of whose elements have this property.*

Proof. By [**4**, proposition 1] there is a sphere $\mathfrak{S}$ about T such that for all U in $\mathfrak{S}$, U is bi-continuous. But T is in the interior of the set of elements of $\mathfrak{E}(\mathfrak{X})$ which are not regular [**14**, Corollary 2.2]. Hence for each $U \in \mathfrak{S}$ there is a proper closed linear manifold $\mathfrak{N}$ of $\mathfrak{X}$ such that U is an isomorphism of $\mathfrak{X}$ onto $\mathfrak{N}$ if the radius of $\mathfrak{S}$ is sufficiently small.

3.2. LEMMA. *Let $T \in \mathfrak{E}(\mathfrak{X})$ have range $\mathfrak{X}$ where T is not one-to-one. Then* *there is a sphere in* $\mathfrak{E}(\mathfrak{X})$ *with center* T *each of whose elements has these properties.*

Proof. This is shown in the same way by use of [**4**, Theorem 1] and [**14**, Corollary 3.12].

3.3. LEMMA. *Let* $T \in \mathfrak{E}(\mathfrak{X})$. *Suppose that* $T(T^*)$ *has property* A *while* $T^*(T)$ *does not. Then* T *can be expressed in the form* $T_1 + V$ *where* $V \in \mathfrak{R}(\mathfrak{X})$ *and* T_1 *is bi-continuous* ($T_1(\mathfrak{X}) = \mathfrak{X}$).

Proof. This is contained in [**15**, Theorem 3.13].

3.4. THEOREM. *Let* $T \in \mathfrak{G}$. *Suppose that for each* α $(0 < \alpha \leq 1)$ *either* $T + \alpha U$ *or* $T^* + \alpha U^*$ *has property* A. *Then* $T + \alpha U \in \mathfrak{H}$ $(0 \leq \alpha \leq 1)$ *and* $f(T + U) = 0$.

Proof. Note that $f(T) = 0$. The set $\mathfrak{H} = \pi^{-1}(\mathfrak{G}_1)$ is open in $\mathfrak{E}(\mathfrak{X})$. Thus either all the $T + \alpha U$ $(0 \leq \alpha \leq 1)$ are in $\mathfrak{H}$ or there is a smallest number β $(0 < \beta \leq 1)$ such that $T + \beta U \notin \mathfrak{H}$. In the latter case one of $T + \beta U$, $T^* + \beta U^*$ has property A but not the other. Suppose that $T + \beta U$ has property A. Then, by Lemma 3.3, $T + \beta U$ can be written in the form $T_1 + V$, where $T_1 \in \mathfrak{E}(\mathfrak{X})$ is bi-continuous and $V \in \mathfrak{R}(\mathfrak{X})$. If $T_1(\mathfrak{X}) = \mathfrak{X}$ then $T_1 \in \mathfrak{G}$ and thus $T + \beta U \in \mathfrak{H}$, contrary to the above. Thus $T_1 = T + \beta U - V$ an isomorphism between $\mathfrak{X}$ and a proper closed linear manifold of $\mathfrak{X}$. Consequently, by Lemma 3.1, if $0 < \alpha < \beta$, and $\beta - \alpha$ is sufficiently small, then $T + \alpha U - V$ has this property. But for such α, $T + \alpha U \in \mathfrak{H}$. Also, by Lemma 2.5, $T + \alpha U - V \in \mathfrak{H}$ and

$$f(T + \alpha U) = f(T + \alpha U - V).$$

Since

$$\operatorname{nul}(T + \alpha U - V) = 0, \ \operatorname{nul}(T^* + \alpha U^* - V^*) > 0,$$

then

$$f(T + \alpha U) > 0.$$

However, since $f(T) = 0$, by Lemma 2.4 we have

$$f(T + \alpha U) = 0.$$

This contradiction establishes the result if $T + \beta U$ has property A. If $T^* + \beta U^*$ has property A then we proceed in a same way using dual results (Lemmas 3.2

and 3.3) to see that for $\alpha < \beta$ and close to β,

$$f(T+\alpha U)=0, \quad f(T+\alpha U) < 0.$$

Thus we conclude that $T+\alpha U \in \mathfrak{H}$ $(0 \leq \alpha \leq 1)$. That

$$f(T+U)=f(T)=0$$

follows from Lemma 2.4.

3.5. THEOREM. *The following formulas for* $\pi^{-1}(\mathfrak{P}_1)$ *hold*:

(a) $\pi^{-1}(\mathfrak{P}_1)=\{U \in \mathfrak{E}(\mathfrak{X}) \mid$ *for each* $T \in \mathfrak{G}$ *either* $T+U$ *or* T^*+U^* *has property* $A\}$;

(b) $\pi^{-1}(\mathfrak{P}_1)=\{U \in \mathfrak{E}(\mathfrak{X}) \mid T+U$ *has property* A *for each* $T \in \mathfrak{G}\}$;

(c) $\pi^{-1}(\mathfrak{P}_1)=\{U \in \mathfrak{E}(\mathfrak{X}) \mid T^*+U^*$ *has property* A *for each* $T \in \mathfrak{G}\}$.

Proof. If $T \in \mathfrak{G}$ and $U \in \pi^{-1}(\mathfrak{P}_1)$ then $\pi(T) \in \mathfrak{G}_1$ and

$$\pi(T+U)=\pi(T)+\pi(U) \in \mathfrak{G}_1,$$

by the definition of $\mathfrak{P}_1$. Then $T+U \in \mathfrak{H}$ and it follows that $\pi^{-1}(\mathfrak{P}_1)$ is contained in each of the sets on the right.

Let the set on the right side of (a) be denoted by $\mathfrak{S}$. Then if $T \in \mathfrak{G}$, $U \in \mathfrak{S}$, $\alpha \neq 0$ a scalar, then $\alpha T+U$ or αT^*+U^* has property A. Hence, for each scalar α, $T+\alpha U$ or $T^*+\alpha U^*$ has property A. Theorem 3.4 shows that $T+\alpha U \in \mathfrak{H}$ for all scalars α. Next we show that if $W \in \mathfrak{G}$, $U \in \mathfrak{S}$ then $UW \in \mathfrak{S}$. Both W and W^* have property A. Hence, by the nature of $\mathfrak{S}$ and [**15**, Theorem 3.4], for each $T \in \mathfrak{G}$ either

$$(TW^{-1}+U)W = T+UW$$

has property A or

$$W^*[(TW^{-1})^*+U^*]=T^*+(UW)^*$$

has property A. Hence $UW \in \mathfrak{S}$.

Next let $U_i \in \mathfrak{S}$, $i=1,2$. For each $T \in \mathfrak{G}$, by the above $T+\alpha U_1 \in \mathfrak{H}$ for $0 \leq \alpha \leq 1$ and, by Theorem 3.4, $f(T+U_1)=0$. By Lemma 2.6, $T+U_1$ can be expressed in the form T_1+V, where $T_1 \in \mathfrak{G}$ and $V \in \mathfrak{R}(\mathfrak{X})$. Likewise $T_1+U_2 \in \mathfrak{H}$ and so, by Lemma 2.5,

$$T_1 + U_2 + V = T + (U_1 + U_2)$$

is in $\mathfrak{H}$. This shows that $\mathfrak{S}$ is a linear manifold in $\mathfrak{E}(\mathfrak{X})$ with the further property that if $T_1, T_2 \in \mathfrak{G}$ and $U \in \mathfrak{S}$ then $U(T_1 - T_2) \in \mathfrak{S}$. However, since $\mathfrak{E}(\mathfrak{X})$ is a Banach algebra, an arbitrary element $W \in \mathfrak{E}(\mathfrak{X})$ can be expressed as the difference of two regular elements. Thus $\mathfrak{S}$ is a right ideal in $\mathfrak{E}(\mathfrak{X})$. Consequently $\pi(\mathfrak{S})$ has the property that, for each $\pi(U) \in \pi(\mathfrak{S})$ and each V in $\mathfrak{E}(\mathfrak{X}) - \mathfrak{K}(\mathfrak{X})$, $\pi(I) + \pi(T)V \in \mathfrak{G}_1$. Thus $\pi(\mathfrak{S}) \subset \mathfrak{P}_1$. This completes the proof for formula (a).

The same argument shows that the right sides of (b) and (c) are contained in $\pi^{-1}(\mathfrak{P}_1)$.

3.6. COROLLARY. *Let $\mathfrak{L}$ be a (left or right) ideal in $\mathfrak{E}(\mathfrak{X})$. Suppose that for each $T \in \mathfrak{L}$, either $I + T$ or $I^* + T^*$ has property A. Then for each $T \in \mathfrak{L}$,* nul $(I + T)$ *and* nul $(I^* + T^*)$ *are finite and equal.*

Proof. By Theorem 3.5, $\mathfrak{L} \subset \pi^{-1}(\mathfrak{P}_1)$. Thus $I + T \in \mathfrak{H}$ for each $T \in \mathfrak{L}$. Since $\mathfrak{L}$ is a linear manifold,

$$f(I + T) = f(I) = 0$$

by Lemma 2.4.

This is a direct generalization of Schauder's well-known result [**13**, p. 189] that if U is completely continuous then

$$\text{nul}\,(I + U) = \text{nul}\,(I^* + U^*)$$

since the two-sided ideal $\mathfrak{K}(\mathfrak{X})$ fulfills the conditions of Corollary 3.6.

3.7. COROLLARY. *The following statements are equivalent:*

(1) $\mathfrak{E}(\mathfrak{X}) - \mathfrak{K}(\mathfrak{X})$ *is semi-simple;*

(2) *for $U \in \mathfrak{E}(\mathfrak{X})$, $U \in \mathfrak{K}(\mathfrak{X})$ if and only if $(T + U)(\mathfrak{X})$ is closed in $\mathfrak{X}$ and either* nul $(T + U)$ *or* nul $(T^* + U^*)$ *is finite for each T regular in $\mathfrak{E}(\mathfrak{X})$.*

Proof. Note that $\mathfrak{E}(\mathfrak{X}) - \mathfrak{K}(\mathfrak{X})$ is semi-simple if and only if $\pi^{-1}(\mathfrak{P}_1) = \mathfrak{K}(\mathfrak{X})$. Also $(T + U)(\mathfrak{X})$ is closed if and only if $(T^* + U^*)(\mathfrak{X}^*)$ is closed in $\mathfrak{X}^*$ [**2**, Chapt. 10]. Then Corollary 3.7 follows from Theorem 3.5 and Lemma 2.3.

If $\mathfrak{X}$ is a separable Hilbert space then since, as shown by Calkin [**3**, Theorem 1.4], $\mathfrak{K}(\mathfrak{X})$ is a maximal, two-sided ideal in $\mathfrak{E}(\mathfrak{X})$, (1) holds. For spaces satisfying (1), (2) gives a necessary and sufficient condition for complete continuity which seems to be new (for sufficiency) even in the Hilbert space case.

4. A generalized Schauder nullity theorem. We give here the result (Theorem 4.5) discussed in § 1. The preliminary material, it is felt, is of independent interest and is presented in greater generality than is absolutely necessary for our purposes.

We adopt the following notation. B is a ring with an identity element e. G is the set of regular elements of B (the elements with a two-sided inverse). For each subgroup G_0 of G let $\mathfrak{J}(G_0)$ be the set of "invariant translations" of G_0, namely the set of $x \in B$ such that $G_0 + x = G_0$. It is clear that

$$\mathfrak{J}(G_0) = \{x \in B \mid y \pm x \in G_0 \text{ for every } y \in G_0\}.$$

In the ring B we consider along with the usual algebraic operations also the "circle operation"

$$x \circ y = x + y - xy.$$

For information on this operation see [7, Chapter 22]. It is evident that $G_0 \cap \mathfrak{J}(G_0)$ is empty.

4.1. THEOREM. *For any subgroup G_0 of G, $\mathfrak{J}(G_0)$ is a subring of B which is a group under the circle operation. Conversely if R is a subring of B which is a group under the circle operation then there exists a subgroup G_0 of G such that $R = \mathfrak{J}(G_0)$. If B is a Banach algebra then $\mathfrak{J}(G)$ is the radical of B.*

Proof. It is clear that if $x \in \mathfrak{J}(G_0)$ then so does $-x$. Thus if x_1 and x_2 lie in $\mathfrak{J}(G_0)$, and $y \in G_0$, then both

$$(y + x_1) + x_2 \quad \text{and} \quad (y - x_1) - x_2$$

lie in G_0, so that $x_1 + x_2 \in \mathfrak{J}(G_0)$. Next we show if $x \in \mathfrak{J}(G_0)$, $y \in G_0$, then $yx \in \mathfrak{J}(G_0)$. For let $z \in G_0$. Then

$$z \pm yz = y(y^{-1}z \pm x) \in G_0.$$

Similarly $xy \in \mathfrak{J}(G_0)$. Since

$$y \pm x_1x_2 = (y + x_1)(e \pm x_2) \mp yx_2 - x_1$$

it follows from the above that $x_1x_2 \in \mathfrak{J}(G_0)$ if x_1 and $x_2 \in \mathfrak{J}(G_0)$. Thus $\mathfrak{J}(G_0)$ is a subring of B.

To see that $\mathfrak{I}(G_0)$ is a group under the circle operation note first that for $x_1, x_2 \in \mathfrak{I}(G_0)$ we have

$$x_1 \circ x_2 = x_1 + x_2 - x_1 x_2 \in \mathfrak{I}(G_0).$$

Now the set of all elements of B with an inverse under the circle operation is a group with the zero element θ of B as the identity element [7, p. 456]. Thus it is sufficient to show that x_1 has an inverse in $\mathfrak{I}(G_0)$ under this operation. Since $e - x_1 \in G_0 \subset G$ there exists an element $w \in B$ such that

$$(e - x_1)(e - w) = (e - w)(e - x_1) = e.$$

Then clearly w is the inverse of x_1 under this operation. Let $y \in G_0$. Then, since

$$x_1 w = w x_1 = x_1 + w$$

we have that

$$(y \pm w)(e - x_1) = y \pm w - y x_1 \mp w x_1 = y(e - x_1) \mp x_1$$

is an element of G_0. Since $(e - x_1) \in G_0$ it follows that $w \in \mathfrak{I}(G_0)$.

Next consider a subring R which is a group under the circle operation. Let G_0 be the set of all elements of the form $e - x$, $x \in R$. If $x_1, x_2 \in R$ then

$$(e - x_1)(e - x_2) = e - x_1 \circ x_2 \in G_0.$$

There exists $z \in R$ such that

$$x_1 \circ z = z \circ x_1 = \theta.$$

Then

$$(e - x_1)(e - z) = (e - z)(e - x_1) = e$$

so that G_0 is a group. We show that $\mathfrak{I}(G_0) = R$. Take $x \in \mathfrak{I}(G_0)$. Then $e - x \in G_0$, and, by the definition of G_0, $x \in R$. On the other hand if $x \in R$, $y \in G_0$ then we may write $y = e - x_1$, where

$$x_1 \in R \text{ and } y \pm x = e - x_1 \pm x \in G_0$$

since R is a ring. Thus $x \in \mathfrak{I}(G_0)$ and $\mathfrak{I}(G_0) = R$.

Finally let B be a Banach algebra. If z is an arbitrary element of B then since, for a sufficiently small scalar λ,

$$e - \lambda z = w \in G\cdot$$

we may write z as the sum of two elements in G. By the above we see that for $x \in \mathfrak{J}(G)$, we have $zx \in \mathfrak{J}(G)$ and thus $e - zx \in G$. Hence x lies in the (Jacobson) radical Q of B. Conversely if $x \in Q$, then for each $w \in G$,

$$w \pm x = w(e \pm w^{-1}x) \in G,$$

so that $x \in \mathfrak{J}(G)$. This completes the proof.

4.2. COROLLARY. *In the notation of Theorem* 4.1, $\mathfrak{J}(G_0)$ *is a two-sided ideal in the subring* $R(G_0)$ *of* B *generated by* G_0 *and lies in the radical* Q *of* $R(G_0)$. *Examples exist for which* $\mathfrak{J}(G_0) = Q$ *and also for which* $\mathfrak{J}(G_0) \neq Q$.

Proof. By the arguments of Theorem 4.1, if $y \in R(G_0)$ then yx, $xy \in \mathfrak{J}(G_0)$ for each $x \in \mathfrak{J}(G_0)$ so that $\mathfrak{J}(G_0)$ is a two-sided ideal of $R(G_0)$. Since $e - yx \in G_0$ for every $y \in R(G_0)$, and G_0 is contained in the set of regular elements of $R(G_0)$, $\mathfrak{J}(G_0) \subset Q$. By Theorem 4.1, if B is a Banach algebra then $\mathfrak{J}(G) = Q$. Take next for B the ring of integers modulo 9. For G_0 take the set consisting of 1 and 8. Here $R(G_0) = B$ and the radical Q of B is the set $\{0, 3, 6\}$. On the other hand $\mathfrak{J}(G_0)$ consists of the zero element alone.

Following Kaplansky [8, p. 153] we call B a *metric ring* if to each element x there is associated a real number $|x|$ such that

$$|\theta| = 0,\ |x| > 0 \text{ if } x \neq \theta,\ |-x| = |x|,\ |x+y| \leq |x| + |y|,\ |xy| \leq |x|\,|y|.$$

Here $|x - y|$ is the metric of B. In this context the sets $\mathfrak{J}(G_0)$ possesses certain topological properties. (The metric ring to which the theory is applied is $\mathfrak{C}(\mathfrak{X}) - \mathfrak{K}(\mathfrak{X})$).

4.3. LEMMA. *If* G_0 *is open then* $\mathfrak{J}(G_0)$ *is closed. The following statements are equivalent.*

(1) $\mathfrak{J}(G_0) \subset \overline{G}_0$.

(2) $0 = \inf |y|,\ y \in G_0$.

(3) $\mathfrak{J}(G_0) \cap \overline{G_0}$ *is nonempty.*

Proof. Let G_0 be open. Suppose that $x_n \in \mathfrak{J}(G_0)$ $(n = 1, 2, 3, \cdots)$ and that $x_n \longrightarrow x$. Given any $y \in G_0$ there exists a sphere S of radius, say, $r > 0$ about y such that $S \subset G_0$. Consequently $S \pm x_n \subset G_0$ for each n. Take n so large that $|x - x_n| < r$. Then for such an integer n, $y \pm (x - x_n) \in S$ and thus

$$y \pm x = y \pm (x - x_n) \mp x_n \in G_0 .$$

Hence $x \in \mathfrak{Z}(G_0)$.

If (1) holds then so does (2) since $\theta \in \mathfrak{Z}(G_0)$. If (2) holds then (3) is clear for the same reason. Suppose that (3) holds. Let

$$w \in \mathfrak{Z}(G_0) \cap \overline{G_0},\ w = \lim y_n,\ y_n \in G_0 .$$

By Theorem 4.1, $w \circ x \in \mathfrak{Z}(G_0)$ for each $x \in \mathfrak{Z}(G_0)$. But

$$w \circ x = \lim (x + y_n - y_n x),$$

and by Theorem 4.1, $y_n + x - y_n x \in G_0$. Hence $w \circ x \in \overline{G_0}$. By Theorem 4.1 again there exists an element z in $\mathfrak{Z}(G_0)$ such that $w \circ z = \theta$. Inasmuch as $z \circ x \in \mathfrak{Z}(G_0)$, by the above

$$w \circ (z \circ x) = (w \circ z) \circ x = x$$

lies in $\overline{G_0}$.

For the group G_0 in the metric ring B let G_{0p} be the *principal component*, that is, that which contains e. Arguments of Hille [7, p. 93] show that G_{0p} is a subgroup of G_0.

4.4. LEMMA. *If* $\mathfrak{Z}(G_{0p}) \subset \overline{G}_{0p}$ *then* $\mathfrak{Z}(G_0)$ *is connected and* $\mathfrak{Z}(G_0) \subset \overline{G}_{0p}$. *If* $\mathfrak{Z}(G_0)$ *is connected, then* $\mathfrak{Z}(G_0) \subset \mathfrak{Z}(G_{0p})$.

Proof. Suppose that $\mathfrak{Z}(G_{0p}) \subset \overline{G}_{0p}$. Then by Lemma 4.3, $\theta \in \overline{G}_{0p}$. Take $x \in \mathfrak{Z}(G_0)$. The set xG_{0p}, being a continuous image of a connected set, is connected; moreover, xG_{0p} lies in $\mathfrak{Z}(G_0)$ by Corollary 4.2. Since θ lies in the closure of xG_{0p}, the set

$$F = xG_{0p} \cup \{\theta\}$$

is a connected subset of $\mathfrak{Z}(G_0)$ which contains x and θ. Hence each element of $\mathfrak{Z}(G_0)$ lies in a connected subset containing θ. Thus $\mathfrak{Z}(G_0)$ is connected.

Suppose that $\mathfrak{Z}(G_0)$ is connected. Then for each $z \in G_{0p}$, $z + \mathfrak{Z}(G_0)$ is a connected subset of G_0 containing z. Hence

$$z + \mathfrak{I}(G_0) \subset G_{0p} \text{ and } \mathfrak{I}(G_0) \subset \mathfrak{I}(G_{0p}).$$

In the statement of the following theorem, the group to which the symbol $\mathfrak{I}$ is applied lies in the Banach algebra $\mathfrak{E}(\mathfrak{X}) - \pi^{-1}(\mathfrak{P}_1)$.

4.5. THEOREM. *Let $\mathfrak{S}$ be any set in $\mathfrak{E}(\mathfrak{X})$ containing the identity I. Let π and τ be the canonical homomorphisms of $\mathfrak{E}(\mathfrak{X})$ onto $\mathfrak{E}(\mathfrak{X}) - \mathfrak{R}(\mathfrak{X})$ and $\mathfrak{E}(\mathfrak{X}) - \pi^{-1}(\mathfrak{P}_1)$, respectively. Suppose that $\pi(\mathfrak{S})$ is a multiplicative group in $\mathfrak{E}(\mathfrak{X}) - \mathfrak{R}(\mathfrak{X})$ and that the closure of the component of $\mathfrak{S}$ containing I contains an element of $\pi^{-1}(\mathfrak{P}_1)$. Then for each $T \in \mathfrak{S}$, $U \in \tau^{-1}\,\mathfrak{I}[\tau(\mathfrak{S})]$ we have*

$$f(T) = f(T + U).$$

Furthermore, $\tau^{-1}\,\mathfrak{I}[\tau(\mathfrak{S})] \supset \pi^{-1}(\mathfrak{P}_1)$, and is the inverse image under π of a subring of $\mathfrak{E}(\mathfrak{X}) - \mathfrak{R}(\mathfrak{X})$ which is a group under the circle operation.

Proof. Consider $\tau(\mathfrak{S})$. By Lemma 2.1 it is a subgroup of the set of regular elements $\mathfrak{G}_1$ of $\mathfrak{E}(\mathfrak{X}) - \pi^{-1}(\mathfrak{P}_1)$. Since τ is continuous, by our hypothesis the principal component of $\tau(\mathfrak{S})$ contains the zero element of $\mathfrak{E}(\mathfrak{X}) - \pi^{-1}(\mathfrak{P}_1)$ in its closure. Hence in this algebra, by Lemmas 4.3 and 4.4, $\mathfrak{I}[\tau(\mathfrak{S})]$ is connected. By Lemma 2.4, f is continuous on $\mathfrak{G}_1$; and if $T_1 \in \tau(\mathfrak{S})$, $U_1 \in \mathfrak{I}[\tau(\mathfrak{S})]$ then since T_1 and $T_1 + U_1$ lie in the same component of $\mathfrak{G}_1$, we have

$$f(T_1 + U_1) = f(T_1).$$

Thus $f(T + U) = f(T)$ if $T \in \mathfrak{S}$ and $U \in \tau^{-1}[\mathfrak{I}(\tau(\mathfrak{S})]$.

Let

$$\tau^{-1}\,\mathfrak{I}[\tau(\mathfrak{S})] = \mathfrak{S}_1 \text{ and } \pi(\mathfrak{S}_1) = \mathfrak{S}_2.$$

Clearly $\pi^{-1}(\mathfrak{S}_2) = \mathfrak{S}_1$ since $\mathfrak{S}_1 \supset \mathfrak{R}(\mathfrak{X})$ which is the kernel of π. By Theorem 4.1, $\mathfrak{I}[\tau(\mathfrak{S})]$ is a subring of $\mathfrak{E}(\mathfrak{X}) - \pi^{-1}(\mathfrak{P}_1)$ which is a group under the circle Operation. Then $\mathfrak{S}_1$ is a subring of $\mathfrak{E}(\mathfrak{X})$, and $\mathfrak{S}_2$ a subring of $\mathfrak{E}(\mathfrak{X}) - \mathfrak{R}(\mathfrak{X})$. We next show that $\mathfrak{S}_2$ is a group under the circle operation. As $\mathfrak{S}_2$ is a subring, it is closed under that operation. Let $T_1 \in \mathfrak{S}_2$, $T_1 = \pi(T)$, $T \in \mathfrak{S}_1$. Then there exists $V \in \mathfrak{S}_1$ such that

$$[\tau(I) - \tau(V)][\tau(I) - \tau(T)] = [\tau(I) - \tau(T)][\tau(I) - \tau(V)] = \tau(I).$$

Then by Lemma 2.1, $I - T$ has a two-sided inverse $I - W$ modulo $\mathfrak{R}(\mathfrak{X})$. Since

$$T_1 \circ \pi(W) = \pi(W) \circ T_1 = 0$$

it suffices to show that $\pi(W) \in \mathfrak{S}_2$. Now $\tau(W) = \tau(V)$ since the two-sided inverse of $\tau(I - T)$ in $\mathfrak{E}(\mathfrak{X}) - \pi^{-1}(\mathfrak{P}_1)$ is unique. Therefore $W \in \mathfrak{S}_1$ and thus $\pi(W) \in \mathfrak{S}_2$.

5. Functionals on semi-groups. Atkinson [**1**] has shown that on $\mathfrak{H}$ the equation

$$f(TU) = f(T) + f(U)$$

is valid. By an entirely different analysis we show how such functionals can be obtained in a semi-group and then apply the results to $\mathfrak{H}$.

5.1. NOTATION. Let S be any semi-group, the product of two elements x, y in S being denoted by xy. Let g and g^* be real-valued functions defined on S, where

$$(1) \qquad \begin{aligned} &g(x_2) \leq g(x_1 x_2) \leq g(x_1) + g(x_2) \\ &g^*(x_1) \leq g^*(x_1 x_2) \leq g^*(x_1) + g^*(x_2) \end{aligned}$$

for all x_1, x_2 in S. Let

$$h(x) = g^*(x) - g(x),$$

and let S_+ (S_-) be the subset of S for which $h(x) \geq 0$ ($h(x) \leq 0$). Suppose that there is a reflexive and symmetric relation $\sim$ on S defined for certain pairs of elements of S such that $x \sim y$ implies $h(x) = h(y)$, and where for each $x \in S$ there exists $y \in S$, $x \sim y$ with either $g(y) = 0$ or $g^*(y) = 0$, The relation $\sim$ need not be transitive. Since g and g^* are nonnegative on S it follows that the existence of y, $x \sim y$, where $g(y)=0$ ($g^*(y)=0$), is equivalent to $x \in S_+$ ($x \in S_-$).

5.2. THEOREM. *Suppose that, in the notation of* 5.1,

(a) $x_i \sim z_i$ $(i = 1, 2)$ *implies that* $h(x_1 x_2) = h(z_1 z_2)$ *holds. Then the formula*

$$(2) \qquad h(x_1 x_2) = h(x_1) + h(x_2)$$

is valid either for all $x_1 \in S_+$ *or for all* $x_2 \in S_-$. *If also*

(b) *there exist* y_1, y_2 *in* S, *where* $h(y_1) > 0$ *and* $h(y_2) < 0$, *then formula* (2) *is valid on* S.

Formula (2) *is valid on* S *if* (a) *holds and*

(c_1) *for each* $x \in S_+$ *there exists* $y \in S$ *such that* $xy \in S_-$,

(c_2) *for each* $x \in S_-$ *there exists* $y \in S$ *such that* $yx \in S_+$.

Proof. We remark that (a) is a necessary condition for (2) since, from (2),

$$h(x_1 x_2) = h(x_1) + h(x_2) = h(z_1) + h(z_2) = h(z_1 z_2).$$

From (1) we obtain

$$g^*(x_1) - g(x_1) - g(x_2) \le g^*(x_1 x_2) - g(x_1 x_2) \le g^*(x_1) + g^*(x_2) - g(x_2)$$

or

$$(3) \qquad h(x_1) - g(x_2) \le h(x_1 x_2) \le h(x_2) + g^*(x_1)$$

Now suppose that (a) holds. Then

$$(4) \qquad h(x_1) \le h(x_1 x_2) \le h(x_1) + h(x_2) \qquad x_1, x_2 \in S_+,$$

$$(5) \qquad h(x_1) + h(x_2) \le h(x_1 x_2) \le h(x_2) \qquad x_1, x_2 \in S_-,$$

$$(6) \qquad h(x_1 x_2) = h(x_1) + h(x_2) \qquad x_1 \in S_+,\ x_2 \in S_-.$$

To show (4) we may assume that

$$g(x_i) = 0,\ g^*(x_i) = h(x_i) \qquad (i = 1, 2).$$

Then (4) follows from (3). For (5) we may assume that

$$-g(x_i) = h(x_i),\ g^*(x_i) = 0 \qquad (i = 1, 2),$$

and again use (3). In the last situation, (3) yields

$$h(x_1) + h(x_2) \le h(x_1 x_2) \le h(x_1) + h(x_2).$$

Next we observe that (c_1) and (c_2) cannot both be false. If, for example, (c_1) is false then for some $x_1 \in S_+$ we have $x_1 y \in S_+$ for all $y \in S$, which yields (c_2).

Suppose now that (a) and (c_2) hold. We show that (2) holds for all x_1, x_2 where $x_2 \in S_-$. By (6) we may suppose that $x_1 \in S_-$. There exists $w \in S$ such that $h(wx_1) \ge 0$. For case 1 we take $w \in S_-$. Then by (5),

$$h(w) + h(x_1) \le h(wx_1) \le h(x_1) \le 0.$$

This implies that $h(x_1) = 0$. Then (2) follows from (6). For case 2 we take $w \in S_+$. This gives, by (6),

$$h(wx_1) = h(w) + h(x_1), \tag{7}$$

$$h(wx_1 x_2) = h(wx_1) + h(x_2). \tag{8}$$

Now (5) shows that $x_1 x_2 \in S_-$. Then, by (6),

$$h(wx_1 x_2) = h(w) + h(x_1 x_2). \tag{9}$$

A combination of (7), (8), and (9) yields (2).

Suppose next that (a) and (c_1) hold. Entirely analogous arguments using (4) in place of (5) show that (2) holds for all x_1, x_2 where $x_1 \in S_+$.

Now assume (a) and (b). We show that (c_1) and (c_2) hold. If (c_1) does not hold then (c_2) must hold and there exists $x \in S_+$ such that $xy \in S_+$ for all $y \in S$. Select y such that $h(y) < 0$. By (a) and (c_2) and the above, $h(y^n) = n\, h(y)$ for any positive integer n and thus $y^n \in S_-$. Also

$$0 \leq h(xy^n) = h(x) + n\, h(y).$$

This is impossible if n is chosen sufficiently large. Thus (c_1) holds. Similarly (c_2) holds.

To conclude the proof we show that (a), (c_1), and (c_2) imply (2). By the above our assumptions give the validity of (2) for any pair x_1, x_2 where either $x_1 \in S_+$ or $x_2 \in S_-$. The remaining case involves $x_1 \in S_-$ and $x_2 \in S_+$. We may select, by (c_2), $w \in S$ such that $wx_1 \in S_+$. If $w \in S_-$ then, as shown above, $h(x_1) = 0$ so that (2) is valid for x_1, x_2. Supposing that $w \in S_+$, we obtain (7), (8), and (9), which again yield (2) for x_1, x_2.

We return to $\mathfrak{G}(\mathfrak{X})$ and start with the following simple result:

5.3. LEMMA. *Let $T_i \in \mathfrak{G}(\mathfrak{X})$ $(i = 1, 2)$ have finite nullity. Then*

$$\text{nul}(T_2) \leq \text{nul}(T_1 T_2) \leq \text{nul}(T_1) + \text{nul}(T_2). \tag{10}$$

This follows from the fact, readily established, that

$$\text{nul}(T_1 T_2) = \text{nul}(T_2) + \dim [T_2(\mathfrak{X}) \cap T_1^{-1}(0)].$$

5.4. LEMMA. *Suppose that $T \in \mathfrak{H}$ and $f(T) \geq 0$ (≤ 0). Then there exists $V \in \mathfrak{H}$ such that $V - T \in \mathfrak{K}(\mathfrak{X})$, $f(T) = f(V)$, and* $\text{nul}(V) = 0$ $(\text{nul}(V^*) = 0)$.

The existence of the transformation V with the indicated property of the nullity follows from [**15**, Theorem 3.13]. That $f(T) = f(V)$ follows from Lemma 2.5.

5.5. COROLLARY. *Let $T_i \in \mathfrak{H}$ $(i = 1, 2)$. Then $f(T_1 T_2) = f(T_1) + f(T_2)$, and f defines a homomorphism of the group of regular elements of $\mathfrak{E}(\mathfrak{X}) - \mathfrak{K}(\mathfrak{X})$ into the additive group of integers.*

We show that this result of Atkinson follows from the above. In the notation of 5.1, set

$$S = \mathfrak{H},\ g^*(T) = \text{nul}\ (T^*),\ \ g(T) = \text{nul}\ (T).$$

Since

$$(T_1 T_2)^* = T_2^* T_1^*,$$

Lemma 5.3 shows that formula (1) is valid. For the relation $T_1 \sim T_2$ we take $T_1 - T_2 \in \mathfrak{K}(\mathfrak{X})$. Lemmas 3.2, 2.4, and 5.4 and the relation

$$f(T) = \text{nul}\ (T^*) - \text{nul}\ (T)$$

show that Theorem 5.2 may be applied to give the first conclusion. The second conclusion is an immediate consequence.

Following ideas of Mackey [**10**, p. 171] we shall say that the Banach space $\mathfrak{X}$ is *stable* if there exists a continuous isomorphism of $\mathfrak{X}$ onto a closed subspace $\mathfrak{X}_1$ of deficiency one. We say that $\mathfrak{X}$ is *stable-like* if there exists a continuous isomorphism of $\mathfrak{X}$ onto a closed subspace $\mathfrak{X}_1$ of finite deficiency.

5.6. THEOREM. *The functional f is non-trivial if and only if $\mathfrak{X}$ is stable-like.*

Proof. If $\mathfrak{X}$ is stable-like, consider the isomorphism T of $\mathfrak{X}$ onto $\mathfrak{X}_1$ of deficiency n. Then nul $(T^*) = n$ and nul $(T) = 0$, so that $f(T) = n$.

Suppose that f is non-trivial. Then there exists $T \in \mathfrak{H}$ such that $f(T) \neq 0$. Since T has a two-sided inverse V modulo $\mathfrak{K}(\mathfrak{X})$, and $f(V) = -f(T)$ by Corollary 5.5, we may assume $f(T) = n > 0$. By Lemma 5.4, there exists a bicontinuous isomorphism U where nul $(U^*) = n$. Then $U(\mathfrak{X})$ is a closed subspace of deficiency n.

Whether or not every infinite-dimensional Banach space must be stable or even stable-like seems to be an open question (see [**10**, p. 205]). This subject is pursued a bit further in Theorem 6.7 and 6.9.

If $\mathfrak{X}$ is finite-dimensional then (10) can be replaced by the more specific rule, known as Sylvester's law of nullity [9, p. 11] which states that

$$\max[\mathrm{nul}\,(T_1), \mathrm{nul}\,(T_2)] \leq \mathrm{nul}\,(T_1 T_2) \leq \mathrm{nul}\,(T_1) + \mathrm{nul}\,(T_2).$$

We show that the validity of Sylvester's rule for all $T_i \in \mathfrak{F}$ where $\mathfrak{X}$ is infinite-dimensional implies that $\mathfrak{X}$ is not stable-like. For suppose otherwise. Consider

$$T_2 \in \mathfrak{F},\ f(T_2) = n > 0,\ \mathrm{nul}\,(T_2) = 0.$$

Then by [14, Theorem 3.15] there exists $T_1 \in \mathfrak{B}(\mathfrak{X})$ such that $T_1 T_2 = I$. Since I and $T_2 \in \mathfrak{F}$, by [15, Theorem 5.4] we see that $T_1 \in \mathfrak{F}$. By Sylvester's rule, $\mathrm{nul}\,(T_1) = 0$, so that T_1 is regular in $\mathfrak{B}(\mathfrak{X})$ and therefore so is T_2, which is a contradiction.

Another generalization of Schauder's theorem may be obtained as follows. Yosida and Kakutani [16] have considered the collection $\mathfrak{Q}(\mathfrak{X})$ of all *quasi-completely continuous* transformations in $\mathfrak{B}(\mathfrak{X})$ i.e. the class of all $T \in \mathfrak{B}(\mathfrak{X})$ such that there exists $V \in \mathfrak{K}(\mathfrak{X})$ and an integer n such that $\|T^n - V\| < 1$.

5.7. THEOREM. *Let $T \in \mathfrak{F}$, and let V be a two-sided inverse of T modulo $\mathfrak{K}(\mathfrak{X})$. Suppose that there exists $W \in \pi^{-1}(\mathfrak{P}_1)$ and an integer m such that $V^m U - W \in \mathfrak{Q}(\mathfrak{X})$. Then $T^m + U \in \mathfrak{F}$, and*

$$f(T^m + U) = mf(T).$$

Proof. Let $V^m U = R_1$ and $R_1 - W = R_2$. By hypothesis there is an integer n such that $I - R_2^n$ is of the form $S_1 + S_2$, where $S_1 \in \mathfrak{G}$ and $S_2 \in \mathfrak{K}(\mathfrak{X})$. Since $\pi^{-1}(\mathfrak{P}_1)$ is a two-sided ideal, there exists $S_3 \in \pi^{-1}(\mathfrak{P}_1)$ such that

$$I - R_1^n = S_1 + S_3.$$

But, by Lemma 2.5, $S_1 + S_3 \in \mathfrak{F}$. Therefore $I - R_1^n$ has a two-sided inverse modulo $\mathfrak{K}(\mathfrak{X})$. Since

$$I - R_1^n = (I - R_1)(I + R_1 + \cdots + R_1^{n-1}) = (I + R_1 + \cdots + R_1^{n-1})(I - R_1)$$

then $I - R_1 \in \mathfrak{F}$. Since the hypothesis on U is satisfied by all αU, $|\alpha| \leq 1$, it follows from Theorem 3.4 that

$$f(I - R_1) = f(I + R_1) = 0.$$

Applying Corollary 5.5, we obtain

$$f(T^m + U) = f[T^m(I + R_1)] = mf(T).$$

6. On the images of left and right regular elements. We make here a detailed study of the images of the sets $\mathfrak{G}$, $\mathfrak{G}^l$, and $\mathfrak{G}^r$ under π. In view of Lemma 2.1, the results also hold for the mapping τ. In particular, we show the following:

6.1. THEOREM. *The canonical homomorphism π has the following properties*:

(1) $\pi(\mathfrak{G}^l \cup \mathfrak{G}^r) = \pi(\mathfrak{G}^l) \cup \pi(\mathfrak{G}^r) = \mathfrak{G}_1^l \cup \mathfrak{G}_1^r$;

(2) $\pi(\mathfrak{G}) = \pi(\mathfrak{G}^l) \cap \pi(\mathfrak{G}^r)$;

(3) *the sets $\pi(\mathfrak{G})$, $\pi(\mathfrak{G}^l)$, and $\pi(\mathfrak{G}^r)$ are open and closed in the sets $\mathfrak{G}_1$, $\mathfrak{G}_1^l$, and $\mathfrak{G}_1^r$, respectively*;

(4) *$\pi(\mathfrak{G})$ is a normal subgroup of $\mathfrak{G}_1$; either $\pi(\mathfrak{G}) = \mathfrak{G}_1$ or $\mathfrak{G}_1/\pi(\mathfrak{G})$ is isomorphic, as a topological group, to the additive group of integers in the discrete topology.*

The interest of (1) lies in the fact that if $\mathfrak{X}$ is stable-like, then $\pi(\mathfrak{G}^l) \neq \mathfrak{G}_1^l$ and $\pi(\mathfrak{G}^r) \neq \mathfrak{G}_1^r$ (see Lemma 6.3). And for (2), even though $\mathfrak{G} = \mathfrak{G}^l \cap \mathfrak{G}^r$ this does not of itself imply that

$$\pi(\mathfrak{G}^l \cap \mathfrak{G}^r) = \pi(\mathfrak{G}^l) \cap \pi(\mathfrak{G}^r).$$

In the course of the proof the following notation is used. $\mathfrak{H}_0$ is the subset of $\mathfrak{H}$ consisting of those T for which $f(T) = 0$ and $\mathfrak{H}_+$ ($\mathfrak{H}_-$) of those T for which $f(T) > 0$ ($f(T) < 0$). The minus sign for sets in $\mathfrak{E}(\mathfrak{X}) - \mathfrak{R}(\mathfrak{X})$ is used in the set-theoretic sense. From the definitions we have $\pi(\mathfrak{H}) = \mathfrak{G}_1$.

The following lemmas are part of the proof of Theorem 6.1.

6.2. LEMMA. $\pi(\mathfrak{G}) = \{T_1 \in \mathfrak{E}(\mathfrak{X}) - \mathfrak{R}(\mathfrak{X}) \mid \pi^{-1}(T_1) \subset \mathfrak{H}_0\}$, and $\pi(\mathfrak{G}) = \pi(\mathfrak{H}_0)$.

Proof. The second statement follows immediately from the first. Suppose that $T_1 = \pi(T)$, $T \in \mathfrak{G}$. Then $\pi^{-1}(T_1) = T + \mathfrak{R}(\mathfrak{X})$, so that for each $U \in \pi^{-1}(T_1)$, $f(U) = f(T)$ by Lemma 2.5. Since $f(T) = 0$, we see that $\pi(\mathfrak{G})$ is contained in the right-hand set. Next assume that T_1 is in the right-hand set. Let $\pi(T) = T_1$. Then $T \in \mathfrak{H}_0$, and $f(T) = 0$. By Lemma 2.6 there exists $V \in \mathfrak{R}(\mathfrak{X})$ such that $T + V \in \mathfrak{G}$. But $\pi(T + V) = T_1$.

6.3. LEMMA. $\pi(\mathfrak{G}^l) = \mathfrak{G}_1^l - \pi(\mathfrak{H}_-)$.

Proof. Clearly $\pi(\mathfrak{G}^l) \subset \mathfrak{G}_1^l$. We shall show that $\pi(\mathfrak{G}^l) \cap \pi(\mathfrak{H}_-)$ is empty. Suppose contrariwise that $T_1 \in \pi(\mathfrak{G}^l) \cap \pi(\mathfrak{H}_-)$. Then there exists $T \in \mathfrak{G}^l$, $U \in \mathfrak{H}_-$ such that $\pi(T) = \pi(U) = T_1$. Then there exists $W \in \mathfrak{K}(\mathfrak{X})$ such that $T = U + W$. Hence, by Lemma 2.5, $f(T) = f(U) < 0$. But from the definition of f, nul $(T) > 0$. Therefore T cannot be one-to-one and this contradicts $T \in \mathfrak{G}^l$. We conclude that $\pi(\mathfrak{G}^l) \subset \mathfrak{G}_1^l - \pi(\mathfrak{H}_-)$.

Suppose that $T_1 \in \mathfrak{G}_1^l - \pi(\mathfrak{H}_-)$ and $\pi(T) = T_1$. By [15, Theorem 5.4], T has property A. Since $T \notin \mathfrak{H}_-$, either nul (T^*) is not finite or nul $(T^*) < \infty$ and $f(T) \geq 0$. Then by [15, Theorem 3.13] there exists $V \in \mathfrak{K}(\mathfrak{X})$ such that $T + V$ is a bi-continuous mapping of $\mathfrak{X}$ into $\mathfrak{X}$. Moreover, by [15, Theorems 5.3 and 5.4], there exists a projection of $\mathfrak{X}$ onto $(T + V)(\mathfrak{X})$. Therefore, by [14, Theorem 3.15], $T + V \in \mathfrak{G}^l$. However, $\pi(T + V) = \pi(T) = T_1$. Thus $\mathfrak{G}^l - \pi(\mathfrak{H}_-) \subset \pi(\mathfrak{G}^l)$.

6.4. LEMMA. $\pi(\mathfrak{G}^r) = \mathfrak{G}_1^r - \pi(\mathfrak{H}_+)$.

In references cited in the proof of Lemma 6.3, dual results exist to those used in 6.3 which enable one to conduct the proof in the same way.

6.5. LEMMA. $\pi(\mathfrak{H}_-) \subset \pi(\mathfrak{G}^r)$ *and* $\pi(\mathfrak{H}_+) \subset \pi(\mathfrak{G}^l)$.

Proof. Suppose that $T \in \mathfrak{H}_-$. By [15, Theorem 3.13] there exists $V \in \mathfrak{K}(\mathfrak{X})$ such that $(T + V)(\mathfrak{X}) = \mathfrak{X}$. Also, by Lemma 2.4, nul $(T + V) < \infty$. Hence [14, Theorem 3.18] shows that $T + V \in \mathfrak{G}^r$. However, $\pi(T + V) = \pi(T)$. The other statement is proved using dual results.

6.6. LEMMA. $\mathfrak{H}_0$, $\mathfrak{H}_+$, $\mathfrak{H}_-$ *are open and closed as subsets of* $\mathfrak{H}$. *These sets are disjoint.*

Proof. Since $f(T)$, by Lemma 2.5, is a continuous integral-valued function on $\mathfrak{H}$, the sets are open and closed subsets of $\mathfrak{H}$.

We turn now to the statements of Theorem 6.1.

Consider (1). By Lemmas 6.3 and 6.4,

$$\pi(\mathfrak{G}^l \cup \mathfrak{G}^r) = \pi(\mathfrak{G}^l) \cup \pi(\mathfrak{G}^r) = [\mathfrak{G}_1^l - \pi(\mathfrak{H}_-)] \cup [\mathfrak{G}_1^r - \pi(\mathfrak{H}_+)].$$

By Lemma 6.5,

$$\pi(\mathfrak{H}_-) \subset \pi(\mathfrak{G}^r), \ \pi(\mathfrak{H}_+) \subset \pi(\mathfrak{G}^l),$$

so that

$$\pi(\mathfrak{G}^l) \cup \pi(\mathfrak{G}^r) = \mathfrak{G}_1^l \cup \mathfrak{G}_1^r.$$

As for (2), note first that $\pi(\mathfrak{G}) = \pi(\mathfrak{H}_0)$ by Lemma 6.2. By Lemmas 6.3 and 6.4,

$$\pi(\mathfrak{G}^l) \cap \pi(\mathfrak{G}^r) = \mathfrak{G}_1^l \cap \mathfrak{G}_1^r - [\pi(\mathfrak{H}_-) \cup \pi(\mathfrak{H}_+)].$$

But $\mathfrak{G}_1^l \cap \mathfrak{G}_1^r = \mathfrak{G}_1 = \pi(\mathfrak{H})$. Also the sets $\pi(\mathfrak{H}_+)$, $\pi(\mathfrak{H}_-)$ and $\pi(\mathfrak{H}_0)$ are disjoint since if, for example, $T_1 \in \pi(\mathfrak{H}_+) \cap \pi(\mathfrak{H}_-)$, $T_1 = \pi(T)$, $T \in \mathfrak{H}_+$ and $T_1 = \pi(V)$, $V \in \mathfrak{H}_-$, then $\pi(T - V) = 0$ so that $T - V \in \mathfrak{R}(\mathfrak{X})$; whence, by Lemma 2.4, $f(T) = f(V)$ which is impossible. Hence

$$\pi(\mathfrak{G}^l) \cap \pi(\mathfrak{G}^r) = \pi(\mathfrak{H}) - [\pi(\mathfrak{H}_-) \cup \pi(\mathfrak{H}_+)] = \pi(\mathfrak{H}_0) = \pi(\mathfrak{G}).$$

The mapping π is a continuous linear mapping of the Banach algebra $\mathfrak{E}(\mathfrak{X})$ onto the Banach algebra $\mathfrak{E}(\mathfrak{X}) - \mathfrak{R}(\mathfrak{X})$. Consequently it takes open sets into open sets. Since $\mathfrak{G}$, $\mathfrak{G}^l$, $\mathfrak{G}^r$, $\mathfrak{G}_1$, $\mathfrak{G}_1^l$ and $\mathfrak{G}_1^r$ are open (see, for example, [**12**]) the statement of (3) on openness follows. Likewise, from Lemma 6.6, $\pi(\mathfrak{H}_-)$ is open in $\mathfrak{G}_1 \subset \mathfrak{G}_1^l$. Since

$$\pi(\mathfrak{G}^l) = \mathfrak{G}_1^l - \pi(\mathfrak{H}_-)$$

by Lemma 6.3, $\pi(\mathfrak{G}^l)$ is closed in $\mathfrak{G}_1^l$. Similarly $\pi(\mathfrak{G}^r)$ is open and closed in $\mathfrak{G}_1^r$. Now

$$\mathfrak{G}_1 = \pi(\mathfrak{H}) = \pi(\mathfrak{H}_0) \cup \pi(\mathfrak{H}_-) \cup \pi(\mathfrak{H}_+)$$

and (as noted above) the latter sets are disjoint and also open by Lemma 6.6. But $\pi(\mathfrak{G}) = \pi(\mathfrak{H}_0)$ by Lemma 6.2. Thus $\pi(\mathfrak{G})$ is open and closed in $\mathfrak{G}_1$ and the proof of (3) is complete.

Only (4) remains to be shown. Either $\pi(\mathfrak{G}) = \mathfrak{G}_1$ or $\pi(\mathfrak{G})$ is properly contained in $\mathfrak{G}_1$. Suppose that the latter holds. By Lemma 6.2, $\pi(\mathfrak{H}_0) = \pi(\mathfrak{G})$. But $\pi(\mathfrak{H}) = \mathfrak{G}_1$. Thus $\mathfrak{H} \neq \mathfrak{H}_0$ and the function f defined on $\mathfrak{H}$ (and on $\pi(\mathfrak{H})$) is not identically zero. Since f is integral valued there is an integer $m > 0$ and $T \in \mathfrak{H}$ such that $|f(T)| = m$ and m is minimal with respect to this property. By Corollary 5.5, f is a homomorphism of $\pi(\mathfrak{H}) = \mathfrak{G}_1$ into the additive group J of integers. If we define f_1 on $\mathfrak{G}_1$ by the rule $f_1 = m^{-1}f$ then f_1 is a homomorphism of $\mathfrak{G}_1$ onto J. The kernel of this homomorphism is $\pi(\mathfrak{H}_0) = \pi(\mathfrak{G})$ (Lemma 6.2). If J is given the discrete topology then f_1 is an open mapping. Since the kernel is open in $\mathfrak{G}_1$ by (3), the inverse image under f_1 of any subset of J is open in $\mathfrak{G}_1$. Hence, [**11**, p. 64], $\mathfrak{G}_1/\pi(\mathfrak{G})$ is isomorphic, as a topological group, to J. This completes the proof of Theorem 6.1.

6.7. THEOREM. *The following statements are equivalent*:

(1) $\mathfrak{X}$ *is not stable-like*; (2) $\mathfrak{H} = \mathfrak{H}_0$; (3) $\pi(\mathfrak{G}) = \mathfrak{G}_1$.

Proof. The equivalence of (1) and (2) is given by Theorem 5.9. In the course of the proof of Theorem 6.1 it was shown that if $\pi(\mathfrak{G}) \neq \mathfrak{G}_1$ then $\mathfrak{H} \neq \mathfrak{H}_0$ so that (2) implies (3). If $\pi(\mathfrak{G}) = \mathfrak{G}_1$ then, by Lemma 6.2, $\pi(\mathfrak{H}_0) = \pi(\mathfrak{H})$. This shows that any element T of $\mathfrak{H}$ differs from an element of $\mathfrak{H}_0$ by a completely continuous transformation in $\mathfrak{E}(\mathfrak{X})$. Therefore, from Lemma 2.4, $\mathfrak{H} = \mathfrak{H}_0$.

6.8. DEFINITION. We say that $\mathfrak{X}$ is *projection-stable* if there exists an isomorphism in $\mathfrak{E}(\mathfrak{X})$ of $\mathfrak{X}$ onto a proper closed linear manifold $\mathfrak{N}$ where there is a (continuous) projection of $\mathfrak{X}$ on $\mathfrak{N}$.

Clearly if $\mathfrak{X}$ is stable-like then $\mathfrak{X}$ is projection-stable. Whether or not the converse is true is an open question. The notion just defined is connected with the notions of Theorem 6.1 by the following result.

6.9. THEOREM. *The following statements are equivalent*:

(1) $\mathfrak{X}$ *is not projection-stable*;

(2) $\mathfrak{G}^l = \mathfrak{G}^r = \mathfrak{G}$;

(3) $\pi(\mathfrak{G}) = \mathfrak{G}_1$ *and* $\mathfrak{G}_1 = \mathfrak{G}_1^l = \mathfrak{G}_1^r$.

Proof. If $\mathfrak{X}$ is not projection-stable then, by [**14**, Theorem 3.15], $\mathfrak{G}^l \subset \mathfrak{G}$ so that $\mathfrak{G}^l = \mathfrak{G}$. But then also $\mathfrak{G} = \mathfrak{G}^r$; for if $T \in \mathfrak{G}^r$, $TU = I$, then $U \in \mathfrak{G}$ and $T = U^{-1} \in \mathfrak{G}$. Thus (1) implies (2). Assume (2). By Theorem 6.1 we see that

$$\pi(\mathfrak{G}) = \pi(\mathfrak{G}^l \cup \mathfrak{G}^r) = \mathfrak{G}_1^l \cup \mathfrak{G}_1^r.$$

But $\pi(\mathfrak{G}) \subset \mathfrak{G}_1$. Hence

$$\mathfrak{G}_1^l = \mathfrak{G}_1^r = \mathfrak{G}_1 \quad \text{and} \quad \pi(\mathfrak{G}) = \mathfrak{G}_1.$$

Assume (3). If $\mathfrak{X}$ were projection-stable then by [**14**, Theorem 3.15] there exists $T \in \mathfrak{G}^l$, $T \notin \mathfrak{G}$. But $\pi(T) \in \mathfrak{G}_1^l = \mathfrak{G}_1$. Hence $T \in \mathfrak{H}$. By its nature $f(T) > 0$. However, from Theorem 6.7, $\mathfrak{H} = \mathfrak{H}_0$, which is a contradiction.

REFERENCES

1. F. V. Atkinson, *The normal solubility of linear equations in normed spaces*, Mat. Sbornik 28 (1951), 3-14. (Russian).

2. S. Banach, *Théorie des opérations linéaires*, Monografje Matematyczna, T.I. Warsaw, 1932.

3. J. W. Calkin, *Two-sided ideals and congruences in the ring of bounded operators in Hilbert space*, Ann. of Math. 42 (1941), 839-873.

4. J. Dieudonné, *Sur les homomorphisms d'espaces normés*, Bull. Sci. Math. (2) 67 (1943), 72-84.

5. N. Dunford and B. J. Pettis, *Linear operations on summable functions*, Trans. Amer. Math. Soc. 47 (1940), 323-392.

6. I. C. Gohberg, *On linear equations in normed spaces*, Doklady Akad. Nauk SSSR (N.S.) 76 (1951), 477-480. (Russian).

7. E. Hille, *Functional analysis and semi-groups*, Amer. Math. Soc. Colloquium Publications, vol. 31, New York, 1948.

8. I. Kaplansky, *Topological rings*, Amer. J. Math. 69 (1947), 153-183.

9. C. C. MacDuffee, *The theory of matrices*, Chelsea Publishing Company, New York, 1946.

10. G. W. Mackey, *On infinite-dimensional linear spaces*, Trans. Amer. Math. Soc. 57 (1945), 155-207.

11. L. Pontrjagin, *Topological groups*, Princeton University Press, 1939.

12. C. E. Rickart, *The singular elements of a Banach algebra*, Duke Math. J. 14 (1947), 1063-1077.

13. J. Schauder, *Über lineare vollstetige Funktionaloperationen*, Studia Math. 2 (1930), 183-196.

14. B. Yood, *Transformations between Banach spaces in the uniform topology*, Ann. of Math. 50 (1949), 486-503.

15. ———, *Properties of linear transformations invariant under addition of a completely continuous transformation*, Duke Math. J. 18 (1951), 599-612.

16. K. Yosida and S. Kakutani, *Operator-theoretic treatment of Markoff's process and mean ergodic theorem*, Ann. of Math., 42 (1941), 188-228.

UNIVERSITY OF OREGON

Pacific Journal of Mathematics

Vol. 4, No. 4 August, 1954